PROCEEDINGS

of the XLI International Scientific Conference
on Information, Communication and Energy Systems and Technologies

2006 ICEST

KING

ICEST 2006 **Proceedings of the XLI International Scientific Conference on Information, Communication and Energy Systems and Technologies**
organized by the Faculty of Communication Technologies,
Technical University of Sofia, 29 June – 1 July 2006, Sofia, Bulgaria

Edited by D. DIMITROV

ISBN-10: 954-9518-37-X
ISBN-13: 978-954-9518-37-5

Typesed by KING Ltd.
Printed in Bulgaria

Dear Colleagues, Dear Friends,

It's our privilege to welcome you to the 41-st International Conference on Information, Communication and Energy Systems and Technologies – ICEST 2006 in Sofia. This is one regular International Conference, organized by The Faculty of Communication Technique and Technologies of Technical University of Sofia, the Technical Faculty of "Saint Kliment Ohridski" University of Bitola and The Faculty of Electronics of the University of Nis. The conference is one important scientific event in the Balkan region and we hope that new Universities from the rest countries in this region will join. We hope that this conference will be one important scientific event in Europe, also.

The Conference program includes 119 papers. More than 220 authors from Bulgaria, Serbia, Macedonia, Greece, Egypt, Czech Republique, Poland, USA, China, Spain, France and Vietnam have contributed for ICEST 2006. The scientific level of the papers is high and they have been accepted for presentation on the base of positive reviewer's reports.

On behalf of the Conference Scientific Committee we would like to thanks to all authors for their contributions. Many thanks to the reviewers for their advices and recommendations. We would like to thanks especially to the Session chairmen and to the members of Conference Organizing committee for their permanent efforts during the preparation of conference. Let's take a chance to thanks to the members of Conference Scientific Committee for the help in the process of organizing of conference.

We hope the conference sessions will enable you to experiment with engineering methods for investigation in the areas of information, communication and energy systems and technologies, presented by participants of the Conference, and to take practical ideas and solutions with you. The Conference will also consolidate the links among the part of European Universities and also create new ones.

We would like to wish you successful presentations.

Be welcome and enjoy your stay in Sofia.

Assoc. Prof. Dr. Emil S. Altimirsky
Dean of the Faculty of Communication
Technique and Technologies

Assoc. Prof. Dr. Dimiter Tz. Dimitrov
Conference Chairman

2006

XLI INTERNATIONAL SCIENTIFIC CONFERENCE ON INFORMATION, COMMUNICATION AND ENERGY SYSTEMS AND TECHNOLOGIES

organized by

Faculty of Communications and Communication Technologies
Technical University of Sofia, Bulgaria

Faculty of Technical Sciences
University „St. Kl. Ohridski", Bitola, Macedonia

Faculty of Electronic Engineering
University of Niš, Serbia and Montenegro

in co-operation with

- ✓ IEEE Bulgaria Section
- ✓ IEEE Macedonia Section
- ✓ IEEE Serbia and Montenegro Section

INTERNATIONAL PROGRAM COMMITTEE

General Chairman:
D. Dimitrov, *Technical University of Sofia, Bulgaria*

Vice Chairmen::

B. Milovanovic	*University of Niš, Serbia and Montenegro*
C. Mitrovski	*University "St.Kliment Ohridski"- Bitola, Macedonia*

Members:

E. Altimirski	*Technical University of Sofia, Bulgaria*
A. Angelov	*Technical University of Sofia, Bulgaria*
D. Antic	*University of Niš, Yugoslavia*
R. Arnaudov	*Technical University of Sofia, Bulgaria*
G. Arsov	*Electrotechnical Faculty, Scopje, Macedonia*
A. Bekiarski	*Technical University of Sofia, Bulgaria*
M. Bogdanov	*Electrotechnical Faculty, Scopje, Macedonia*
B. Bonev	*Technical University of Sofia, Bulgaria*
P. Borovska	*Technical University of Sofia, Bulgaria*
V. Ceselkovska	*Faculty of Technical Sciences-Bitola, Macedonia*
V. Dimcev	*Electrotechnical Faculty Skopje, Macedonia*
B. Dimitrijevic	*University of Niš, Serbia and Montenegro*
R. Dinov	*Technical University of Sofia, Bulgaria*
N. Dodov	*Technical University of Sofia, Bulgaria*
D. Dobrev	*Technical University of Sofia, Bulgaria*
I. Genovski	*Technical University of Sofia, Bulgaria*
M. Gusev,	*Faculty of Natural Sciences and Mathematics, Macedonia*
A. Grigorov	*Technical University of Sofia, Bulgaria*
E. Ferdinandov	*Technical University of Sofia, Bulgaria*
H. Hristov	*Technical University of Sofia, Bulgaria, IEEE CAS Chapter Chair*
M. Hristov	*Technical University of Sofia, Bulgaria*
L. Jordanova	*Technical University of Sofia, Bulgaria*
R. Kountchev	*Technical University of Sofia, Bulgaria*
S. Lishkov	*Technical University of Sofia, Bulgaria*
P. Merdjanov	*Technical University of Sofia, Bulgaria*
I. Milentijevic	*University of Niš, Serbia and Montenegro*
S. Mirtchev	*Technical University of Sofia, Bulgaria*
P. Mirtevski	*Faculty of Technical Sciences- Bitola, Macedonia*
M. Momtchejikov	*Technical University of Sofia, Bulgaria*
B. Pankov	*Technical University of Sofia, Bulgaria*
E. Pentcheva,	*Technical University of Sofia, Bulgaria*
A. Popova	*Technical University of Sofia, Bulgaria*
V. Poulkov	*Technical University of Sofia, Bulgaria*
H.-P. Schade	*Technical University of Ilmenau, Germany*
B. Shishkov	*BAS, Bulgaria*
L. Song	*Harbin University of Science and Technology, China*
D. Sotirov	*Technical University of Sofia, Bulgaria*
M. Stefanovic	*University of Niš, Serbia and Montenegro*
M. Stojcev	*University of Niš, Serbia and Montenegro*
G. Stoyanov	*IEEE Bulgaria Section Chair*
St. Tabakov	*Technical University of Sofia, Bulgaria*
D. Tasic	*University of Niš, Serbia and Montenegro*

ORGANIZING COMMITTEE

Chairman:
D. Dimitrov, *Technical University of Sofia, Bulgaria*

Members:

I. Dochev	*Technical University of Sofia, Bulgaria*
V. Georgieva	*Technical University of Sofia, Bulgaria*
R. Goleva	*Technical University of Sofia, Bulgaria*
G. Iliev	*Technical University of Sofia, Bulgaria*
S. Kolev	*Technical University of Sofia, Bulgaria*
P. Koleva	*Technical University of Sofia, Bulgaria*
L. Lubih	*Technical University of Sofia, Bulgaria*
V. Manevska	*Faculty of Technical Sciences- Bitola, Macedonia*
R. Mironov	*Technical University of Sofia, Bulgaria*
C. Mitrovski	*Faculty of Technical Sciences- Bitola, Macedonia*
T. Mitsev	*Technical University of Sofia, Bulgaria*
D. Najdenov	*Faculty of Technical Sciences- Bitola, Macedonia*
P. Petkov	*Technical University of Sofia, Bulgaria*
A. Popova	*Technical University of Sofia, Bulgaria*
M. Radevska	*Faculty of Technical Sciences- Bitola, Macedonia*
P. Rizov	*Technical University of Sofia, Bulgaria*
Z. Stankovic	*University of Niš, Serbia and Montenegro*
L. Trpezanovski	*Faculty of Technical Sciences-Bitola, Macedonia*
Z. Peric	*University of Niš, Serbia and Montenegro*

CONFERENCE SECRETARIAT

Ludvig Lubih *Technical University of Sofia, Bulgaria*

Veska Georgieva *Technical University of Sofia, Bulgaria*

Technical University of Sofia
Faculty of Communications and Communication Technologies
Kl. Ohridski Blvd. 8, 1000, Sofia, Bulgaria
Phone: (+359 2) 965 3153
Phone: (+359 2) 965 3998
Fax: (+359 2) 965 3095
e-mail: icest@tu-sofia.bg

CONFERENCE INTERNET SITE

http://eedu.tu-sofia.bg/icest/

LIST OF ICEST 2006 REVIEWERS

Assoc. Prof. **Mihail Anchev**, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. **Anton Andonov**, PhD
University of Chemical Technology and Metallurgy, Sofia, Bulgaria

Assoc. Prof. **Anna Andonova**, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. **Roumen Arnaudov** PhD
Technical University of Sofia, Bulgaria

Assist. Prof. **Georgi Balabanov**
Technical University of Sofia, Bulgaria

Assoc. Prof. **Alexander Bekiarski**, PhD
Technical University of Sofia, Bulgaria

Prof. **Eduardo Boemo**, ScD
Univesrity Authonoma di Madrid, Spain

Assoc. Prof. **Jovica Bogdanovic**, PhD
"Sistel" Nis, Serbia & Montenegro

Prof. **Gancho Bojilov**, ScD
Technical University of Sofia, Bulgaria

Assoc. Prof. **Ognian Bumbarov**, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. **Veselin Demirev**, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. **Krasimir Denishev**, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. **Dimiter Dimitrov**, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. **Emil Dimitrov**, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. **Rozalina Dimova**, PhD
Technical University of Varna, Bulgaria

Assoc. Prof. **Todor Djamijkov**, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. **Jelena Djordjević**, PhD
University of Niš, Serbia and Montenegro

Assoc. Prof. **Nikola Dodov**, PhD
Technical University of Sofia, Bulgaria

Assist. Prof. **Ivo Dochev**
Technical University of Sofia, Bulgaria

Vasil Donev, PhD
Technical University of Varna, Bulgaria

Assoc. Prof. **Lila Donevska**, PhD
Technical University of Sofia, Bulgaria

Prof. Dr. **Miglena Dontschewa**
University of Applied sciences, Dorbirn, Austria

Assoc. Prof. **Mario Gachev**, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. **Elisaveta Gadjeva**, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. **Peter Goranov**, PhD
Technical University of Sofia, Bulgaria

Prof. **Nikola Gradinarov**, ScD
Technical University of Sofia, Bulgaria

Prof. **Marin Hristov**, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. **Georgi Iliev**, PhD
Technical University of Sofia, Bulgaria

Assoc. **Prof. Ilia G. Iliev**, PhD
Technical University of Sofia, Bulgaria

Assoc.Prof. **Valja Kukenska**, PhD
Technical University of Gabrovo, Bulgaria

Assoc. Prof. **Mihail Iliev**, PhD
Technical University of Rousse, Bulgaria

Prof. **Racho Ivanov**, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. **Todor Ivanov**, PhD
University "Assen Zlatarov" , Bourgas, Bulgaria

Assist. Prof. **Toni Janevski**, PhD
Technical University of Skopje, Macedonia

Assoc. Prof. **Kiril Koichev**, PhD
Technical University of Gabrovo, Bulgaria

TABLE OF CONTENTS

PART ONE

MICROWAVE TECHNIQUE AND ANTENNAS I

MICROWAVE TECHNIQUE AND ANTENNAS II

RADIO COMMUNICATION SYSTEMS

CABLE COMMUNICATION SYSTEMS

TELECOMMUNICATION SYSTEMS

SIGNAL PROCESSING

PART TWO

IMAGE PROCESSING I

IMAGE PROCESSING II

SOUND PROCESSING

COMPUTER SYSTEMS AND INTERNET TECHNOLOGIES I

COMPUTER SYSTEMS AND INTERNET TECHNOLOGIES II

ELECTRONICS I

ELECTRONICS II

CONTROL SYSTEMS I

CONTROL SYSTEMS II

ENERGY SYSTEMS I

ENERGY SYSTEMS II

MEDICAL COMMUNICATIONS

ENGINEERING EDUCATION

A square microstrip antenna with polarization switching by commutation elements

Milko V. Stefanov[1], Nikola I. Dodov[2]

Abstract – **Reconfigurable square microstrip patch antenna with adjusting its polarization sense is designed. Four rectangular slits with switching elements are made to adjust the polarization by switching one of the two pair elements. Axial Ratio versus frequency and radiation patterns are examined. It's suggested a method of decreasing of the crosspolarization.**

I. INTRODUCTION

Microstrip patch antennas are widely used in wireless applications due to their low-profile structure. They are extremely compatible for embedded antennas in handheld wireless devices such as cellular phones, pagers, GPS receivers etc [1].

Antennas with switching of polarization sense are used for realization of polarization diversity. This is a method to double the channel capacity in one connection because RHCP (Right Hand Circular Polarization) and LHCP (Left Hand Circular Polarization) are orthogonal. The simplest method to switch the polarization is to construct two separate feeding systems for the two polarization senses, but they have to be independent and not to influence each other. The design and fabrication are complicated. Because of that methods with including additional elements on the patch are used. In [1], a four-element polarization-agile microstrip array is presented, where each square patch is loaded with four varactor diodes, one on each edge; microstrip arrays with a phase-shift circuit; using of in-plane biased ferrite substrates are presented in [2]. In this article the switching properties of PIN diodes are used. These properties are determined by the direction of the forward biased dc current with very low level [3]. In switch applications the PIN diode should ideally control the RF signal level without introducing distortion which might change the shape of the RF signal.

Polarization switching can be realized by microstrip patches with different shape. It has to be kept the condition to radiate circularly polarized wave – generation of two orthogonal field components that are in phase quadrature

One common configuration is a square patch with a small difference in length of the two opposite sides (asymmetry). This asymmetry can be also realized by two rectangular slits in the two opposite sides of the square patch. Therefore the orthogonal surface current path lengths are changed, respectively the resonant frequencies of the two modes have a small difference and a circularly polarized wave is radiated. If such slits are made in the other two sides, full symmetry is achieved and the antenna is linearly polarized. To restore the asymmetry a PIN diode in switching mode is placed to the end of each slit. When two opposite PIN diodes are forward biased, the other diodes are reverse biased. The surface current paths to the end of two adjoining sides of the patch are different. Therefore the resonant frequencies of the orthogonal modes are different. This difference depends on the slit size, which is designed to achieve phase shift of 90 degrees between the radiated fields.

When the other couple of PIN diodes are forward biased, circularly polarized wave with the same parameters is radiated, but with the opposite sense of polarization.

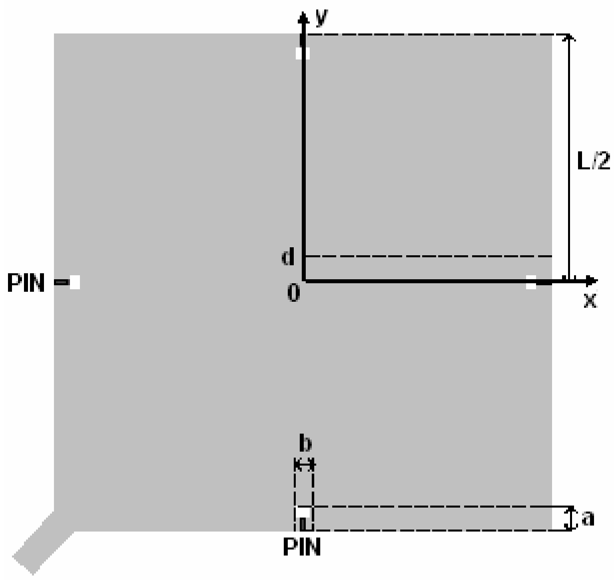


Fig. 1. Circularly polarized microstrip antenna with switching of the polarization sense

II. DESIGN OF SQUARE PATCH WITH POLARIZATION SWITHCING

Dimensions design of a square patch antenna includes calculation of the resonant length, for central working frequency $f = 5GHz$. There are not simple analytical expressions for synthesis and analysis of the circularly polarized microstrip antenna because of the complex character of the field in the antenna. That's why numerical methods for analysis as Moments Method, Finite Difference Method etc. have to be used. In the current analysis a model of a square microstrip antenna with side slits by the medium of a software package, working with Moments Method has been designed. Iteratively the nominal patch size L is determined.

[1] Milko V. Stefanov is with the Faculty of Communications and Communication Technologies, Technical University of Sofia, 8, Kliment Ohridski Str, 1756 Sofia, Bulgaria,
E-mail: mvs@mail.bg

[2] Nikola I. Dodov is with the Faculty of Communications and Communication Technologies, Technical University of Sofia, 8, Kliment Ohridski Str, 1756 Sofia, Bulgaria,
E-mail: ndodov@tu-sofia.bg

The antenna consists of dielectric substrate with $\varepsilon_r = 2.2$ and thickness h=0.5 mm. It is fed by a microstrip line included diagonally to one of the corners of the patch.

Slits dimensions (a and b) are also determined iteratively by contiguous changing of the length and width of each one until the necessary radiation characteristics are achieved.

Polarization switching is easy modeled by the substitution metal bridges for the forward biased PIN diodes, but slits for the reverse biased diodes.

The slits can be made in the center of each side as well as shifted each one to a distance d. That is a method for decreasing the crosspolarization radiation and minimal level of Axial Ratio (AR).

III. RESULTS FROM THE ANALYSIS

On the basis of models of square microstrip antenna with side slits and commutation elements the variations of the axial ratio with the frequency (fig. 2 and 4) and radiation patterns for the necessary polarization and crosspolarization (fig. 3 and 5) are obtained.

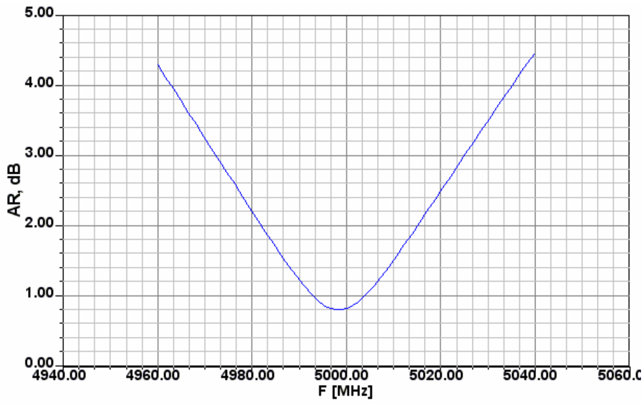


Fig. 2. Axial Ratio versus frequency, d=0, RHCP and LHCP

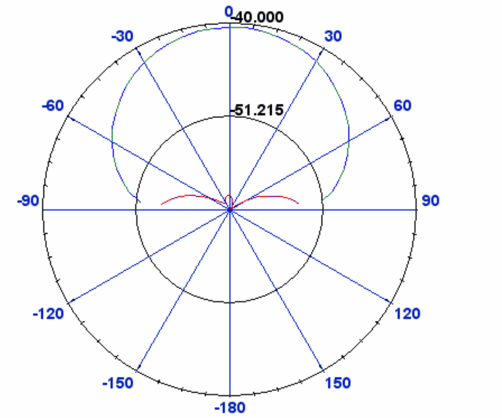


Fig. 3. Radiation pattern, d=0

TABLE I
NUMERICAL RESULTS

L, mm	a, mm	b, mm	d, mm	ARBW, MHz	min AR, dB
19.57	0.4	1	0	52	0.8
19.57	0.4	1	0.2	53	0.2

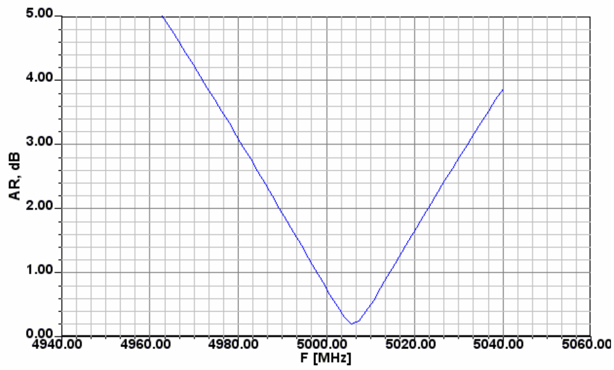


Fig. 4. Axial Ratio versus frequency, d=0.2 mm, RHCP and LHCP

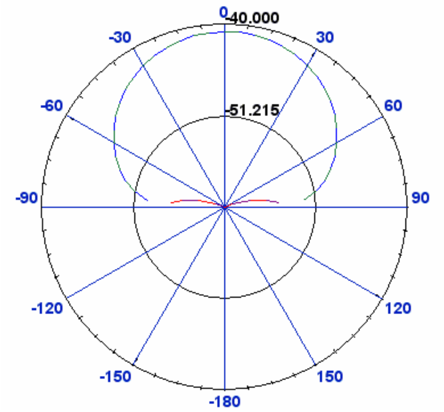


Fig. 5. Radiation pattern, d=.0.2

In the first case the slits are made in the center if each side of the patch (d=0). From Fig. 2 Axial Ratio Bandwidth (ARBW=52 MHz) can be determined. Minimal Axial Ratio is AR=0.8 dB. Fig. 3 shoes that the difference between the polarization and crosspolarization level is 20 dB

In the second case the slits are shifted toward the center to a distance d=0.2 mm. Therefore minimal AR decreases to 0.2 dB, ARBW is little increased, but the crosspolarization difference is increased to 22 dB.

IV. CONCLUSION

In this paper a circularly polarized antenna with polarization switching has been analyzed from the view point of the polarization characteristic – Axial Ratio and respectively Axial Ratio Bandwidth. By the switching of the respective couple of PIN diodes RHCP or LHCP wave is radiated.

Minimal AR and crosspolarization level are decreased by shifting of the slits to a small distance.

REFERENCES

[1] S. Zhong, X. Yang, S. Gao, "Polarization-Agile Microstrip Antenna Array Using a Single Phase-Shift Circuit," *IEEE Trans. Antennas Propag*., vol. 52, pp. 84-87, 2004.

[2] T. Fukusako, A. Imahase, N. Mita, "Polarization characteristics of patch antenna using in-plane and weakly biased ferrite substrate," *IEEE Trans. Antennas Propag*., vol.52, 1, pp.325-327. Jan. 2004

[3] F. Yang, Y. Rahmat-Samii, "Patch Antennas with Switchable Slots (PASS) in Wireless Communications: Concepts, Designs and Applications", *IEEE Trans. Antennas Propagat.*, vol. 47-2, pp. 13-28, 2005.

Compact Wide-band Overlapped Patches Microstrip Antenna

A. A. Shaalan[1], H. M. Abdel-Salam[1], S. H. Zainud-Deen[2], and K. H. Awadalla[2]

Abstract — **This paper presents the design, simulation, and measurements of a novel wideband patch antenna. In this paper, the bandwidth of a single layer microstrip patch antenna is enhanced by using multiple resonances without significantly enlarging the size of the proposed antenna. The validity of the design concept is demonstrated by two examples with 51.4% and 56.8% bandwidths.**

Keywords — **Multi-resonance, Patch antenna, Wide band.**

I. INTRODUCTION

MICROSTRIP patch antennas are widely used because of their many advantages, such as the low profile, light weight, and conformity. However, patch antennas have a main disadvantage i.e. a narrow bandwidth. Researchers have made many efforts to overcome this problem and many configurations have been presented to extend the bandwidth. The conventional method to increase the bandwidth is using parasitic patches. In [1], the authors presented a multiple resonator wide-band microstrip antenna. The parasitic patches are located on the same layer with the main patch. In [2], an aperture-coupled microstrip antenna is described with parasitic patches stacked on the top of the main patch. However, these methods typically enlarge the antenna size, either in the antenna plane or in the antenna height. With the rapid development of wireless communications, single-patch wide-band antennas have attracted many researchers' attention [3]–[5]. In [6], the authors presented a wide-band E-shaped patch antenna and demonstrated that its bandwidth could exceed 30%. Wong [7] presents a good survey of available candidate designs that offer a broadband while maintaining the compactness of microstrip antennas. In [8], the impedance bandwidth of a single-layer microstrip patch antenna is enhanced by using multi-resonance technique. This microstrip antenna employs three square patches that are overlapped along their diagonals, this antenna has five distinct resonance frequencies and is designed to operate from 5.09 to 8.61 GHz. It achieves 51.4 percent bandwidth for return loss < -10 dB. It has been noticed that unfortunately the authors failed to give the coordinates and the specifications of the coaxial probe feed, in this paper a trial and error method was used to find out these missing values.

Another antenna is proposed which is in fact a modification of the overlapped patches microstrip antenna. A slot is incorporated into the complex patch to expand the antenna bandwidth. It achieves 56.8 percent bandwidth for return loss < -10 dB. Simulations will be shown and compared with the measured results.

II. ANTENNAS STRUCTURE, MEASUREMENTS AND SIMULATION RESULTS

A. Overlapped patches microstrip antenna (OPMA)

For a conventional rectangular Microstrip patch antenna of length L and width W, the resonance frequency for any TM_{mn} mode is given by James and Hall [9] to be dependent on the length L, the width W, and the effective dielectric constant of the substrate. But for the dominant TM_{10} mode, the resonance frequency is only dependent on the length L, and the effective dielectric constant. Therefore, it is clear that the resonance frequency of the rectangular microstrip patch antenna is a function of its length (L), so if the microstrip patch antenna has multiple lengths it will be multi-resonance antenna i.e. for every different length there will be a different resonant frequency, hence the bandwidth of the microstrip patch antenna can be enhanced. This technique is utilized in the design of the following microstrip patch antenna.

Three square patches are overlapped along their diagonals to form a non-regular single patch as shown in Fig. 1, The dimensions of the patches are (W_1 X W_1), (W_2 X W_2) and (W_3 X W_3), respectively. S_1 and S_3 indicate the overlapping dimensions of the patches. The structure has five different resonant lengths as follows: L_1, L_2, L_3, L_4 and L_5. As an example, an antenna with the following dimensions was designed: three square patches of dimensions (7.5 X 7.5) mm, (13.5 X 13.5) mm and (7.1 X 7.1) mm with overlapping dimensions S_1=6.4 mm and S_3=4.6 mm, a dielectric substrate of relative permittivity ε_r =2.35 and thickness h =3.175 mm was used [8].

This antenna is fed by a coaxial probe at position (X_f, Y_f) as shown in Fig. 2. The probe feed location and its radius were adjusted in such a way that one can obtain satisfactory performance. Using trial and error, it has been found that at X_f = 4 mm, Y_f = 8 mm, and a probe diameter=1.25mm, the widest bandwidth of this antenna is obtained. The FDTD method full wave simulator *FIDELITY* is used to simulate the overlapped patches microstrip antenna (OPMA) and the obtained results have been compared to other results produced using *IE3D*, a commercial simulator based on the method of moment and good agreements have been found

[1] Department of Electronics and Communications Engineering, Faculty of Engineering, Zagazig University, Egypt

[2] Department of Electronics and Communications Engineering, Faculty of Electronic Engineering, Menoufia University, Egypt

between the two generated results as shown in Fig. 3. For a return loss less than –10 dB the frequency band ranges from 5.09 to 8.61GHz. It achieves 51.4 percent bandwidth for return loss < -10 dB.

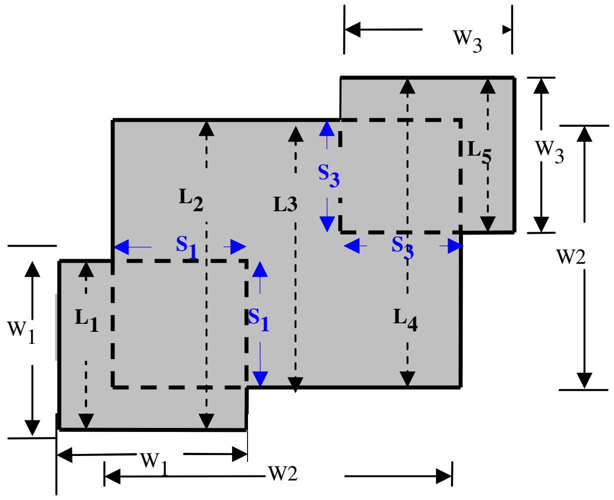


Fig. 1. Geometry of the multi-resonance wideband patch.

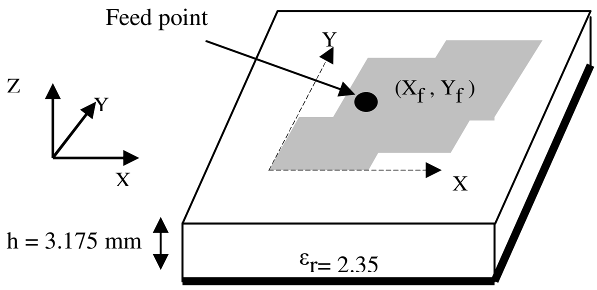


Fig. 2. The OPMA configuration.

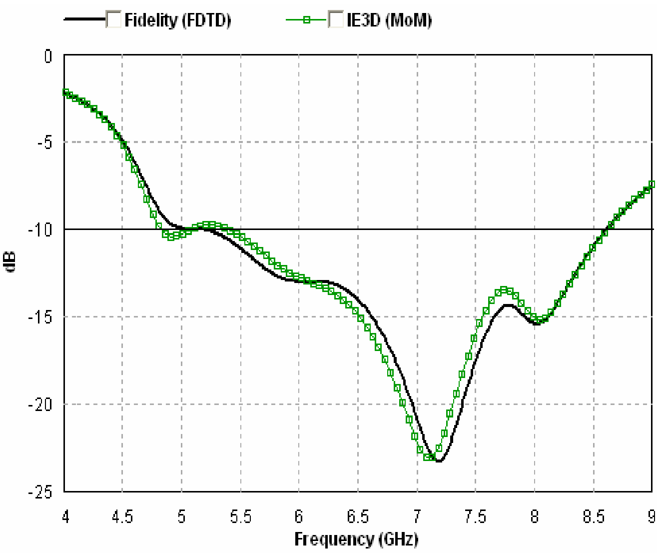


Fig. 3. The return loss of the OPMA.

B. Slotted overlapped patches microstrip antenna (SOPMA)

In the SOPMA, a slot is incorporated into the patch to expand its bandwidth. The OPMA structure described earlier is reused here but modified by inserting a slot. The slot is selected to be 5.1 mm X 0.5 mm and its lower left point is located at (4.625 mm, 5.3 mm). The new SOPMA structure is shown in Fig. 4.

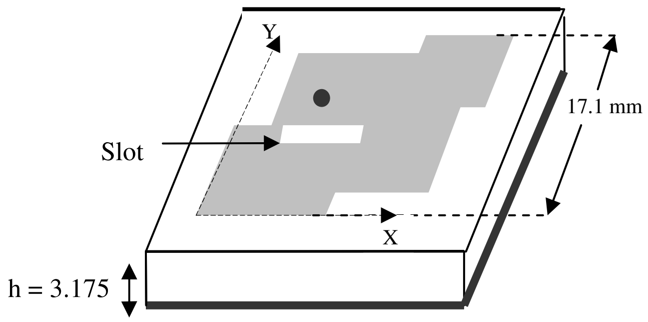


Fig. 4. The SOPMA configuration.

Fig. 5 shows the measured and calculated return loss of the proposed antenna. The SOPMA has 56.8 % bandwidth compared with 51.4 % of the OPMA i.e. wider bandwidth. Also it is clear that the value of S_{11} at resonance is improved by the inserted slot.

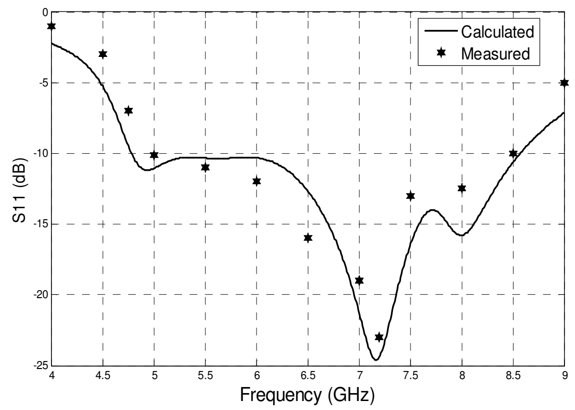


Fig. 5 Measured and calculated return loss of the proposed antenna.

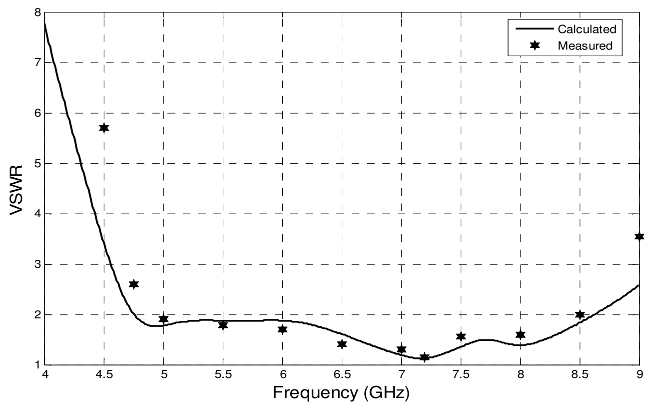


Fig. 6. Measured and calculated VSWR of the proposed antenna

Fig. 6 shows the measured and calculated VSWR of the proposed antenna.The antenna frequency bandwidth with

VSWR<2 covers the fequency range of 4.78-8.57 GHz. This agrees with the less than –10 dB band of the return loss. It has a bandwidth of 56.8 % with the center frequency 6.675 GHz.

The variation of the real part of the input impedance of the SOPMA is shown in Fig. 7. It can be observed that the input resistance is compatible with the 50 ohm characteristics of the input feed line.

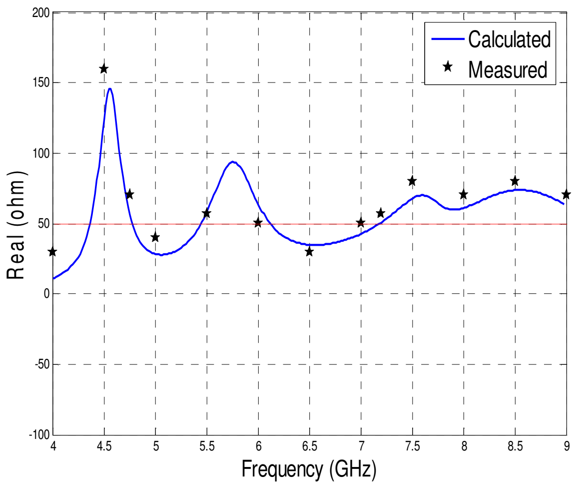


Fig. 7. Measured and calculated real part of the Input Impedance of the proposed antenna

Fig. 8 shows that The SOPMA has five distinct resonance frequencies where the imaginary part of the input impedance equals zero (however no perfect matching is attained). The upper four resonances has VSWR < 2, but the lowest, which is at 4.6 GHz has VSWR close to 3.

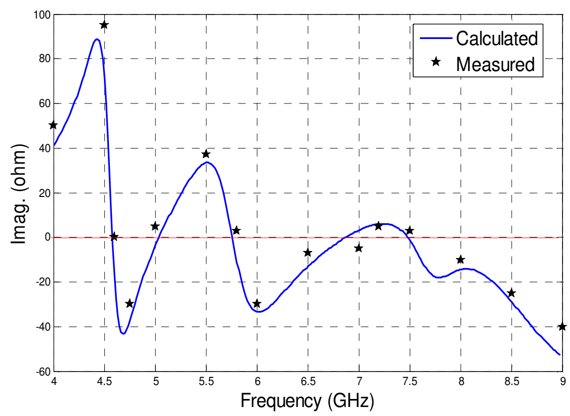


Fig. 8. Measured and calculated imaginary part of the Input Impedance of the proposed antenna.

Since a microstrip patch antenna radiates normal to its patch surface, the elevation pattern for $\phi = 0$ and $\phi = 90$ degrees would be important. Fig. 9 shows the gain Pattern of the OPMA in the XZ-plane ($\phi = 0$ deg.) at different frequencies, it is apparent that this antenna provides stable far field radiation characteristics in the entire operating band with relatively high gain. It is quite clear that the radiation pattern is not symmetrical because of the asymmetry of the patch. It is noticed that at 6.21 GHz the maximum gain is obtained in the broadside direction and this is measured to be 5.63 dBi for both, $\phi = 0$ and $\phi = 90$ degrees. The back-lobe radiation is sufficiently small and is measured to be -14 dBi for the above plot. This low back-lobe radiation is an added advantage for using this antenna in a cellular phone, since it reduces the amount of electromagnetic radiation, which travels towards the users head.

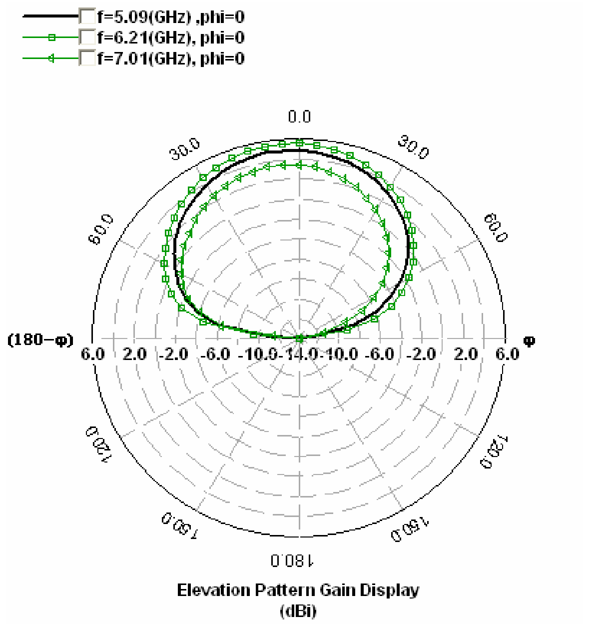


Fig. 9. Radiation Pattern of the SOPMA in the XZ-plane.

Fig. 10 shows the radiation pattern of the SOPMA in the YZ-plane ($\phi = 90$ deg.) at different frequencies, it is Clear that this pattern is also not symmetrical due to the same cause. However, the beamwidth in both planes is wide enough.

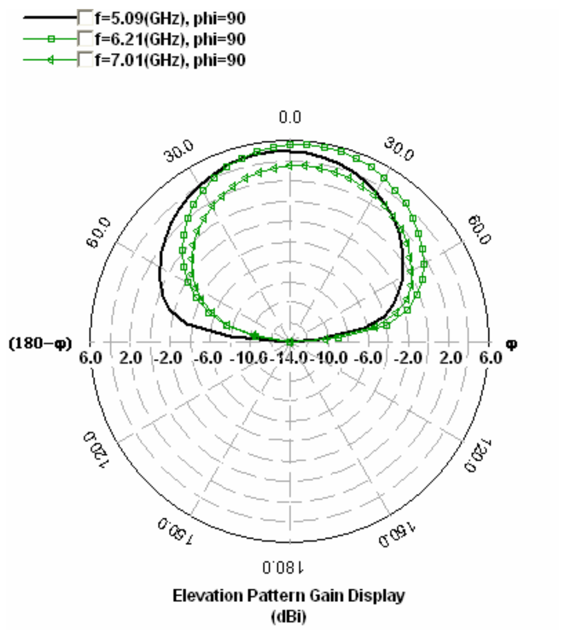


Fig. 10. Radiation Pattern of the SOPMA in the YZ-plane.

Fig. 11 shows that the resonant frequencies of the SOPMA are lower than the resonance frequencies of the OPMA. This is of course, the effect of the slot which adds an inductance to the equivalent circuit of the patch, this added inductance naturally lowers the resonance frequencies as indicated in Fig. 11.

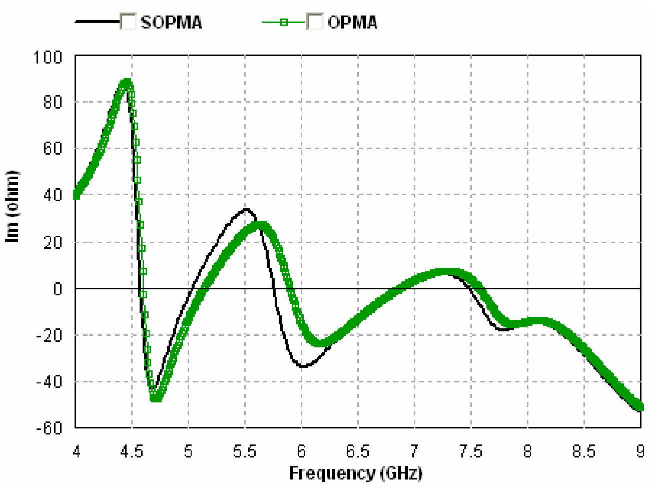


Fig. 11. The imaginary part of the input impedance of both the OPMA and the SOPMA

Fig. 12 illustrates the gain of the OPMA against frequency, the gain is greater than 2 dBi in a frequency range (4.31-7.88 GHz), and the gain variations are less than about 4 dBi across the operating frequency. Due to the fact that the radiating apertures of the two edge patches are relatively smaller compared to those of the main patch, the gain decreases at higher frequencies.

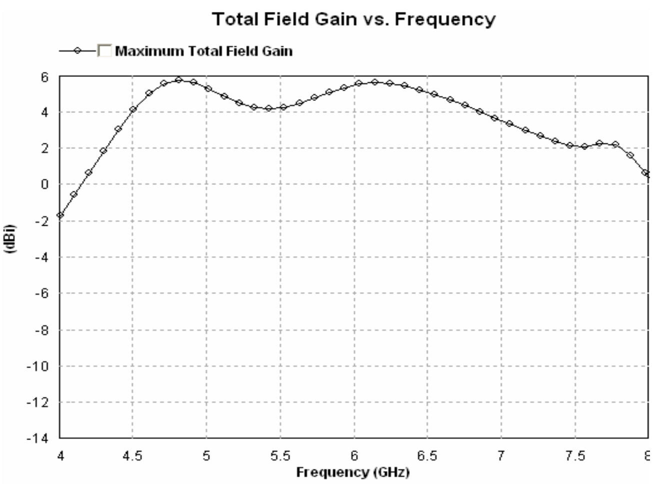


Fig. 12. Gain of the SOPMA

III. Conclusion

In this paper, two designs for small-size wide-bandwidth microstrip patch antennas have been presented. The first design employs three square patches that are overlapped along their diagonals and has been simulated using two commercial field solvers, the obtained bandwidth was 51.4%. In the second design which has been fabricated, a slot is incorporated into the complex patch to expand its bandwidth, it achieves 56.8 percent bandwidth for return loss < –10 dB. Simulations have been shown and compared with the measured results and good agreements have been found. Each structure of these designs can be easily fabricated on a single-layer and relatively thin substrate for applications in hand-held devices. It has been shown that these antennas can easily be used in other frequency bands with different substrate materials.

References

[1] G. Kumar and K. C. Gupta, "Directly coupled multiple resonator wide-band microstrip antenna", *IEEE Trans. Antennas Propagat.*, vol. AP-33, pp. 588–593, June 1985.

[2] S. D. Targonski, R. B. Waterhouse, and D. M. Pozar, "Design of wide-band aperture-stacked patch microstrip antennas", *IEEE Trans. Antennas Propagat.*, vol. 46, pp. 1245–1251, Sept. 1998.

[3] Y. Jang, J. Yoon, H. Shin, "A large bandwidth T-Shaped microstrip-fed ground plane slot antenna", *Microwave Journal*, pp. 92-103, January 2002.

[4] Hassan M. Elkamchouchi, Gehan Abouelseoud, "A compact broadband sierpinski gasket patch microstrip antenna", *21st National Radio Science Conf.*, pp. 1-8, Egypt, March 2004.

[5] A. Shackelford, K. Lee, K. Luk, "Design of small size wide bandwidth microstrip patch antennas", *IEEE Antennas and Propagation Magazine*, vol. 45, pp. 75-82, February 2003.

[6] F. Yang, X.X. Zhang, X. Ye and Y. Rahmat-Samii, "Wide-band E-shaped patch antennas for wireless communications", *IEEE Trans. Antennas Propagat.*, vol.49, pp. 1094-1100, July 2001.

[7] K.Wong, *Compact and Broadband Microstrip Antennas*, John Wiley and Sons, New York, 2002.

[8] K. Rambabu, M. Alam, J. Bornemann and M. A. Stuchly, "Compact wide-band dual-Polarized microstrip patch antenna", *IEEE Antenna and Propagation Society International Symposium,* vol. 2, pp. 1955-1958, June 2004.

[9] J. R. James, and P. S. Hall, *Handbook of Microstrip Antennas*, Peter Peregnnus Ltd., London, United Kingdom, 1989.

New Circularly Polarized Capacitively Probe-Fed Wideband Microstrip Antenna

A. A. Shaalan[1]

Abstract — **This paper presents the design, simulation, and measurements of a novel circularly polarized wideband probe-fed microstrip patch antenna with capacitive feed mechanism. The proposed antenna is designed to achieve three targets; wide bandwidth up to 23%, perfect matching at the input (Zin ≈ 50 ohms), and circular polarization (CP) at resonance. It is designed to operate at 1.8 GHz. This antenna is applicable to Personal Communication System (PCS) which uses the frequency range from 1850-1990 MHz. It can be claimed that this is the first time to realize such microstrip antenna to achieve the three mentioned targets together. Calculated as well as measured results for this antenna are included.**

Keywords — **Patch antennas, Wideband, Polarization.**

I. INTRODUCTION

Many communications systems, such as mobile, satellite, radar, etc., require broadband circularly polarized (CP) antennas. Several microstrip antenna (MSA) configurations are available to generate CP, which can be obtained by either single feed or dual feed. Examples of single feed MSAs are nearly square, corner-chopped square, square or circle with a slot, circular patch with a notch, elliptical, pentagon, modified triangular patches, etc. [1-7]. However, these antennas have a very small axial ratio (AR) bandwidth (BW). A square or circular MSA fed at two orthogonal points with equal amplitude and 90^{o} phase difference yields a better AR BW than the single-feed MSAs. However these antennas have narrow impedance BW.

In this paper, a novel circularly polarized wideband probe-fed microstrip patch antenna with capacitive feed mechanism is proposed, this antenna is designed to achieve wideband performance, perfect matching at the input due to the capacitive feed, and wide axial ratio bandwidth due to the dual feed.

The IE3D package from Zeland Software Incorporated, which is based on the method of moments, was used for the design process and calculated results.

II. ANTENNA GEOMETRY AND OPERATION

The capacitively probe-fed microstrip antenna presented in [8] is now modified to produce circular polarization [9]. This is obtained by properly introducing another small rectangular probe-fed patch, which is also capacitively coupled to the radiating element.The proposed antenna is shown in Fig. 1.

[1]Department of Electronics and Communications Engineering, Faculty of Engineering, Zagazig University, Egypt

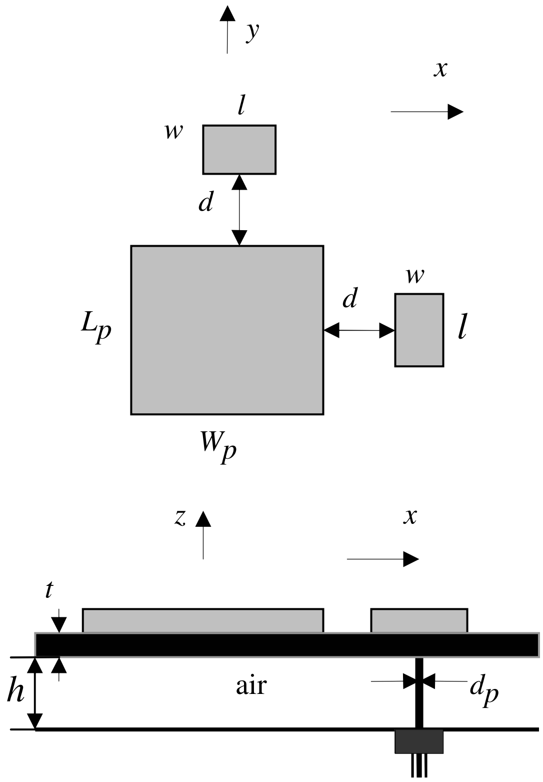


Fig. 1. The proposed circularly polarized wideband microstrip antenna

Each feeding probe is positioned in the center of the two small patches. Both the radiating element and the small patches are supported by a layer of FR-4, with t = 1.6 mm, ε_r = 4.4, and loss tangent tan δ = 0.02, which is suspended in air, at h = 15 mm above a copper ground plane of 100 X 100 mm. Each probe has a diameter of dp = 0.9 mm. The other parameters are Lp= Wp = 51 mm, d = 8 mm, l = 10 mm, w = 5 mm.

The proposed antenna was constructed and studied. Using a Vector Network Analyzer (Agilent 8719ES), which covers the frequency range of 50 MHz up to 13.5 GHz. A photo of the proposed antenna is shown in Fig. 2.

III. EXPERIMENTAL RESULTS

Fig. 3 shows the measured and calculated return loss of the proposed antenna. It is clear that for a return loss (S_{11}) less than –10 dB the frequency band ranges from 1.67 to 2.09 GHz. It has a bandwidth of 22.3 % with the resonance frequency at 1.76 GHz which is very close to 1.8 GHz used in modern wireless communications.

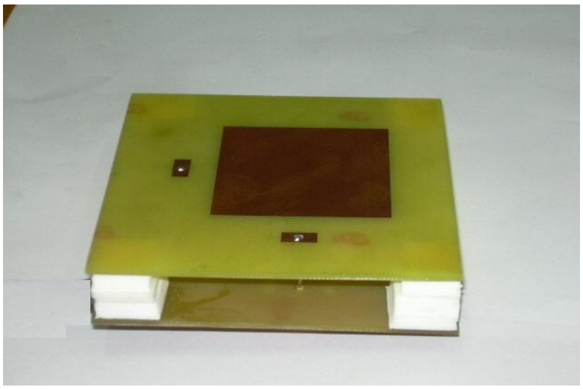


Fig. 2. A photo of the capacitively probe-fed wideband microstrip antenna.

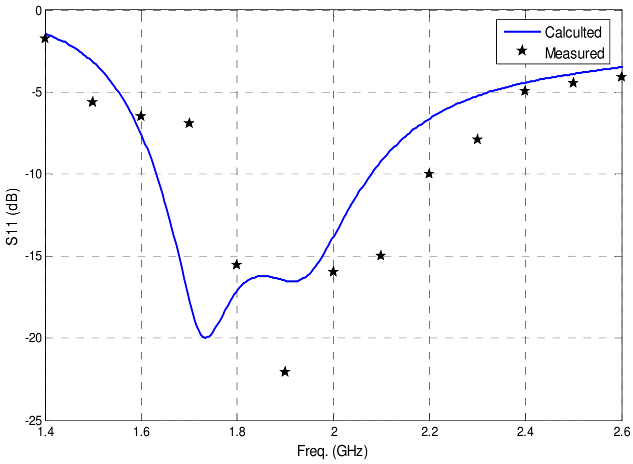


Fig. 3. Measured and calculated return loss of the proposed antenna.

Fig. 4 shows the measured and calculated VSWR of the proposed antenna. The antenna frequency bandwidth with VSWR < 2 covers the frequency range of 1.65 -2.11 GHz. This agrees with the less than –10 dB band of the return loss

The variation of the real part of the input impedance of the proposed antenna is shown in Fig. 5. It is clear that at 1.8 GHz, The real part of the input impedance is very close to 50 ohm, and the imaginary part (Fig. 6) is very close to zero ohms. This is of course, the effect of the capacitive feed mechanism which compensates for the inductive component associated with the probe.

The circular polarization can be verified by looking at the radiation patterns (Fig. 7) of the proposed antenna. When the 2-ports are excited with 90^{o} phase difference, the antenna yields good circularly polarized pattern with cross-polarization below -13 dB from theta = -60^{o} to 60^{o} at 1.78 GHz.

Figure 8 shows that the axial ratio is below 3 dB from 1.5 to 2.02 GHz, the axial ratio has an acceptable value of 2.2 dB at the resonant frequency of 1.8 GHz.

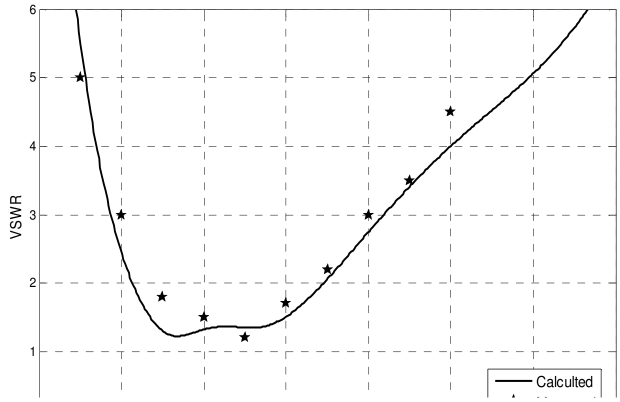


Fig. 4 Measured and calculated VSWR of the proposed antenna.

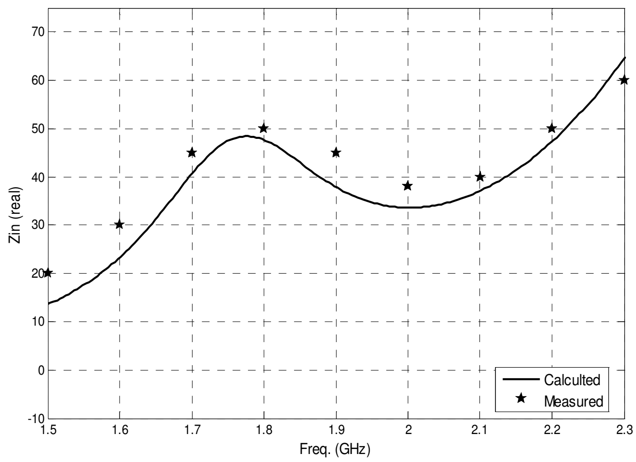


Fig. 5 Measured and calculated real part of the input impedance of the proposed antenna.

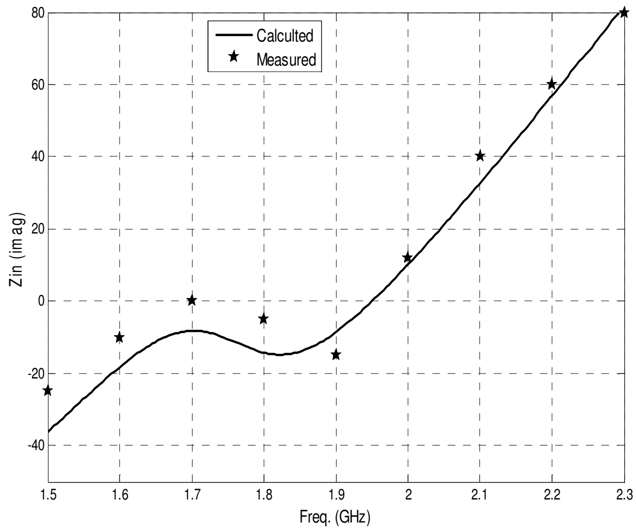


Fig. 6 Measured and calculated imaginary part of the input impedance of the proposed antenna.

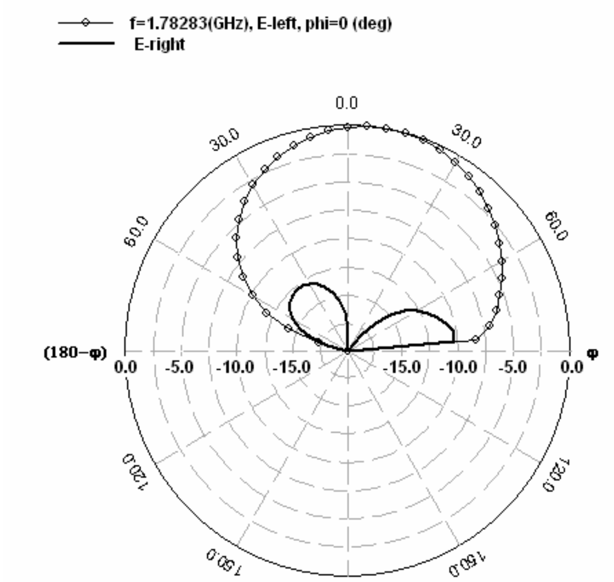


Fig. 7 The radiation pattern of the proposed circularly polarized capacitively probe-fed wideband microstrip antenna.

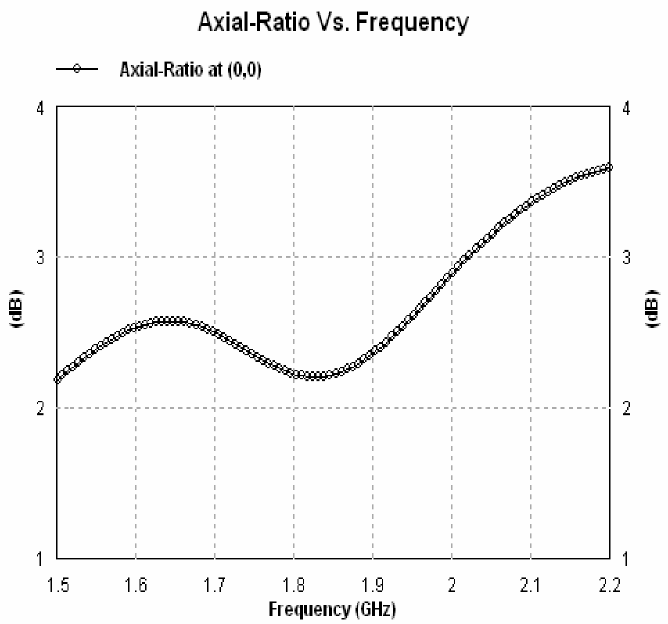


Fig. 8 The axial ratio of the proposed circularly polarized capacitively probe-fed wideband microstrip antenna.

The effective bandwidth (where both the S_{11} is less than -10 dB AND the axial ratio is less than 3 dB) is from 1.67 to 2.02 GHz. The effective bandwidth equals 18.9 %.

Figure 9 shows that the directivity of the proposed circularly polarized capacitively probe-fed wideband microstrip antenna has an acceptable value of 7.7 dBi at the resonant frequency of 1.8 GHz. Figure 10 shows that the radiating efficiency of the proposed antenna has a good value of 93 % at the resonant frequency of 1.8 GHz.

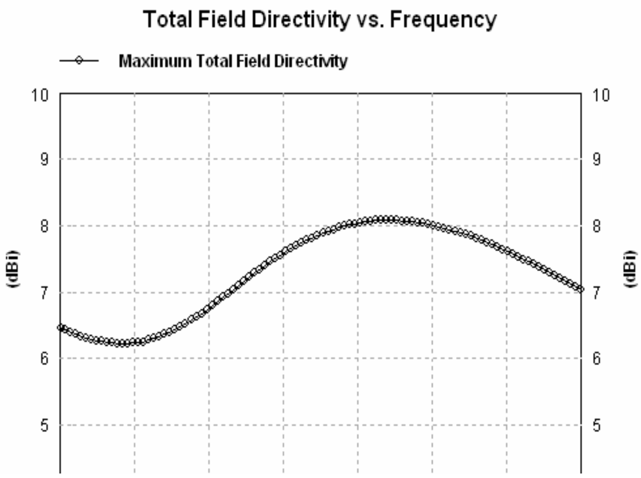


Fig. 9 The directivity of the proposed circularly polarized capacitively probe-fed wideband microstrip antenna.

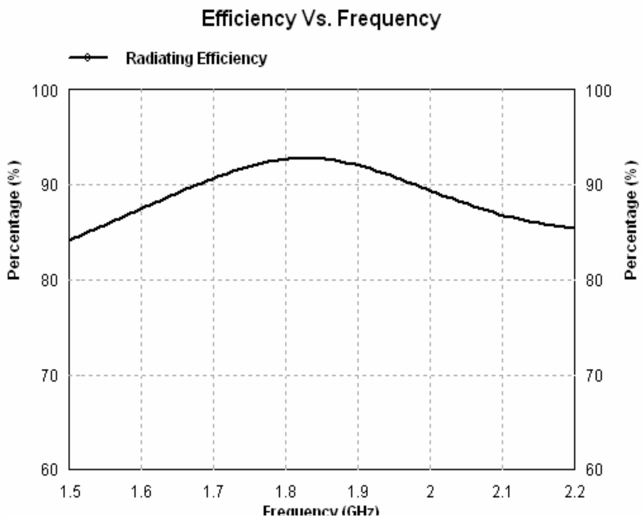


Fig. 10 The radiating efficiency of the proposed circularly polarized capacitively probe-fed wideband microstrip antenna.

IV. CONCLUSION

In this paper, a new circularly polarized capacitively probe-fed wideband microstrip antenna is presented. This antenna consists of two small probe-fed rectangular patches, which are capacitively coupled to the radiating element. The proposed antenna is designed to achieve three targets; wide bandwidth up to 23 %, perfect matching at the input (Zin ≈ 50 ohms), and circular polarization at the resonance. It can be claimed that this is the first time to realize such microstrip antenna to achieve the three mentioned targets together.

REFERENCES

[1] J. R. James and P. S. Hall, *Handbook of microstrip antennas*, Peter Pereginus Ltd., London, England, Vol. 1, 1989.
[2] P. c. Sharma, K. C. Gupta, Analysis and optimized design of sigle feed circularly polarized microstrip antennas, *IEEE Trans Antennas Propagat.*, *Ap-31*, 1983, pp. 949-955
[3] H. Iwaski, T. Nakajima, and Y. Suzuki, Gain improvement of circularly polarized array antenna using linearly polarized elements, *IEEE Trans Antennas Propagat.*, 1995, pp. 604-608
[4] A. Bhattacharya and L. Shafai, A wider band microstrip antenna for circular polarization, *IEEE Trans Antennas Propagat.,* 1998, pp. 157-163
[5] C. H. Tang, J. H. Lu, and K. H. Wong, Circularly polarized equilateral-triangular microstrip antenna with truncated tip, *Electron Lett.* 13, 1998, pp. 1277-1278.
[6] W. H. Hsu, and K. L. Wong, Circularly polarized disk sector microstrip antenna, *Electron Lett. 34*, 1998, pp. 2188-2190.
[7] K. Wong, *Compact and broadband microstrip antennas,* John Wiley and Sons, New York, 2002.
[8] Gordon Mayhew-Ridgers, John W. Odendaal, and John Joubert, Single- layer capacitive feed for wideband probe-fed microstrip antenna elements, *IEEE Transactions on Antennas and Propagat*., Vol. 51, no. 6, June 2003, pp. 1405-1407.
[9] Hussein Mahmoud Abd El-Salam, Wideband microstrip antennas, M. Sc. Thesis. Dept. of Electronics & Communications Eng., Zagazig University, March 2006.

Evaluation of the Influence of Substrate Electrical Parameters over Matching of Microstrip Resonator

Iva N. Stefanova[1] and Nikola I. Dodov[2]

Abstract – **Impedance matching of rectangular transmission line - fed microstrip patch is investigated in this paper. For evaluation of impedance matching the reflection coefficient (S_{11}) is used. On the basis of the variation of S_{11} the impedance bandwidth of the antenna can be defined. The varying of S_{11} is determined by the input impedance R_{in} of the antenna.**

Keywords – **microstrip antennas, antenna feeds, transmission lines.**

I. Introduction

Microstrip patch antennas are the most dynamic developing section of the antenna field in recent years. In communication systems the low weight, simple manufacturing, easy installation, aerodynamic profile and naturally low cost are important requirements to antennas. To a considerable degree the microstrip patch antennas meet those needs. They are inexpensive to manufacture, low-profile, conformable to planar and nonplanar surfaces, easy to produce using the widespread printed-circuit technology, mechanically robust. Major disadvantage of the microstrip antennas are their not very good electrical characteristics – low efficiency, poor polarization purity, low power, spurious feed radiation and narrow frequency bandwidth [1]. There are methods, such as increasing the height of the substrate, which can be used to increase the efficiency and the bandwidth [2].

The microstrip antenna consist of a very thin metallic strip (patch) ($t << \lambda_0$, where λ_0 is the free space wavelength), placed a small fraction of wavelength above a ground plane ($h << \lambda_0$). They are separated by a dielectric layer called substrate. The microstrip patch is designed so that its pattern maximum is perpendicular to the patch. This is achieved by proper choosing of the field configuration beneath the patch.

There are numerous dielectric materials that can be used for substrates in microstrip antennas and their dielectric constants are in the range of $2 \le \varepsilon_r \le 10$. Most desirable for antennas are thick substrates with dielectric constant in the lower part of the range because they provide higher efficiency, wider bandwidth, but also larger element size [3].

Thin substrates with higher dielectric constants are used in microwave circuitry where tight bounding of the fields is desired in order to minimize the undesired radiation and the element size is smaller. In this case there are greater losses, they are less efficient and have relatively smaller bandwidths [4]. Since microstrip antennas are often integrated with other microwave circuits, a compromise has to be made between good antenna and circuit performance.

The microstrip patch can be rectangular, square, dipole, circular and with other shape [1]. The most often used shapes are rectangular and circular, because they are easy to manufacture and analyze, and have good radiation characteristics.

The most common feed configurations for microstrip line are: with microstrip line, with coaxial probe, aperture coupling and proximity coupling. The microstrip line feed is the first type of feed used for microstrip patches and arrays [5]. It is a conducting strip (of much smaller width than that of the patch), placed above the substrate like the patch. The microstrip line is connected to the patch in a position inside his borders. This method is simple to fabricate, easy to match by controlling the inset position, and simple to model. The coaxial-line feed is also very common, it is also easy to do and match, and has low spurious radiation. The noncontacting feed methods use electromagnetic coupling to send energy between the patch and the feed line. The most common noncontacting feed techniques are the aperture coupling and the proximity coupling. In the both there are two substrates which allow independent optimization of the antenna and feed mechanisms, but make the fabrication difficult and the antenna price higher. Taking all this into consideration a survey of microstrip line feed is made.

Impedance matching of rectangular transmission line - fed microstrip patch is investigated in this paper. For evaluation of impedance matching the reflection coefficient (S_{11}) is used. On the ground of the variation of S_{11} the impedance bandwidth of the antenna can be defined. The varying of S_{11} is determined by the input impedance R_{in} of the antenna.

II. Evaluation of the Influence of Substrate Electrical Parameters

The purpose of the present investigation is to determine the influence of the substrate parameters over the microstrip antenna matching. The effect over the bandwidth has been researched.

Two are the basic substrate parameters that have effect. These are the substrate thickness h and its dielectric constant ε_r. They have great influence over the antenna performance, and the choice of dielectric for a substrate is a very important part of the antenna design.

[1] Iva N. Stefanova is with the Faculty of Communications and Communication Technologies, Technical University, Kliment Ohridski 8, 1000 Sofia, Bulgaria, e-mail: ivastefanova@mail.bg

[2] Nikola I. Dodov is with the Faculty of Communications and Communication Technologies, Technical University, Kliment Ohridski 8, 1000 Sofia, Bulgaria, e-mail: ndodov@tu-sofia.bg

The structure considered in this analysis is a rectangular microstrip patch, fed by a microstrip line which has a characteristic impedance of $Z_c = 100\Omega$ (Fig. 1). The dimensions of the rectangular patch are calculated for work at frequency of 10GHz. The analysis is done by means of electromagnetic simulator. Innitially a case where the resonator is calculated with a standart dielectric – Duroid5880 (with thickness $h = 1{,}588$ *mm,* and dielectric constant $\varepsilon_r = 2{,}2$) is analysed. After that a parametric study is made by changing the values of the thickness h and the dielectric constant ε_r only and keeping all the other values and dimensions constant.

The analysis made so far is incorrect from the point of view of the proper patch calculation for work at 10 GHz, which is also confirmed by the obtained results. It can be seen that the resonant action of the rectangular patch is shifted away from the presupposed frequency as expected. A linear shift of the resonance towards lower frequencies is observed with the increase of both the thickness h and the dielectric constant ε_r.

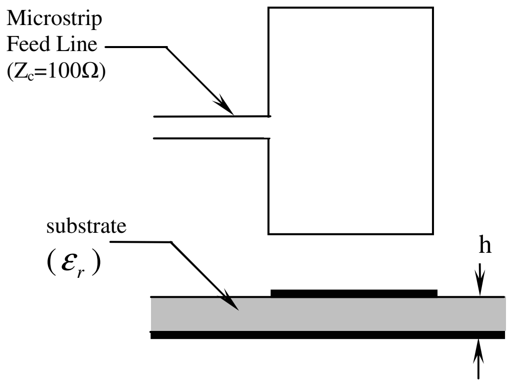


Fig. 1

With the aim of higher accuracy of the parametric study in the second part the case where recalculating the patch dimensions for every single value of the thickness h and the dielectric constant ε_r is analyzed. It can be expected that in this way the antenna resonance will remain about 10 GHz.

III. Results

The input impedance R_{in} and the reflection coefficient S_{11} are shown on Figs. 2 - 5 for the parametric study where only the values of the thickness h and the dielectric constant ε_r are changed. All the other values and sizes are kept unchanged. The same parameters for the second case where the dimensions of the patch are recalculated for every change of the thickness h and the dielectric constant ε_r, are shown on Figs. 6 – 9.

When ε_r is changed in the limits of 2-2,5 (Fig. 6) this has a small effect over the input impedance R_{in}. When ε_r is increased, R_{in} decreases. This can be explained with the fact that at higher values of ε_r there is less radiation, and smaller values of the radiation conductance G_{rad}. The maximum value of R_{in} is getting smaller with higher values of ε_r because of the smaller impedance and the smaller losses.

There is certain shifting of the resonance towards lower frequencies with increase of ε_r as well as similar shifting of the maximum of the input impedance R_{in}. In all of the cases the maximum of R_{in} is not at the resonant frequency for which the resonator has been calculated.

The considerations made so far can be viewed also at S_{11} (Fig. 7). This can be expected because of the direct dependency of S_{11} and R_{in} [1]:

$$\left|S_{11}\right| = \left|\frac{R_{in} - Z_c}{R_{in} + Z_c}\right|. \tag{1}$$

Since the microstrip feed line has a characteristic impedance of $Z_c = 100\ \Omega$, the minimum of the reflection coefficient S_{11} can be seen for frequencies for which $Z_c \neq R_{in}$. For these frequencies for which R_{in} has maximum value, it is different from $Z_c = 100\ \Omega$, and the matching is worse.

With increase of ε_r lower values of S_{11} can be seen.

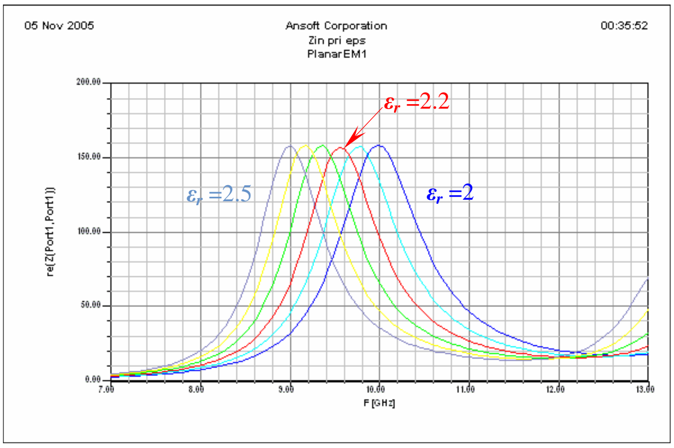


Fig. 2. Rin(f), ε_r = (2, 2.1, 2.2, 2.3, 2.4, 2.5)

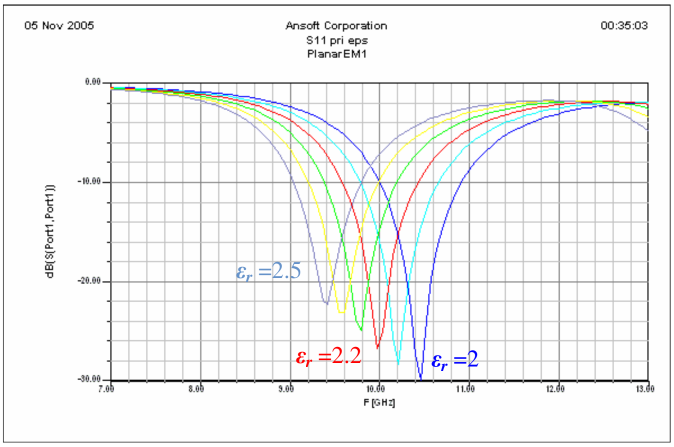


Fig. 3, S_{11}(f), ε_r = (2, 2.1, 2.2, 2.3, 2.4, 2.5)

The influence of the substrate thickness h is much stronger than the influence of ε_r. The higher the thickness h (Fig. 8), the less the radiation, and lower the values of the radiation conductance G_{rad}. The maximum values of R_{in} are all at the same frequency (9,6 GHz), which is lower than the expected for the resonator (10 GHz). The same can be seen also for the reflection coefficient (Fig. 9).

amplitudes of R_{in} and S_{11} become more pronounced. The reason for this effect is that besides the already mentioned influences there is also the influence of the patch dimensions. This influence is also responsible for the shift of the resonant frequency with increase of the thickness h in Figs. 3 and 4

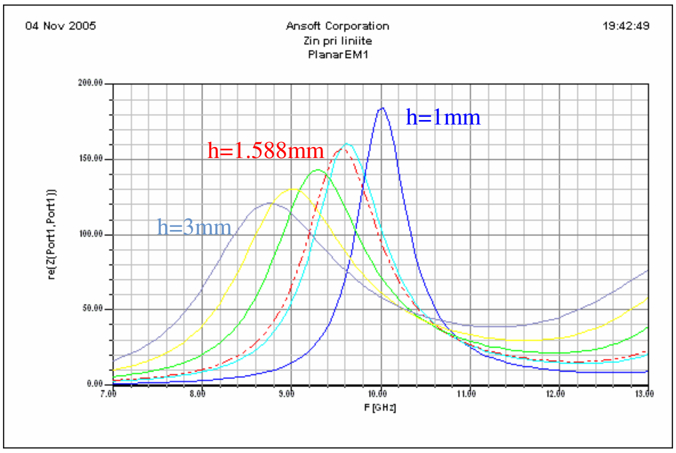


Fig. 4. $R_{in}(f)$, h = (1, 1.5, 1.588, 2, 2.5, 3) mm

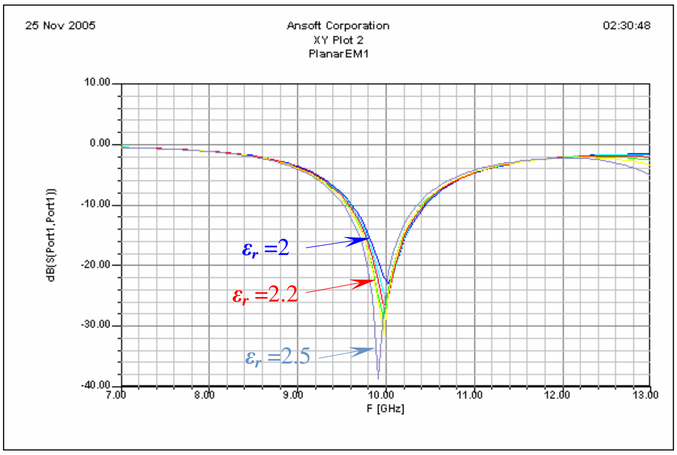


Fig. 7, $S_{11}(f)$, ε_r = (2, 2.1, 2.2, 2.3, 2.4, 2.5)

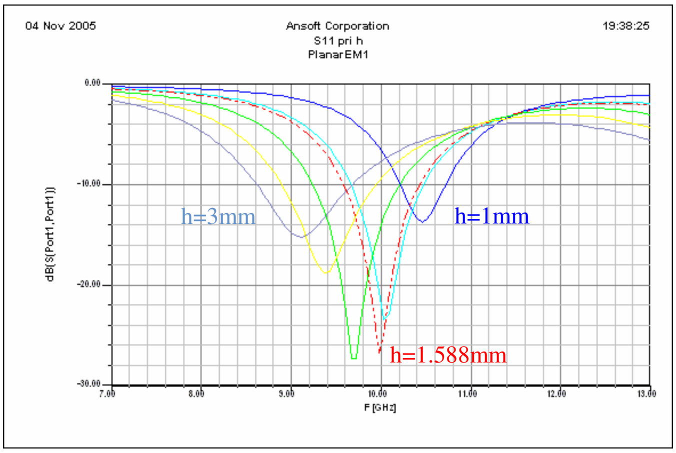


Fig. 5. $S_{11}(f)$, h = (1, 1.5, 1.588, 2, 2.5, 3) mm

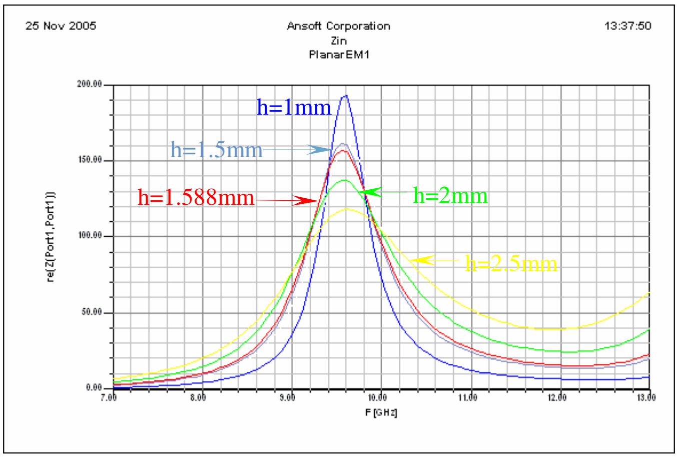


Fig. 8. $R_{in}(f)$, h = (1, 1.5, 1.588, 2, 2.5) mm

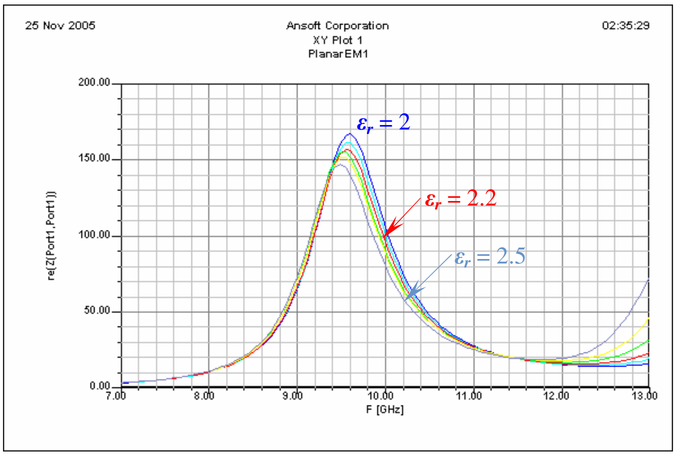


Fig. 6. $R_{in}(f)$, ε_r = (2, 2.1, 2.2, 2.3, 2.4, 2.5)

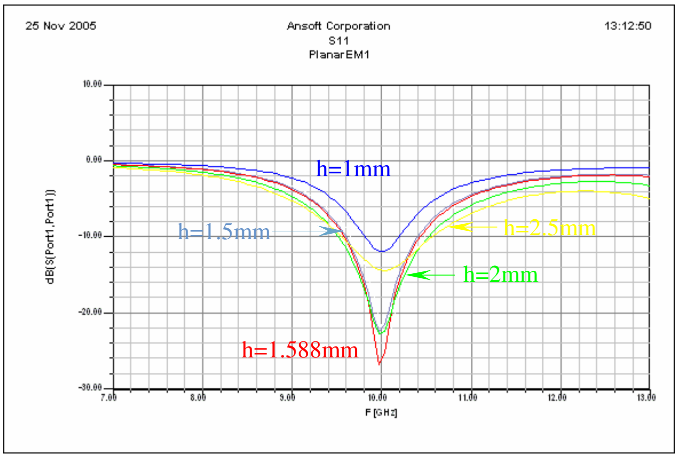


Fig. 9. $S_{11}(f)$, h = (1, 1.5, 1.588, 2, 2.5) mm

As for the first analyzed cases (Figs. 2 - 5), and in comparison with the latter cases (Figs. 6 - 9), it can be seen that the shifting of the resonance frequency and the changes in the

IV. CONCLUSIONS

A parametric study of the influence of the substrate parameters (ε_r и h) has been made in this paper. The effect of their variation over the input impedance and over the reflection coefficient of a rectangular patch has been analyzed. Best matching has been reached at higher values of ε_r and middle values of h. With increasing of ε_r there is a change in the resonant frequency, while with varying of h this effect doesn't take place. On purpose of thoroughness of the analysis, the cases of varying of the examined parameter only are analysed, as well as with recalculating the dimensions of the patch for every value of the parameter analysed.

REFERENCES:

1. C. A. Balanis, J. Willey, *Antenna Theory (Analysis and Design)*, New York, 1998;
2. D. M. Pozar, "A Review of Bandwidth Enhancement Techniques for Microstrip Antennas" in *Microstrip Antennas, The Analysis and Design of Microstrip Antennas and Arrays, D. M. Pozar and D. H. Schaubert (Eds.), New York, IEEE Press*, pp. 157-166, 1995;
3. K. R. Carver, and J. W. Mink, "Microstrip Antenna Technology", *IEEE Trans. Antennas Propaga.*, vol. AP-25, no.1, pp. 2-24, Jan. 1981;
4. D. M. Pozar, "Microstrip Antennas", *Proc. IEEE*, vol. 80, no.1, pp. 76-81, Jan. 1992;
5. R. E. Munson, "Conformal Microstrip Antennas and Microstrip Phased Arrays", *IEEE Trans. Antennas Propaga.*, vol. 22, pp. 74-78, Jan. 1974.

Automated Measurements with Parallel Computation of the Radiations and the Surface Waves from Patch Antennas

Chavdar Levchev[1] and Nikolay Stoyanov[2]

Abstract – **In this paper the block scheme at automated system of control the measurements with parallel computations of the information is shown. The characteristics of measurement far field and near field are examined. One of the main purposes is measure of surface waves in the microstrip antenna arrays. The functions of everyone measurement tools are discussed. The special software about these measurements is made. For example the graphics and diagrams are presented.**
Keywords – **Automated measurements, surface waves, patch antennas, far field, near field**

I. INTRODUCTION

The measurement of electromagnetic radiations realize by using concrete standards. In the last few years the measurements are automated. The antennas, the measurement instruments, amplifiers and cables are calibrated and certificated and their parameters are under permanent control [1]. The antennas under test are put on a special roll-over-azimuth positioner in the anechoic chamber for different measurements. The measurement is realized outside of anechoic chamber. In this work is given short description at automated system of control the measurements with parallel computations of the information.

II. AUTOMATED SYSTEM OF CONTROL THE MEASUREMENTS AND PARALLEL COMPUTATIONS OF THE INFORMATION

The block scheme at automated system for control of the measurements with parallel computations of the information is shown at fig. 1. The screening of the chamber together with microwave absorber can provide suitable environment. This type measurement chamber is trying to simulate the conditions of the free space. The screening reduces the noise level from the surround area and other external influences. The microwave absorbers minimize the unwanted reflected waves from the walls which have influence at the measurements. In the practice, it is easy to reach high levels of attenuation (from 80 *dB* to 140 *dB)* by screening at the interferences in the surrounding area which usually makes these interferences negligible.

The special roll-over-azimuth positioner rotates from 0° to 360° in horizontal plane and can be used for control of antenna under test (AUT) height.

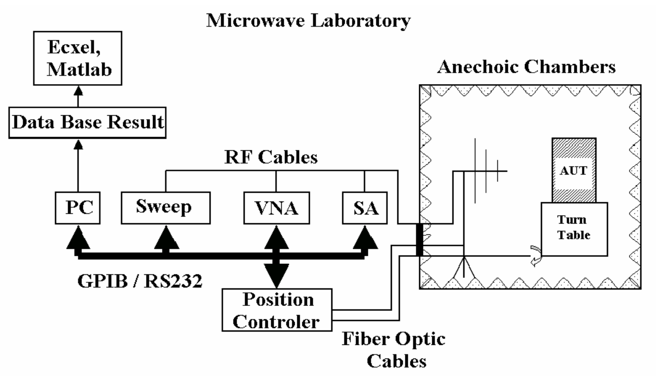


Fig. 1. Block scheme of automated system for control of the measurements

Inside the anechoic chamber, the antenna under test is put on the roll-over-azimuth positioner. In the chamber is mounting a measurement antenna over control mast which allows changing the polarization. The control of the roll-over-azimuth positioner and the mast realized by helping of positioner-controller, which one is controlled by a computer. The connection between positioner and the mast realizes by optic cables which ones together with antenna RF cables go out and go into the camera through a special panel with suitable couplings. Special software for control and treatment of the results is developed and introduced for this whole measurement process. A software fragment which control the positioner is shown at fig. 2.

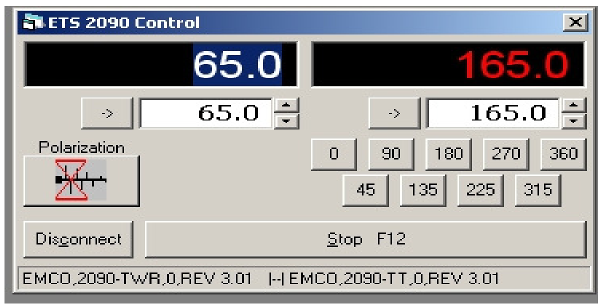


Fig. 2. Positioner Panel

Another basic element of measurement process is the measurement antenna. The antenna must be calibrated standard antenna with parameters – gain (G), antenna factor (AF), VSWR and diagram pattern. A short description of developed methods of measurement at electromagnetic radiation in anechoic chamber is shown in fig. 3.

[1]Chavdar Levchev is with the Ministry of Interior – Bulgaria
[2]Nikolay Stoyanov is with Ministry of Interior – Bulgaria, E-mail: n_m_stoyanov@abv.bg

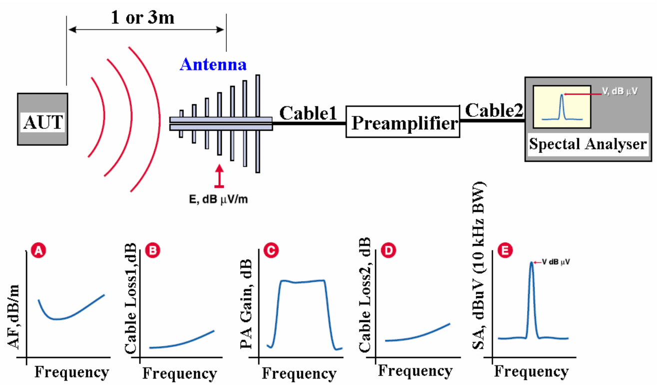


Fig. 3. Methods of measurement at electromagnetic radiation in anechoic chamber

$$E(dB\mu Vm^{-1}) = V(dB\mu V) + CL1(dB) + CL2(dB) + AF(dBm^{-1}) - PAG(dB)$$

A is an antenna factor like. The dimensions are dB/m. The antenna factor connects the voltage of antenna output with conditional power of the field around the antenna. The antenna factor is different for each frequency. B, D are corrections of the signal which are result of cable losses from attenuation and mish-mash.

C - gain of preamplifier like a frequency function,

E – the measurement value about given frequency of SA with RBW=10 kHz, VBW=3kHz In the Anechoic chamber (Fig.3) at the place of measurement antenna put the tested antenna with the known parameters. The antennas are turned directly one opposite other. The signal of the etalon antenna comes from the sweep-generator. The level of received signal at the measurement antenna is measured with spectral analyzer. After that at the place of measurement antenna is put a etalon antenna and measure her level. The levels can be measured with the spectral analyzer and with the vector network analyzer/ (VNA) also.

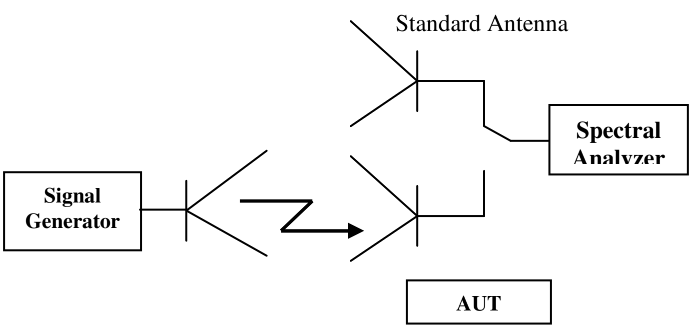


Fig. 4. Comparison between the antennas

$Gaut = L1 - L2 + C$, where Gaut is the gain of the measurement antenna (dBi), L1 – measured level of the received signal from the AUT (dB), L2 – measured level of the received signal from etalon antenna (dB), C –gain of the etalon antenna (dBi).

VSWR must be measured at the cables and the amplifiers which are used in the measurement because there are losses from mish-mash. The attenuation losses in the cables and preamplifiers gain are measured also.

All these measurements make with the VNA /vector network analyzer/ which give a chance be measured the amplitude, the frequency and the phase of (at) the different parameters at the measurement antennas. The mentioned above antenna parameters, amplifiers and cables are include in the computation algorithm about measured electromagnetic radiations. Periodically is necessary to make an examination at diagram pattern of measurement antennas. This examination can be made in the anechoic chamber where at the place of measurement antenna is put an etalon antenna for the necessary frequency diapason. The roll-over-azimuth positioner is rotating with fixed step up to 360°. The measurement is automated and special software is developed. A fragment from this software is shown at fig. 5.

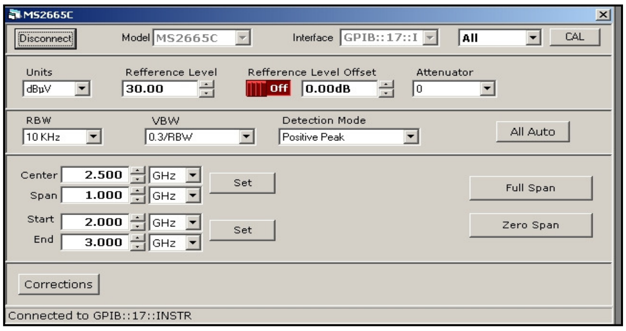


Fig. 5. Spectral Analyzer panel

The basic part in this system is the receiver. In this case this receiver is a spectral analyzer (SA) ANRITSU MS 2665C. The SA have a determine sensitivity which define the general sensitivity at the whole system[4]. SA measures the amplitude-frequency characteristics at the electromagnetic field. The measurement with SA is automated also and a fragment from this software is shown at fig. 6.

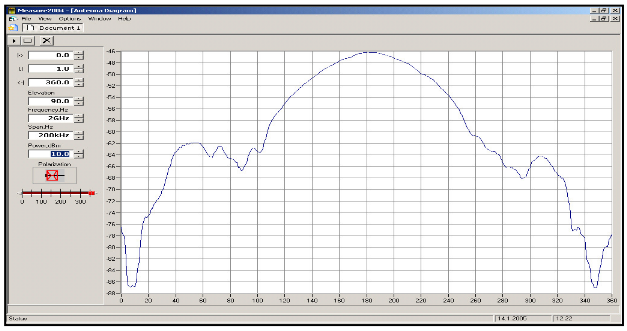


Fig. 6. Measured antenna pattern

The distance between the two antennas can be 1m or 3m and the antenna, which one must be measured, is put on the rotating table at high 0,8m.

The measurement antenna is mounted on a rotating mast and the field strength for the two polarizations, horizontal and vertical respectively, is measured. The rotation at the table is at intervals of 45°. The frequency diapasons are from 30 MHz to 15GHz in the “positive peak” mode at the spectral analyzer. The measurements are making for e two highs. The measurement and processing of the information is parallel. The measurement files are saved in the data base for the different users.

III. NEAR FIELD MEASUREMENT

The radiation field from a transmitting antenna is characterized by the complex Poynting vector **E x H*** in which **E** is the electric field and **H** is the magnetic field. Close to the antenna the Poynting vector is imaginary (reactive) and **(E, H)** decay as 1/r [5]. These two types of fields dominate in different regions in space around the antenna. Based on this characterization of the Poynting vector, can identify three major regions (Fig. 7).

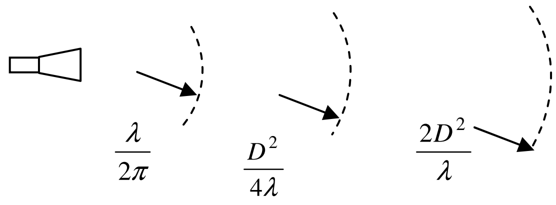


Fig. 7. Radiation regions

– *Reactive field*

This region is the space immediately surrounding the antenna. The extend of this region is $0<r<\lambda/2\pi$, where λ is the wavelength.

– *Radiating Near-Field*

Beyond the immediate neighborhood of the reactive field the radiating field begins to dominate. The extend of this region is $\lambda/2\pi<r<2D^2/\lambda$, where D is the largest dimension of the antenna. This region can be divided into two subregions. For $\lambda/2\pi<r<D^2/4\lambda$ the fields decay more rapidly than 1/r and the radiation pattern (relative angular distribution of the field) is dependent on r. For $D^2/4\lambda<r<2D^2/\lambda$ the fields decay as 1/r, but the radiation pattern is dependent on r.

– *Radiating Far-Field*

Beyond the radiating Near-Field region $r>2D^2/\lambda$ or $r>10\lambda$ (criterion for small antennas) the Poynting vector is real (only radiating fields) and has only two components in spherical coordinates (θ, φ)

IV. PLANAR SCANNING TECHNIQUE

In the planar scanning technique, a probe antenna is moved in a plane situated in front of the AUT and the received signal (amplitude and phase) is record. The position of the probe is characterized by the coordinates (x, y, z_0) in the xyz coordinate system of the AUT. During the scanning, z_0, is kept constant, while x and y are varied. The distance z_0 is approximately 3λ to avoid the sampling of the reactive energy of the AUT. The dimensions of the Near-Field scanning aperture must be large enough to accept all significant energy from the AUT. The scan dimensions, D, have to meet the criterion $D_s > D+2z_0 \tan\theta$, where D is the largest AUT dimension and θ is the maximum processed radiation pattern angle (Fig. 8). For a specific scanner with an allowable scan area D_p , this criterion determines the maximum and minimum AUT size ($D_{min} \approx 2\lambda$).

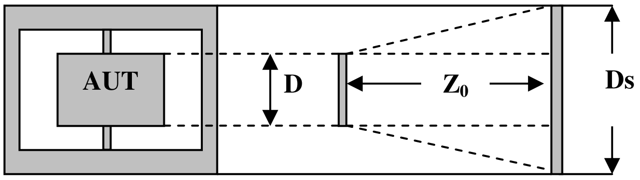


Fig. 8. Scan Size

The measured Near-Field data ***E*** (x, y, z_0) is transformed into the plane wave spectrum ***E*** (k_x, k_y) in the K-space, by a two-dimensional Fourier transform. The coordinates (k_x, k_y) in the K-space are related to the spherical coordinates (θ, φ), through the relationships $k_x = k\sin\theta\cos\varphi$, $k_y = k\sin\theta\sin\varphi$ and $k=2\pi/\lambda$. The antenna plane wave spectrum is distorted by the angular response of the probe. This effect can be deconvoluted from the AUT angular response by taking the ratio of the total plane wave spectrum to the probe spectrum. This operation is known as **probe correction**. The plane wave spectrum in the visible range, $-k< k_x, k_y<k$, is proportional to the radiation pattern F (θ, φ). Accordingly, the radiation pattern can be considered as a spatially band-limited function in the K-space on which Nyquist sampling theory applies and the sampling space can be chosen as $\Delta x<\Delta y< \lambda/2$.This sampling criterion ensures that no aliasing occurs in the visible range. For high-gain antennas are interested only in a limited angular sector around the AUT main beam.

V. MEASUREMENT OF THE SURFACE WAVE

The planar scanning technique is very expensive for experimental measurements. That is why, another technique for measuring various properties of electromagnetic surfaces have been developed [2, 3]. The presence of surface wave modes is detected in the transmission between two antennas positioned near the surface. These antennas can be simple monopole probes, or they can be specifically designed to efficiently launch surface waves. By varying the polarization of the antennas, one can distinguish between TM and TE modes [4]. The measurements on the textured surface indicate a frequency band in which there are no propagating surface waves.

The reflection phase can be measured using a pair of horn antennas directed toward the surface. The phase of the reflected wave is measured with respect to a surface with known reflection properties, such as a flat metallic surface. Within the surface wave band gap, the textured surface reflects in-phase, rather than out-of-phase.

- *TM Surface Waves*

In TM surface waves, the electric field forms that extend vertically out of the surface. TM waves cam be measured using a pair of small monopole antennas oriented normally with the respect to the surface, as is shown in fig. 9.The vertical electric field of the probe couples to the vertical electric field of the TM surface waves.

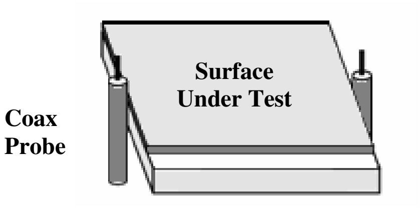


Fig. 9. TM surface wave measurement using monopole probe antennas

For improved signal, a flared parallel plate waveguide structure functions as a more effective TM surface wave antenna. This type of antenna can be made from a triangular piece of microwave circuit board material, with cooper cladding on both sides. This structure provides a smooth transition between the mode in the coaxial cable and the surface wave mode, producing a stronger transmission signal than the small monopole.

- *TE Surface Waves*

In TE surface waves, the electric field is parallel to the surface and the magnetic field forms vertical loops that arc out of the surface. They can be measured with a pair of small monopole probes oriented parallel to the sheet as shown in the Fig. 10. The horizontal electric field of the antenna couples to the horizontal electric field to the TE waves. Since this configuration lacks the symmetry of the vertical monopole where will be much greater cross-coupling to TM waves that may complicate the measurement.

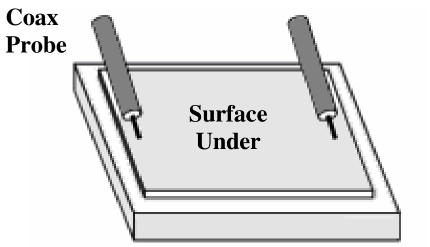


Fig. 10. TE surface wave measurement using monopole probe antennas

The typical E and H surface wave measurements using the methods shown at fig.9 and fig.10 were made by authors with vector network analyzer Anritsu 37247D. For the purposes of the measurements were used flat metal surface and FR4 dielectric slab with thickness 0,5mm and 2,5mm respectively. The frequency range is from 0.04GHz to 5GHz.

The obtained experimental results for E- and H-field are shown on fig. 11 and fig. 12, respectively. This experimental results shows that, the increasing of dielectric slab thickness, leads to increasing of mutual coupling between the ports. This effect is due to surface waves, which propagate on the material surface.

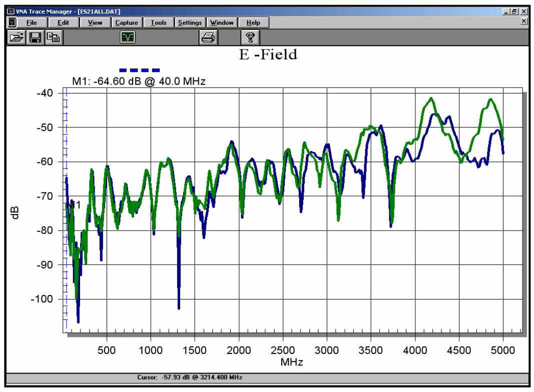


Fig. 11. E-field measurements

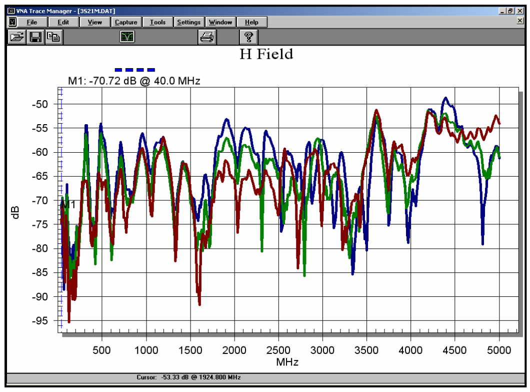


Fig. 12. H-field measurements

VI. Conclusion

Automated control system of measurements and parallel computation of the information, which is created, give a chance for continuously measurement elaboration. The measurement precision improves and the measurement time and the following treatment of the received results reduce. This type of work allows receiving enough experimental data about aposterior analysis and measurement microwave surface waves.

The accumulated measurement dates and mathematical treatments and modeling allow realizing the real action of improvement at diagram pattern of the microstrip antennas and suppressing of surface waves.

References

[1] Orbit/RF Microwave Test and Measurement Product Catalog, 1998

[2] A. Tankielun, U. Keller, W. John, H.Garbe, “Complex Deconvolution for Improvement of Standard Monopole Near-Field Measurement Results,” *Proceedings of the 16th International Zurich Symposium on electromagnetic Compatibility,* Zurich, 2005

[3] D.Baudry, A. Louis, B. Mazari, “A study and improvement of open-ended coaxial probe used for near-field measurements,” *Proceedings of the 16th International Zurich Symposium on electromagnetic Compatibility,* Zurich, 2005

[4] Yablonovitch, U. S. provisional patent application, serial number60/079953, filed on March 30, 1998

Bandwidth Improvement of an Aperture-Coupled Stacked Microstrip Antenna

Slavi R. Baev[1], Nikola I. Dodov[2]

Abstract – **The impedance bandwidth of a microstrip antenna is enhanced by a combination of planar and stacked parasitic elements. The proposed antenna configuration is compared with a typical aperture-coupled stacked microstrip antenna. A simulation analysis is performed and a comparison between some characteristics of the two antennas is presented.**

Keywords – **microstrip antenna, parasitic elements, bandwidth.**

I. INTRODUCTION

Microstrip antennas are widely used in the communication systems due to their advantages, which make them preferable to other types of antennas. These advantages are low profile, light weight, small size, and easy production process. The application of these antennas is limited because of their drawbacks – most important are the inherently low gain and narrow bandwidth.

There are many publications and books, which describe the variety of approaches for bandwidth improvement of the microstrip antennas [1], [2]. These approaches can be classified into three main categories. First is the impedance matching where a matching circuit is used between the feed line and the radiator. Next, the introduction of multiple resonances in the impedance characteristic of the antenna can be applied by cutting slots, excitation of several resonant modes or addition of parasitic elements with slightly different dimensions. The third approach includes insertion of loss to the system, which lowers the quality factor and therefore improves the bandwidth.

By excitation of multiple closely spaced resonances in the impedance characteristic the antenna bandwidth can be enlarged to a great extent. As already mentioned, several resonances can be excited by the use of parasitic elements. In this case the antenna construction includes two or more radiating elements with slight difference in their dimensions. Usually one element is directly fed and the others are coupled to the electromagnetic field radiated from it. The parasitic elements can be placed in horizontal direction (planar configuration) [3], [4], [5] or in vertical direction (stacked configuration) [6], [7], [8] compared to the main radiator.

[1]Slavi R. Baev is with the Department of Radiotechnics, Faculty of Communications and Communications Technologies, Technical University – Sofia, 8 "Sv. Kliment Ohridski" Blvd., Sofia 1000, Bulgaria, E-mail: sbaev@tu-sofia.bg

[2]Nikola I. Dodov is with the Department of Radiotechnics, Faculty of Communications and Communications Technologies, Technical University – Sofia, 8 "Sv. Kliment Ohridski" Blvd., Sofia 1000, Bulgaria, E-mail: ndodov@tu-sofia.bg

The aim of the present publication is to propose a new antenna construction, which combines the advantages of the planar and stacked configurations in order to improve the impedance bandwidth of a typical aperture-coupled stacked microstrip antenna. For the purpose two pairs of microstrip dipoles with different size are placed above the main radiator, instead of one wide rectangular patch. In this way two additional resonances are introduced into the impedance characteristic, instead of one.

In Section 2 the construction and the principles of operation of the two antennas are described. In Section 3 a simulation investigation of the two configurations is performed and a set of graphs for some of the antenna characteristics are presented.

II. ANTENNA DESCRIPTION

The authors of this publication use as a reference antenna a typical aperture-coupled stacked microstrip antenna. The main advantages of this antenna are: small geometrical area; symmetrical construction, which keeps the phase center constant and the radiation pattern symmetrical in the working frequency band; very broad impedance bandwidth. The main disadvantages are the thick construction and the high level of back radiation. The second effect appears when the feeding aperture resonates and acts as a radiator in the antenna bandwidth. This drawback can be alleviated by making the slot width short enough to move its resonant frequency away from the working band. This implies that the bandwidth is narrowed because the number of resonances is reduced by one.

The structure of a typical aperture-coupled stacked micro-strip antenna with its main parameters is shown in Fig. 1.

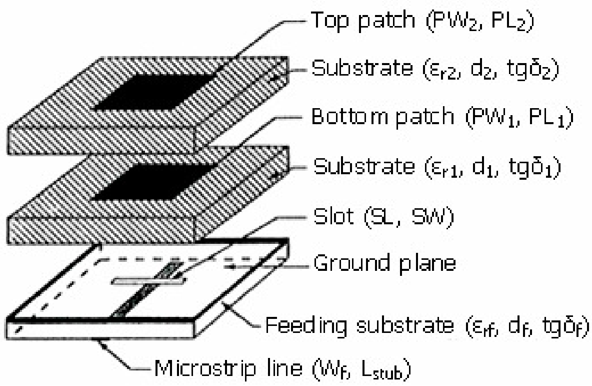


Fig. 1. Aperture-coupled stacked microstrip antenna

The designations are as follows:

PL_i, PW_i	Length and width of the patches,
ε_{ri}, d_i, $tg\delta_i$	Relative dielectric permittivity, thickness and loss tangent of the substrate materials,
SL, SW	Length and width of the coupling slot,
W_f, L_{stub}	Width of the feedline and stub length of the line beyond the center of the slot.

The antenna is fed by a microstrip line located at the bottom of the lowest substrate. The aperture in the ground plane is excited by the currents flowing on the feedline. The electromagnetic field radiated by the feedline couples to the bottom patch through the coupling slot. The electromagnetic connection between the two patches is capacitive in nature. A detailed analysis of the aperture-coupling feeding technique is described in [9].

There are two resonances in the impedance characteristic of the aperture-coupled stacked microstrip antenna when the feeding aperture does not act as a radiator. The resonances result from the mutual operation of the resonant elements in the structure – the two patches and the slot. This factor determines the complex nature of each resonance. Reference [6] contains detailed parameter study on the effect that the geometrical dimensions of an aperture-coupled stacked microstrip patch has on the antenna characteristics.

In the present publication a new configuration that introduces a third resonance in the antenna impedance characteristic is presented. The working band is improved by positioning several parasitic elements displaced in the horizontal direction above the main patch, instead of just one as in typical construction.

The proposed antenna structure is identical with that shown in Fig. 1. The only difference is in the upper conductor layer, the topology of which along with the main dimensions is shown in Fig. 2.

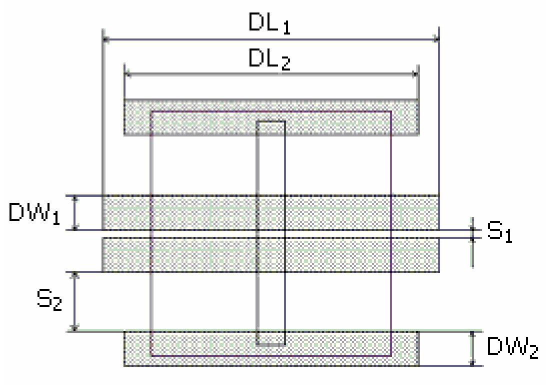


Fig. 2. Upper conductor layer layout of the proposed antenna and the contours of the bottom patch and the aperture located underneath

The designations are as follows:

DL_i, DW_i	Length and width of the two pairs of microstrip dipoles,
S_i	Dipole spacing.

The use of two pairs of microstrip dipoles as parasitic radiators is determined by the following considerations. The narrow width of the dipoles gives almost the same occupied area as the typical rectangular patch. In this way the small horizontal size of the antenna is preserved. The upper layer consists of two pairs of dipoles symmetrically positioned in relation to the center of the lower patch. Therefore the antenna geometry remains symmetrical, which predetermines a constant phase center and a symmetrical radiation pattern in the bandwidth of interest.

The geometries of the microstrip dipole and the rectangular patch being similar, their characteristics are also expected to be similar. The longitudinal surface current distributions are similar for both types of radiators, which determine the similar radiation patterns and gains. These properties of the microstrip dipole and microstrip patch are crucial for the relatively constant shape of the radiation pattern in the antenna bandwidth. However, the input impedance, bandwidth and cross-polar radiation can differ considerably. One major advantage of the dipole is the lower cross-polar radiation because of its narrow width, which decreases the transverse surface current component. A more detailed description of the printed dipoles is given in [1] and the feeding technique that uses an electromagnetically coupled microstrip line is investigated in [10] and [11]. This kind of feeding is used in the proposed structure, since the parasitic dipoles are excited by the electromagnetic field of the main patch underneath.

The addition of several parasitic elements increases the number of geometrical parameters that can be varied. This gives an additional freedom in the optimization process but at the same time makes this procedure more complex.

The dipole lengths are chosen to differ from one pair to another and also compared to the resonant length of the main patch. This leads to the presence of three resonant elements which interact and give three resonances in the system.

III. Simulation Analysis

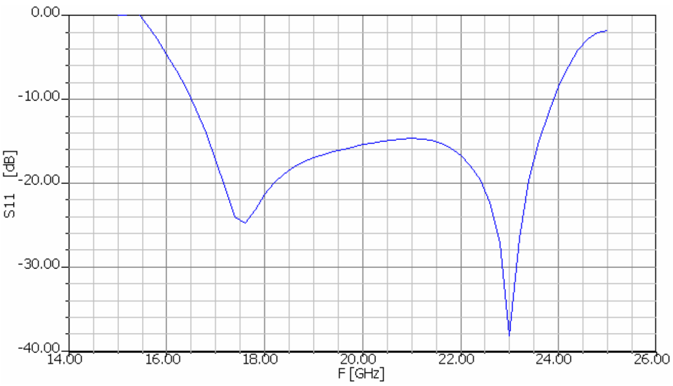
As a typical aperture-coupled stacked microstrip antenna the authors of this publication consider the Antenna #1, which is proposed and investigated in [6]. This antenna shall be called the "reference antenna" throughout the paper. The choice of this particular antenna is dictated by the good impedance characteristics and wide bandwidth. The experimental results measured in [6] show a bandwidth BW=32,5 % (17,1 – 23,75 GHz) for return loss S_{11}<-10 dB and a center frequency f_0=20,42 GHz. The aim of the present publication is to propose antenna with enhanced bandwidth (determined for S_{11}<-10 dB) compared to the reference antenna. The same materials are used in both antenna structures and slight changes in some geometrical parameters are only made for the proposed antenna in order to achieve the desired goal.

The parameters for both antennas are:

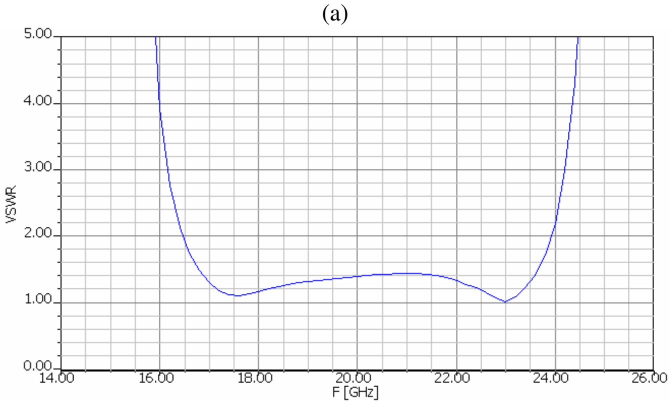
Reference antenna: PL_2=PW_2=3,8 mm; ε_{r2}=2,33; d_2=0,787 mm; $tg\delta_2$=0,0012; PL_1=PW_1=3,5 mm; ε_{r1}=2,2; d_1=0,508 mm; $tg\delta_1$=0,0009; SL=3,2 mm; SW=0,4 mm; ε_{rf}=2,2; d_f=0,508 mm; $tg\delta_f$=0,0009; W_f=1,55 mm; L_{stub}=1,8 mm;

Proposed antenna: DL_1=4,8 mm; DW_1=0,5 mm; DL_2=4,2 mm; DW_2=0,5 mm; S_1=0,1 mm; S_2=0,85 mm; ε_{r2}=2,33; d_2=1 mm; $tg\delta_2$=0,0012; PL_1=3,44 mm; PW_1=3,5 mm; ε_{r1}=2,2; d_1=0,508 mm; $tg\delta_1$=0,0009; SL=3,2 mm; SW=0,4 mm; ε_{rf}=2,2; d_f=0,308 mm; $tg\delta_f$=0,0009; W_f=1,55 mm; L_{stub}=1,6 mm;

For the purposes of a precise comparison both antennas are analyzed with the simulation software Ansoft Designer. This approach excludes the inaccuracy caused by differences in the test setups. In Ansoft Designer the surface of the patch is divided into a mesh of triangles and rectangles, the size of which depends on the working frequency. To generate a solution this software employs the mixed-potential integral equation (MPIE) method. The method of moments (MoM) is applied to the MPIE to obtain the current distribution on the surface mesh. Then the antenna characteristics like return loss (S_{11}), voltage standing-wave ratio (*VSWR*), gain, radiation pattern are calculated from this current distribution.



(a)



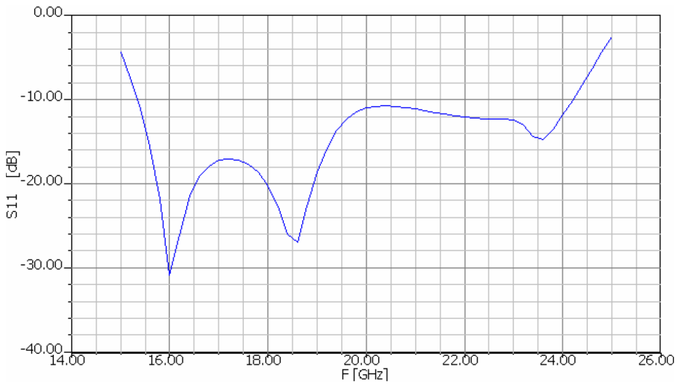
(b)

Fig. 3. Impedance characteristics of the reference antenna as a function of frequency: (a) S_{11}; (b) *VSWR*.

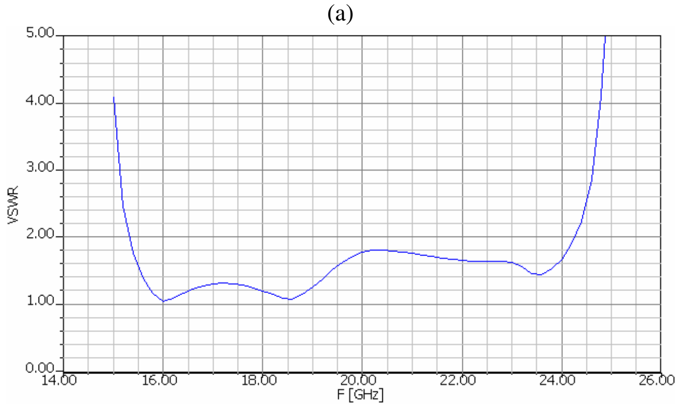
Figs. 3(a) and (b) present S_{11} and *VSWR* as a function of frequency for the reference antenna and Figs. 4(a) and (b) - for the proposed antenna. As expected the reference antenna has only two minima in the impedance characteristic, and for the proposed antenna there are three resonances.

From Fig. 3(a) it is seen that the reference antenna satisfies the condition S_{11}<-10 dB in the range 16,5 – 23,9 GHz or this corresponds to an impedance bandwidth of 36,6 % centered on f_0=20,2 GHz. Fig 4(a) shows the following results for the proposed antenna: S_{11}<-10 dB in the range 15,3 – 24,2 GHz or a bandwidth of 45,1 % centered on f_0=19,75 GHz. The resultant absolute value of bandwidth improvement for the proposed antenna compared to the reference antenna is 8,5 % and the relative enhancement is about 23 %. This large improvement is due to the introduction of a third resonance in the impedance characteristic of the patch antenna.

As mentioned above a given resonance can not be related only to a particular antenna element [6]. Therefore it is difficult to determine a cause for the excitation of each one resonance in the system. The study performed on this problem allows noting some specific features. When the inner pair of dipoles is removed the first resonance (for the lowest frequency – Fig. 4(a)) disappears. Only one of the low frequency resonances leaves. Hence the first resonance is determined to a great extent by the inner pair of dipoles. This behavior may be expected because these dipoles are with maximum length and therefore minimum resonant frequency. When the outer dipoles are removed the bandwidth is narrowed from its high frequency end due to the loss of the third resonance. However there remains a minimum at about 21 GHz, which is related to the coupling aperture. Therefore the third resonance is a result from the mutual operation of the outer dipoles and the slot. This is in correspondence with the high level of S_{11} for the third resonance. The outer dipoles are located above the ends of the aperture where the coupled field is weaker compared to the center. This implies weaker excitation of these dipoles and the resonance related to them.



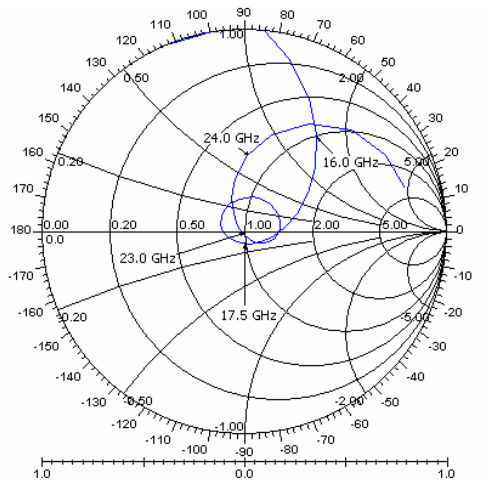
(a)



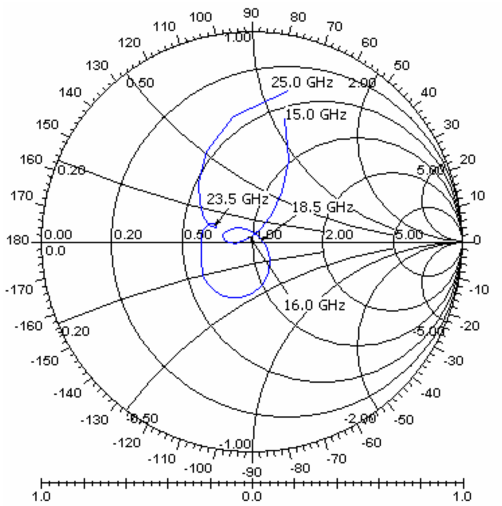
(b)

Fig. 4. Impedance characteristics of the proposed antenna as a function of frequency: (a) S_{11}; (b) *VSWR*.

Figs. 5(a) and (b) show the Smith charts for the two antennas. As expected for the reference antenna there is only one well formed closed loop in the center of the chart. This is a typical behavior for a double-resonant antenna. As seen in Fig. 5(b) there are two loops centered on the Smith chart for the proposed antenna. The bigger one corresponds to the resonances at 16 GHz and 18,5 GHz and the smaller one – to the weaker resonance at 23,5 GHz.



(a)



(b)

Fig. 5. Smith chart: (a) Reference antenna (b) Proposed antenna.

IV. Conclusion

In this paper an antenna, which combines the advantages of the planar and the stacked parasitic configurations in order to achieve bandwidth enhancement, is presented. This is realized through the substitution of the standard upper patch in an aperture-coupled stacked microstrip antenna by two pairs of microstrip dipoles with different lengths. The insertion of an additional resonant element in the antenna construction leads to the excitation of a third resonance in the impedance characteristic, while for a typical aperture-coupled stacked patch there are only two resonances. This causes improvement of the impedance bandwidth.

A simulation analysis on the proposed and a typical aperture-coupled stacked microstrip antenna is performed. A set of graphs for some of the main antenna parameters are presented. The results show relative bandwidth enhancement of 23 % and an absolute value of 8,5 %.

The addition of more geometrical parameters in the antenna construction complicates its optimization. The authors do not claim for an optimal design of the presented antenna. The main purpose of the publication is to show that the used bandwidth enhancement approach gives good results and can be realized in practice. The future efforts of the authors will be concentrated on the performance of a parameter analysis which must reveal the relation between the antenna characteristics and its geometry and to outline optimization guidelines for this type of antenna.

References

[1] R. Garg, P. Bhartia, I. Bahl, and A. Ittipiboon, *Microstrip Antenna Design Handbook,* Norwood, Artech House, 2001.

[2] D. M. Pozar, “A Review of Bandwidth Enhancement Techniques for Microstrip Antennas”, in *Microstrip Antennas, The Analysis and Design of Microstrip Antennas and Arrays,* D. M. Pozar, and D. H. Schaubert(Eds.), New York, IEEE Press, pp. 157-166, 1995.

[3] R. Garg, and V. S. Reddy, “A Broad-Band Coupled-Strips Microstrip Antenna”, *IEEE Trans. Antennas Propagat.*, vol. 49, pp. 1344–1345, 2001.

[4] G. Kumar and K. C. Gupta, “Broad-Band Microstrip Antennas Using Additional Resonators Gap-Coupled to the Radiating Edges,” *IEEE Trans. Antennas Propagat.,* vol. 32, pp. 1375-1379, 1984.

[5] G. Kumar and K. C. Gupta, “Nonradiating Edges and Four Edges Gap-Coupled Multiple Resonator Broad-Band Microstip Antennas,” *IEEE Trans. Antennas Propagat.,* vol. 33, pp. 173-178, 1985.

[6] F. Croq and D. M. Pozar, “Millimeter Wave Design of Wide-Band Aperture-Coupled Stacked Microstrip Antennas,” *IEEE Trans. Antennas Propagat.*, vol. 39, pp. 1770–1776, 1991.

[7] S. D. Targonski, R. B. Waterhouse, and D. M. Pozar, “Design of Wide-Band Aperture-Stacked Patch Microstrip Antennas,” *IEEE Trans. Antennas Propagat.*, vol. 46, pp. 1245–1251, 1998.

[8] A. Ittipiboon, B. Clarke, and M. Cuhaci, “Slot-Coupled Stacked Microstrip Antennas” – *IEEE Antennas and Propagation Society International Symposium*, vol. 3, pp. 1108-1111, 1990.

[9] D. H. Schaubert and P. L. Sullivan, “Analysis of an Aperture-Coupled Microstrip Antenna,” *IEEE Trans. Antennas Propagat.,* vol. 34, pp. 977-984, 1986.

[10] H. G. Oltman, and D. A. Huebner, “Electromagnetically Coupled Microstrip Dipoles”, *IEEE Trans. Antennas and Propagation,* vol. 29, pp. 151-157, 1981.

[11] P. B. Katehi, and N. G. Alexopoulos, “On the Modeling of Electromagnetically Coupled Microstrip Antennas – The Printed Strip Dipole”, *IEEE Trans. Antennas and Propagation,* vol. 32, pp. 1179-1186, 1984.

Neural Network – Based DOA Estimation and Beamfoming for Smart Antenna

Maja Sarevska, Bratislav Milovanović, and Zoran Stanković

Abstract – **This paper considers the *Neural Network* (NN) based smart antenna solution for both *Direction Of Arrival* (DOA) estimation and null-steering beamforming, providing a complete NN solution. The main purpose is to present the influence of the imprecise DOA estimations on the NN beamforming performances. Computer simulations for given example will show that the uncertainty in DOAs in the range less than ±0.5° will provide satisfactory NN based beamforming.**

Keywords – **neural network, DOAs, beamforming.**

I. Introduction

A NN [1,2,3] is a powerful tool in signal processing. Due to its strong numerical approximation capability, it is widely used in identification and optimization. The research in antenna arrays is most active due to its military and commercial applications, and also in new solutions for future radiotelescopes. The focus of antenna array signal processing is on DOA estimation and beamforming.

Currently, superresolution algorithms, such as the *Multiple Signal Classification* (MUSIC) and *Estimation of Signal Parameters via Rotational Invariance Technique* (ESPRIT) [4], can be used to perform the direction finding or angle of arrival estimation. One drawback of these algorithms is the difficulty of implementing them in real time because of their intensive computational complexity. NNs, on the other hand, due to their high-speed computational capability, can obtain results in real time. Once the DOAs of sources are available, the beamforming algorithm can be used to track, in real time, sources of interest, and null out the other sources as interference. This is done by controlling the beampattern of an antenna array in an adaptive means. Conventional methods are typically linear algebra-based methods. They require time-consuming matrix inversion computation and cannot meat real-time requirements. Conventional beamformers require highly calibrated antennas with identical element properties. Performance degradation often occurs due to the fact that these algorithms poorly adapt to element failure or other sources of errors. On the other hand, NN-based antenna array do not suffer from these shortcoming. They use simple addition, multiplication, division, and threshold operations in the basic processing element. They possess advantages as massive parallelism, nonlinear property, adaptive learning capability, generalization capability, strong fault-tolerant capability and insensitivity to uncertainty.

Maja Sarevska, Bratislav Milovanović, and Zoran Stanković are with Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18 000 Niš, Serbia and Montenegro, E-mail: maja_sarevska@yahoo.com, [bata,zoran]@elfak.ni.ac.yu

The paper is organized as follows: Section II describes the DOA estimation problem, Section III presents the NN Null-Steering algorithm for beamforming, Section IV presents the NN DOA uncertainty problem, Section V is presenting the results gained from computer simulations, and in Section VI some conclusion remarks are noted.

II. NN DOA Estimation

Let observe a linear antenna array with M elements, let K (K<M) be the number of narrowband plane waves, centered at frequency ω_0 impinging on the array from directions $\{\theta_1\ \theta_2\ .\ .\ .\ \theta_K\}$. Using complex signal representation, the received signal in the ith array element is:

$$x_i = \sum_{m=1}^{K} s_m(t) e^{-j(i-1)K_m} + n_i(t), \quad i = 1,2,...M \tag{1}$$

where $s_m(t)$ is the signal of the m-th wave, $n_i(t)$ is the noise signal received at the i-th sensor and

$$K_m = \frac{\omega_0 d}{c}\sin(\theta_m) \tag{2}$$

where d is the spacing between the elements of the array, and c is the speed of the light in free-space. In vector notation the output of the array is:

$$\boldsymbol{X}(t)=\boldsymbol{A}\boldsymbol{S}(t)+\boldsymbol{N}(t) \tag{3}$$

where $\boldsymbol{X}(t)$, $\boldsymbol{N}(t)$, and $\boldsymbol{S}(t)$ are:

$$\begin{aligned} \boldsymbol{X}(t)&=[\ x_1(t)\ x_2(t)\ .\ .\ .\ x_M(t)]^T \\ \boldsymbol{N}(t)&=[\ n_1(t)\ n_2(t)\ .\ .\ .\ n_M(t)]^T \\ \boldsymbol{S}(t)&=[\ s_1(t)\ s_2(t)\ .\ .\ .\ s_K(t)]^T \end{aligned} \tag{4}$$

In (3) $\boldsymbol{A}$ is the M × K steering matrix of the array toward the direction of the incoming signals:

$$\boldsymbol{A}=[\boldsymbol{a}(\theta_1)\ \boldsymbol{a}(\theta_2)\ .\ .\ .\ \boldsymbol{a}(\theta_K)] \tag{5}$$

where $\boldsymbol{a}(\theta_m)$ is the steering vector associated with direction θ_m:

$$\boldsymbol{a}(\theta_m)=[1\ e^{-jKm}\ e^{-j2Km}\ .\ .\ .\ e^{-j(M-1)Km}]^T \tag{6}$$

The received spatial correlation matrix $\boldsymbol{R}$ of the received noisy signals can be estimated as:

$$\boldsymbol{R}=E\{\boldsymbol{X}(t)\boldsymbol{X}(t)^H\}=\boldsymbol{A}E[\boldsymbol{S}(t)\boldsymbol{S}^H(t)]\boldsymbol{A}^H+E[\boldsymbol{N}(t)\boldsymbol{N}^H(t)] \tag{7}$$

Following the Fig.1, the antenna array is performing the mapping G: $\boldsymbol{R}^K \rightarrow \boldsymbol{C}^M$ from the space of DOAs, $\{\Theta=[\theta_1,\theta_2,\ .\ .\ .\ ,\ \theta_K]^T\}$ to the space of sensor output $\{\mathbf{X}(t)=[x_1(t)\ x_2(t)\ .\ .\ .\ x_M(t)]^T\}$. A neural network is used to perform the inverse mapping F: $\boldsymbol{C}^M \rightarrow \boldsymbol{R}^K$. For this task a *Radial Basis Function*

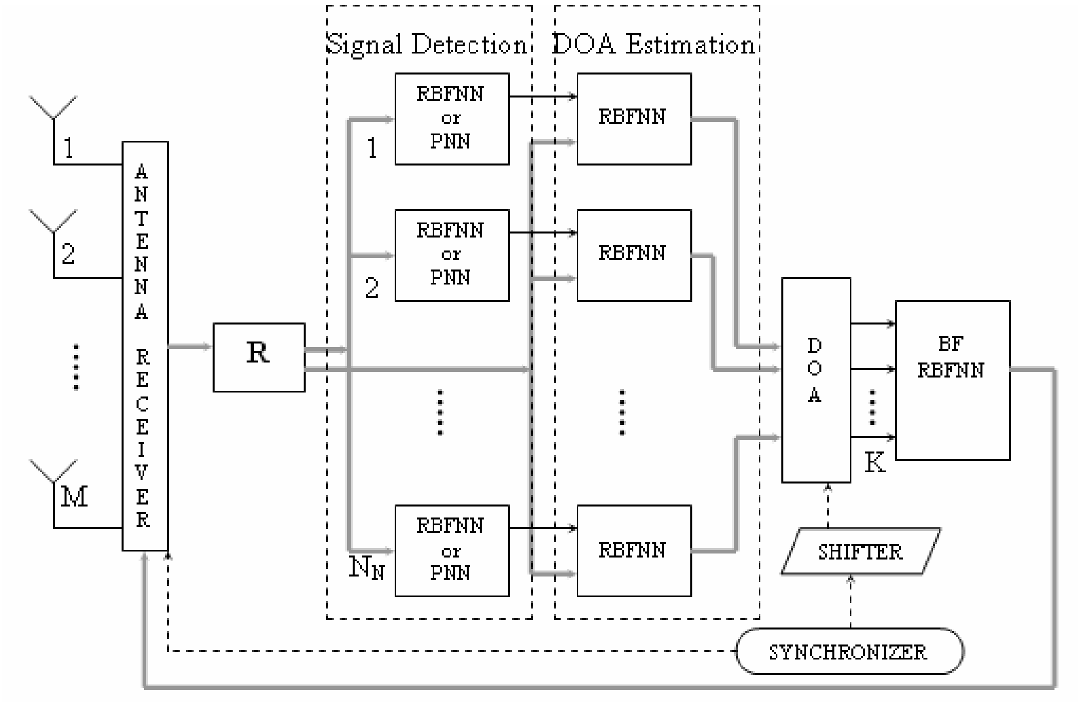


Fig.1 The Block Diagram of NN-Based Smart Antenna

(RBFNN) is used [5], instead of backpropagation neural network because the second is slower in training. In [5] the values of **R** are used at the input of the NN estimated in block **R** in Fig.1. The antenna view is divided in N_N sectors and the algorithms for detection and estimation stage are same, the difference is only in the number of nodes in the output layer. Namely, the number of the nodes in the output layer of the first stage (detection) is one (there is a signal gives one, and no signal gives 0), and the number of the nodes in the output layer of the second stage is determined by the angular resolution of the algorithm and the width of the corresponding sector.

There are a lot of learning strategies that have appeared in the literature to train RBFNN. The one used in [5] was introduced in [6], where an unsupervised learning algorithm (such as *K-means* [2]) is initially used to identify the centers of the Gaussian functions used in the hidden layer. The standard deviation of the Gaussian function of a certain mean is the average distance to the first few nearest neighbors of the means of the other Gaussian functions. This procedure allows us to identify the weights (means and standard deviations of the Gaussian functions) from the input to the hidden layer. The weights from the hidden layer to the output layer are estimated by supervised learning known as delta rule, applied on single layer networks [3]. With this procedure, for training we need 5min in detection stage and about 15min in estimation stage. An alternative is instead of using the same neural networks in both stages, to use different neural network in the first stage. The reason for this is the fact that the task of signal detection is a vector classification problem. Any input vector should be classified as 0 (there is NO signal in the corresponding sector) or 1 (there IS a signal in the corresponding sector). For this task an appropriate neural network is *Probabilistic Neural Network* (PNN), which is proposed in [7].

III. NN-BASED NULL-STEERING BEAMFORMER

Let $\boldsymbol{a}(\theta_l)$ be the steering vector in the direction where unity response is desired, and that $\boldsymbol{a}(\theta_2)$, $\boldsymbol{a}(\theta_3)$,. . ., $\boldsymbol{a}(\theta_K)$ are K-1 steering vectors of interference signal directions. We are trying to put nulls in these K-1 directions and to receive the signal from direction θ_l. We can create the antenna radiation pattern by associating a weight value to each antenna element. The desired weight vector is the solution to the following equations:

$$\boldsymbol{w}^H \boldsymbol{a}(\theta_l)=1 \tag{8}$$

$$\boldsymbol{w}^H \boldsymbol{a}(\theta_i)=0, \quad i=2,\dots,K \tag{9}$$

Using matrix notation this becomes:

$$\boldsymbol{w}^H \boldsymbol{A}=\boldsymbol{e}^T \tag{10}$$

were $\boldsymbol{e}$ is a vector with all zeros except the first element which is one:

$$\boldsymbol{e}=[1\ 0\ .\ .\ .\ 0]^T \tag{11}$$

For K=M, $\boldsymbol{A}$ is square matrix. Assuming that the inverse of $\boldsymbol{A}$ exists, which requires that all steering vectors are linearly independent, the solution for weight vector is:

$$\boldsymbol{w}^H=\boldsymbol{e}^T\boldsymbol{A}^{-1} \tag{12}$$

When steering vectors are not linearly independent $\boldsymbol{A}$ is not invertible and its pseudo inverse can be used. Observing the Eq.(12) it follows that the first row of the inverse of $\boldsymbol{A}$ forms the desired weight vector.

When the number of required nulls is less than M, $\boldsymbol{A}$ is not square matrix. A suitable estimate of weights may be produced using:

$$\boldsymbol{w}^H=\boldsymbol{e}^T\boldsymbol{A}^H(\boldsymbol{A}\boldsymbol{A}^H)^{-1} \tag{13}$$

RBFNN can successfully perform this Beamforming (BF) procedure and it is presented with block BF RBFNN in Fig.1. Unlike the other authors who use **R** at the input of the NN, in our case we use the DOAs at the input of the BF RBFNN. Given combination of DOAs correspond to given radiation pattern (antenna weight vector) that produce unity response in desired direction, since the NN is trained to give unity response only for one DOA (let say the first one). For multi-user detection we can divide the time into K slots, and each slot will correspond to one user. In k-th time slot the position of the desired signal direction: θ_k, in the input vector is first one. This time division multiplexing is synchronized with the antenna array.

The BF RBFNN same as in the DOA estimation stage, is consisted of three layers of neurons. The input layer has K neurons and the number of neurons in the output layer is 2M corresponding to real and imaginary parts of the weights of the antenna array elements. The hidden layer dimension is larger than that of the input layer. The NN weights from the input to the hidden layer are determined by mentioned *K-means* algorithm and the NN weights associated to the neuron connections from the hidden to the output layer are determined with NN training using *delta learning rule*. The BF RBFNN receives input vectors as combinations of DOAs and produces the antenna element weights at the output. Training pairs are produced using Eq, (12). In this case, dividing the space into sectors cannot perform the reduction of the number of training samples. The reason for this is the fact that the antenna element weights are associated to the whole antenna view. Some other means must be developed in order to decrease the number of training samples. Also as discussed in [8,9] limitation should be expected and future interest is to solve these limitations in order a large number of users to be served.

IV. NN DOA UNCERTAINTY PROBLEM

DOA estimation using NN concept is related to some degree of uncertainty. Namely, the actual vector of DOAs: $\Theta=[\theta_1,\theta_2, \ldots, \theta_K]^T$ is presented with estimated vector: $\Theta'=[\theta_1',\theta_2', \ldots, \theta_K']^T$ where:

$$\theta_i'=\theta_i+\Delta\theta_i\,,\quad i=1,2,\ldots,K \tag{14}$$

The parameter $\Delta\theta_i$ receives random values with uniform distribution in the interval [-maxerr, maxerr], where maxerr is maximal angle error in degrees. This maximal error is dependant from the performances of the NN concept in the DOA estimation phase. It is very important to found out the degree of accuracy that is necessary for DOA estimation in order satisfactory beamforming to be performed. This DOA uncertainty can be decreased by appropriate NN training in DOA estimation phase or by additional training in NN beamforming stage, which will probably overburden the total training in beamforming stage. In the next section the worst case will be analyzed, that is when all DOAs are assumed to be imprecise.

V. COMPUTER SIMULATIONS

Many different examples were investigated, here the results for the example when there are K=6 users and M=6 antenna elements are exposed. A regular linear antenna array was used with inter-element spacing of d=0.5 wavelengths. The BF RBFNN has 6 neurons in the input layer, 30 in the hidden, and 12 neurons in the output one. The centers of Gaussian transfer functions in the hidden layer were determined with *K-means* clustering algorithm. The variances were estimated as the mean distance of the three nearest neighboring centers from the corresponding center. The case for ϕ=const. and $\theta\in(0°\div180°)$ was analyzed. The users were placed in the space with mutual distance of 20°.

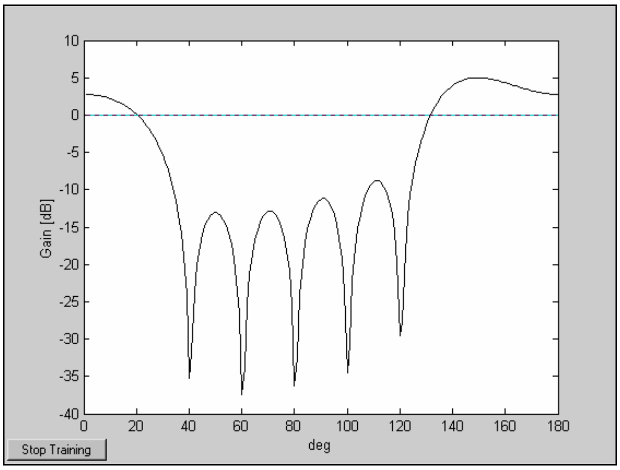


Fig. 2. Radiation pattern for 6 users at mutual distance of 20°

Fig.2 is presenting the results gained for antenna array gain (radiation pattern) for 6 users placed at mutual distance of 20°. It can be seen that NN successfully places five nulls and successfully receives the user of interest placed at 20,4°. The suppression of interference is about 30dB.

Now let assume that there are six users at mutual distance of 20 degrees and that the DOA estimation NN has performed the DOA estimation with accuracy within the range of [-maxerr,maxerr]. Fig.3 is presenting the relative error of the absolute value of the estimated array weights and Fig.4 is presenting the relative error of the argument of the estimated array weights, for the case when maxerr=0.1°. It can be concluded that null-steering NN successfully performs the beamforming almost completely neglecting the DOA uncertainty. Fig.5 and Fig.6 are presenting the results gained for the case when maxerr=0.5°. The influence of DOA uncertainty is obvious. It can be easily concluded that further enlargement of the DOA imprecision (the higher value of maxerr) will largely damage the null-steering NN beamforming performances.

We should mention that we have analyzed the case when the relative error due to imperfect NN beamforming generalization is almost zero in order to observe only the influence of the DOA estimation uncertainty.

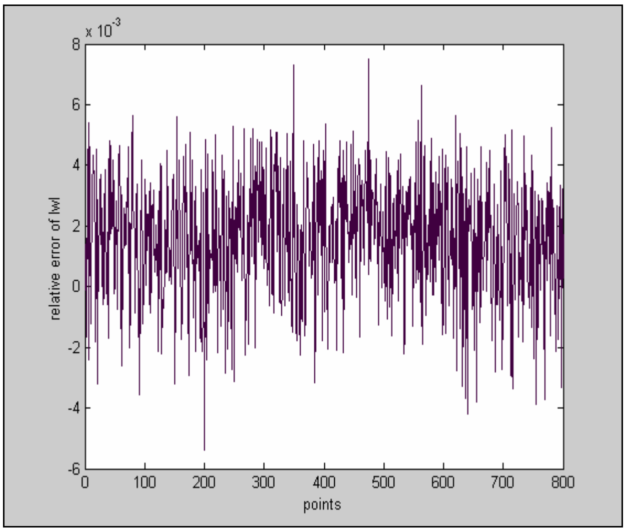


Fig. 3. Relative mean error of |*w*| for maxerr=0.1°

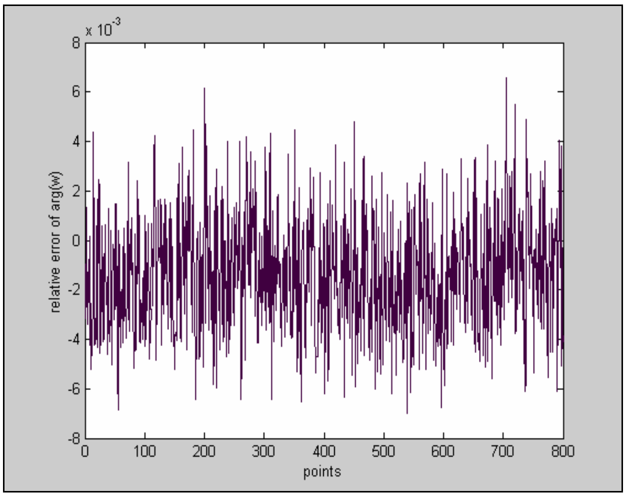


Fig. 4. Relative mean error of arg(*w*) for maxerr=0.1°

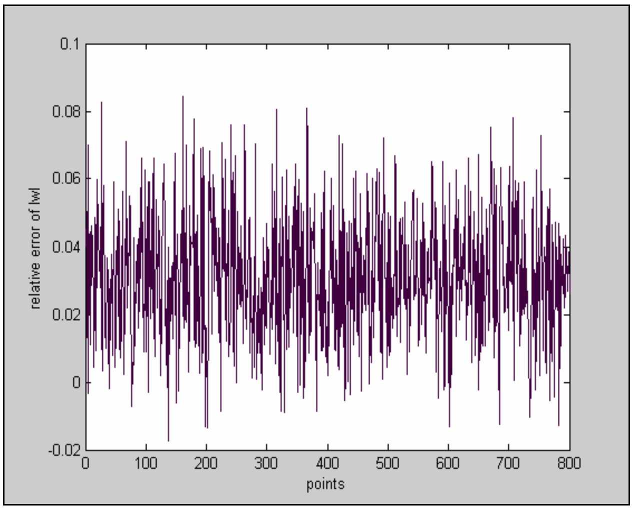


Fig. 5. Relative mean error of |*w*| for maxerr=0.5°

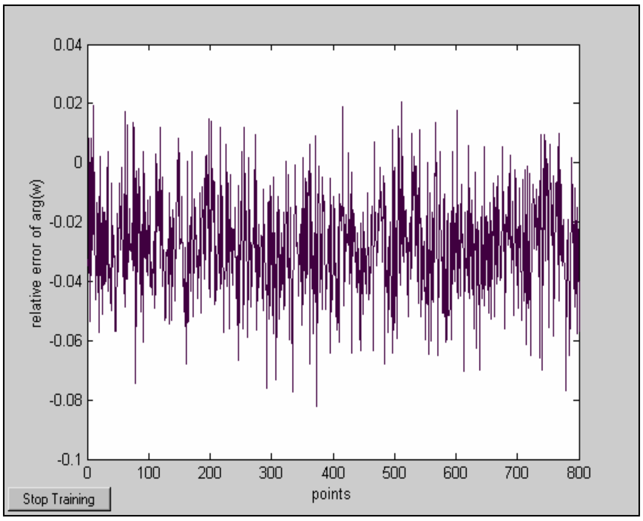


Fig. 6. Relative mean error of arg(*w*) for maxerr=0.5°

VI. Conclusions

A neural network based smart antenna solution was presented, both for DOA estimation and for null-steering beamforming providing a complete NN solution. The main issue was to present the influence of DOA uncertainty while NN estimation, on the NN beamforming performances. The results from computer simulations showed that the small DOA uncertainty doesn't infect the beamforming. For the given example the DOA precision of 0.5 degrees is providing successful NN beamforming. The DOA precision can be increased with appropriate NN training in the DOA estimation phase or additional training in beamforming stage can overcome the DOA uncertainty but in the same time it will probably overburden it.

References

[1] Yu Hen Hu, and Jenq-Neng Hwang, "*Handbook of Neural Network Signal Processing*", CRC Press LLC, Boca Raton Florida, 2001.
[2] M. M. Gupta, L. Jin, and N. Homma, "*Static and Dynamic Neural Networks*",A John Wiley & Sons, Inc., Publication, 2003.
[3] C.G. Christodoulou, and M. Georgiopoulos, "*Applications of Neural networks in Electromagnetics*", Artech House, Inc., 2001.
[4] R.O. Schmidt, "Multiple emitter location and signal parameter estimation", *IEEE Trans. Antennas propagat.*, vol.34, pp. 276-280, March1986.
[5] A.H. El Zooghby, C.G. Christodoulou, and M. Georgiopoulos, "A neural network-based smart antenna for multiple source tracking", *IEEE Trans. Antenna Prop*., vol. 48, no.5, May 2000.
[6] T.J. Moody and C.J. Darken, "Fast learning in networks of locally tuned processing units", *Neural Computat*., vol.1, pp. 281-294, 1989.
[7] M. Sarevska, B. Milovanović, and Z. Stanković, "Alternative Signal Detection For Neural Network-Based Smart Antenna", *Neurel'04*, Belgrade, pp. 85-89, Sept. 2004.
[8] Maja Sarevska, Liljana Gavrilovska "Capabilities of Neural Network – Based Smart Antenna for DOA Estimation", *Conf.* ETAI'05, Ohrid, R.of Macedonia, pp. 94-99, September 2005.
[9] Maja Sarevska, Bratislav Milovanovic, and Zoran Stankovic, "Reliability of Radial Basis Function – Neural Network Smart Antenna", *Conf.*WSEAS'05 on Communications, Athens, Greece, July 2005

Overview of COST 273 Part I: propagation modeling and channel characterization

Irina D. Sirkova[1]

Abstract – **This paper is an overview of COST 273 activities with special emphasis on electromagnetic propagation modeling. A summary is given of the entire project and some of its major achievements are outlined.**

Keywords – **Propagation modeling, channel characterization, MIMO systems, scenarios.**

I. INTRODUCTION

Action 273 is a European project (start May 2001, end May 2005) within the COST (European **CO**operation in the field of **S**cientific and **T**echnical Research, [1]) framework dealing with radio aspects of mobile and wireless networks. This paper is a presentation of Action's final report, [2], which has appeared also as a book, [3]. The first part of the paper describes briefly the structure of the Action, its work groups (WG) and sub-work groups (SWG) and their activities. The Section III overviews the work related to deterministic propagation methods and channel modeling and characterization. Special attention is paid to two major achievements (to the author's opinion) in Action 273: MIMO channel model and reference scenarios for radio network simulation and evaluation. Finally, some trends for the future, as foreseen by the Action participants, are given.

II. COST ACTION 273

The COST framework contains several areas, among them "Telecommunications and Information Society Technologies" (TIST) and Action 273 was part of this area. The Action 273 (Chairman: Prof. Luis M. Correia, Technical University of Lisbon, Lisbon, Portugal) emerged as follow-on from three previous COST projects, namely:

- COST 207: *Digital Land Mobile Radio Communications* (Mar. 1984 – Sep. 1988), which contributed to the development of GSM;
- COST 231: *Evolution of Land Mobile Radio (Including Personal) Communications* (Apr. 1989 – Apr. 1996), which contributed to the deployment of GSM1800 and to the development of DECT, HIPERLAN 1 and UMTS;
- COST 259: *Wireless Flexible Personalized Communications* (Dec. 1996 – Apr. 2000), which contributed to the deployment of DECT and HIPERLAN 1 and to the development of UMTS and HIPERLAN 2, as well as initial inputs to the next generations of HIPERLAN and 4th generation systems.

The main objective of COST 273 is "to increase the knowledge on the radio aspects of mobile broadband multimedia networks, by exploring and developing new methods, models, techniques, strategies and tools towards the implementation of 4th generation mobile communication systems", [4]. It considers frequencies ranging from the upper UHF up to millimetre waves, and data rates essentially higher than 2 Mb/s. It is also expected that it will contribute to the deployment of more or less standardized systems like UMTS and WLANs. The Action collaborates with other European projects in the TIST area, [4]. Results of the previous Actions have been taken by the industry and standardization bodies, e.g., ETSI – European Telecommunications Standards Institute; ITU-R – International Telecommunications Union – Radio Sector; 3GPP – Third Generation Partnership Project; the same is expected for the COST 273. The Action is structured in 3 Working Groups (WGs) and several Sub-Working Groups (SWGs):

- **WG 1** - Radio System Aspects. The GW1 members deal basically with: access techniques, modulation schemes, detection algorithms, signal processing, Multiple-Input Multiple-Out (MIMO) systems, and receiver structures. Joint sessions with Working Group 2 have considered MIMO channel modeling and the capacity of the MIMO channel and fading forecasting for adaptive systems. Achievements of WG 1 are reported mainly in Chapters 2 and 3 of [3]. (Note that, due to the need of joint treatment of numerous problems by experts from different WGs, the book [3] does not follow exactly the topics covered by the WGs as enumerated here.)
- **WG 2** - Propagation and Antennas. Topics addressed in WG2: propagation mechanisms, channel measurements, channel characterization, adaptive antennas, antennas modeling, and MIMO channels. Chapters 4 – 7 of the book are dedicated to the results obtained in WG 2.
  - SWG 2.1 - MIMO channel model. Goal: establishment of a model for the wireless channel that is suitable for MIMO systems, characterized by a set of parameters.
  - SWG 2.2 - Antenna performance of small mobile terminals. Goal: establishment of measuring techniques and of performance relations for antennas on small mobile terminals.
  - SWG 2.3 - Multidimensional channel measurements. Goal: Development of propagation channel measurement techniques and provision of measurement results to support channel modeling.
- **WG 3** - Radio Network Aspects. Topics encompassed in WG3: radio network planning, ad-hoc networks, traffic modeling, capacity and interference modeling, radio resources management, methodologies for performance

[1]Irina D. Sirkova is with the Institute of electronics – BAS, 72 blvd. "Tzarigradsko chaussee", 1784 Sofia, Bulgaria, E-mail: irina@ie.bas.bg

evaluation of radio networks. The work done in WG 3 is summarized in Chapters 7, 8 and 9.

- SWG 3.1: Mobile radio networks reference scenarios. Goal: definition of common reference scenarios to be used in the assessment of Radio Network Planning (RNP) and Radio Resource Management (RRM).

III. ACTIVITIES RELATED TO PROPAGATION MODELING AND CHANNEL CHARACTERIZATION

It is known that UMTS radio networks are very sensitive to variations of the radio environment and the traffic conditions, [5]. This requires UMTS planning to be based on reliable propagation models and detailed traffic predictions. Especially this is true when transition from one environment to another, with different propagation conditions, is due or when the automatic approach for UMTS planning is addressed. Nevertheless, the most accurate methods for wave propagation description available in electromagnetics are still not very often applied by the telecommunication engineers due to their complexity and large amounts of computer resources required. Generally, statistical models are preferred in the planning phase of a mobile radio system, where speed is more important than accuracy, while deterministic models are used in the design phase. Widespread adoption of otherwise physically meaningful, accurate and flexible deterministic propagation models is still very limited also by the cost and the reliability of the input databases (containing details on terrain, buildings etc.) they require. The work done essentially in WG 2 and SWG 3.1 of the Action has tried to reduce these two gaps: low use of accurate deterministic propagation methods and lack of accurate databases for them. The application of complicated methods aims to identify/solve as more problems as possible still in an early planning stage in order to facilitate the next steps.

Some of the well known deterministic methods have been applied to solve problems arising in antennas design and propagation modeling, among them:

1) integral methods (moment methods, physical optics and modal expansion);

2) differential equation methods (FDTD and Parabolic Equation (PE));

3) ray methods (tracing/launching and Gaussian beams).

A great number of Temporary Documents (TDs) reported during the Action have been dedicated to applicative aspects and comparative overviews of these methods including such issues as their database accuracy sensitivity, computation time, diffuse scattering modeling capabilities and performance metrics. Their specific applications, their pros and cons, may be read in Chapter 4 of [3].

The major part of the investigations/applications in the frame of Action 273 as for deterministic methods is dedicated to the ray methods. Ray tracing provides several advantages over traditional prediction methods: accurate site-specific field strength prediction, possibility to incorporate 3D antenna patterns and to obtain wide-band characteristics such as direction of arrival and multipath time delay. In this paper, to illustrate the ray methods applications, more details are given on the results obtained in [5] which demonstrated the real need of more accurate propagation prediction. The authors of [5] quantified the advantage of accurate radio coverage predictions undertaking a case study for a UMTS radio network planning in a 5 km^2 area in Paris, France. The conclusion following from their results is that a conventional propagation model (as the well-known Hata-Okumura model or even its improved version COST231-Hata model) could

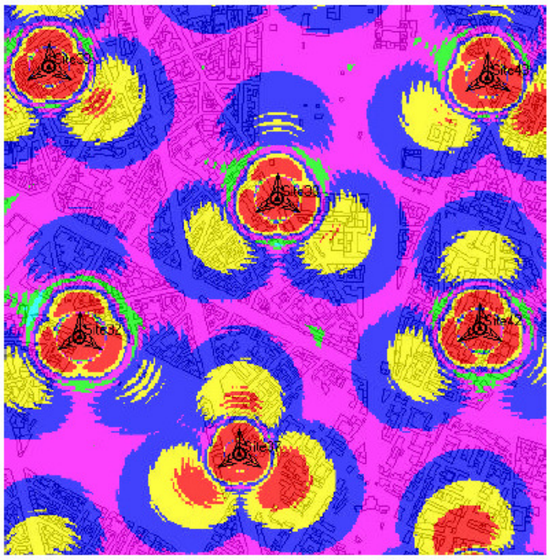


Fig. 1. Signal strength from empirical propagation predictions - COST231-Hata model. The building layout is not taken into account. (Fig. taken from [5])

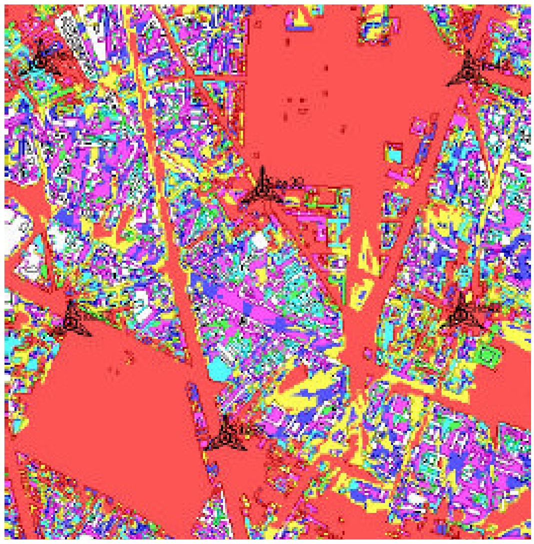


Fig. 2. Signal strength with ray tracing propagation predictions. Clearly seen are the canyon effect of streets and the impact of the buildings on the propagation. (Fig. taken from [5])

lead to erroneous planning with less than expected quality of service, unacceptable interference and building of more base stations than necessary, see Figs. 1 and 2. An accurate site-specific ray-tracing model, [6], integrated in the planning tool allows the radio network planner to reach optimal numbers for the base station deployment and configuration while meeting the expected service level requirements. One of the exemplary quantitative results of the case study for urban environment in [5] is that the ratio of rejected calls determined by the ray-tracing model is 14 times higher compared to the ratio derived from predictions using a simple COST 231-Hata-model.

The ray tracing based prediction techniques have some weaknesses: they are unreliable under specific tropospheric conditions (ducting), suffer difficulties when transiting from one type of region/cell (say, rural) to other (urban). Some of these difficulties could be overcome using even more complicated differential equation based methods as the PE method. The advantage of the PE method is that, as a full-wave method, it accounts simultaneously and accurately for the wave diffraction, refraction and scattering propagation mechanisms. Its disadvantage in comparison to ray tracing is that the PE does not provide the information on angle of departure/arrival in a straightforward manner. The PE method application is described in details in [7].

As the channel characterization is concerned, the COST 273 participants dealt with “traditional” topics characterizing the variability of power received through the mobile radio channel such as:

1) path loss, its distance dependence and small- and large-scale fluctuations;
2) temporal and angular dispersion;
3) fading prediction techniques;
4) experimental channel characterization techniques;
5) statistical processing of measurements;

but in this area special attention was devoted to the MIMO channel model and the reference scenarios for radio network simulation and evaluation developed in Action 273. The specification of antenna arrays at both link ends by setting the number of antenna elements, their geometrical configuration and polarizations turns the propagation model into a MIMO

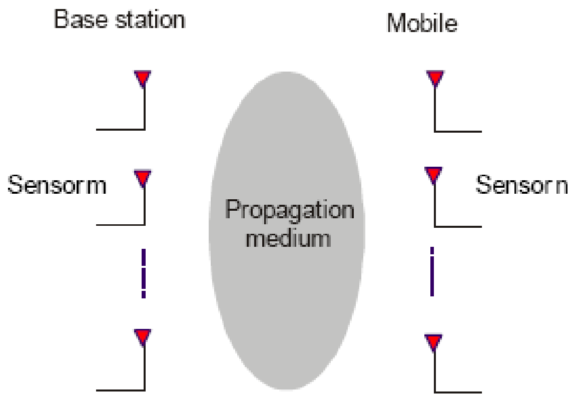


Fig. 3. Example of MIMO configuration (Fig. taken from [8])

channel model, see Fig 3. In order to make use of MIMO advantages, the space-time behaviour of the channel should be modeled. In general, from propagation point of view, a MIMO model has to meet the following requirements (as stated by COST 273):

- separation of the environment from the antenna arrays;
- adaptation to different environments;
- modeling of mobile movement;
- agreement with measurements;
- compliance with former models;
- precise, but by as few parameters as possible, analytical description of the spatial structure of the channel.

A very good example of fulfilment of these criteria may be found in [8], where the space-time behavior of the propagation channel is simulated on a ray-basis.

The reader may have noted the existence of discrepancy between the great efforts devoted to the development of new MIMO models during the last years and the almost lack of their validation. Still, a lot of work on models validation has been done within COST 273 and its novel MIMO model is mainly based on numerous and profound measurement campaigns reported and discussed during the Action. Its model is geometry based on the one hand and stochastic on the other, combining the best of both approaches, see Chapter 6 of [3] and report [9]. Nevertheless, it is to note that this novel concept is still to be validated as a whole.

In COST 273 special attention was paid also to the term *scenario* due to the great diversity in its use. In general, the term reflects a consistent set of parameters. At the propagation level (of which we are concerned in this paper), COST 273 presents a large variety of types of *scenarios*, reflecting the physical environment: one speaks of outdoor macro-, micro- and pico-cellular scenarios with or without line-of-sight, indoor office scenario, etc. In many cases the results of different researchers (both theoretical and experimental) appeared not to be comparable due to the different assumptions and approaches when describing the scenarios. This problem becomes evident especially when dealing with comparison of different RRM techniques and RNP strategies. Finally, a *Spatial Propagation Scenario* in COST 273 was introduced as a set of parameters, defining the spatial and multiuser situation, and the environment, with direct consequences on the propagation channel properties. A scenario like this is conceived to be used by several inter-working parties, in order to be able to compare performance, or to extract conclusions of how the propagation channel affects such performance (the classification of scenarios still needs to be motivated). To cope with scenario related problems, a new initiative raised within COST 273 SWG 3.1: the Mobile Radio Access Network Reference Scenarios (MORANS), which is oriented to provide a set of scenarios, parameters and models to be used as a common simulation platform. Due to the complexity of the Reference Scenarios reflecting the approaches used when evaluating RNP or RRM strategies, it is agreed that two types of scenario elements will be provided: synthetic scenarios, based on simple and regular geometrical lay-outs and simplified models, which make easy the interpretation of the results, and real-world based scenarios, where some data is taken from the real world, thus making their use more complex but providing the possibility to test radio network algorithms under more realistic conditions. In order to be practically relevant, the Reference Scenarios have to satisfy a number of requirements:

- ease of use;
- limited number of cases;
- defined interfaces between different parts of the scenario;
- extendibility.

Finally, it is clear that the success of further research requires a deeper integration of skills among the various disciplines involved in the design and development of mobile communication systems: for instance, the signal processing issues are closely related to propagation issues. In the future, investigations of the inter layers interactions will be needed. Some prospective topics are related to the advanced planning (Hybrid Networks - radio network architecture and automatic planning methods for GSM/UMTS). The creation and maintenance of reference scenarios will continue and their definition will involve the participation of physical layer, propagation, and networking experts. More investigations will be dedicated also to the Ultra Wide Band communications and MIMO communications.

More details on COST 273 may be found on http://www.lx.it.pt/cost273.

IV. Conclusion

This paper has summarized a small part of the huge number of topics the COST 273 participants dealt with. Demonstrated is the need of precise propagation modeling and channel characterization. These two topics have been chosen because of their importance: all the rest of signal processing, coding and deployment hinges on the accurate modeling.

References

[1] http://www.cost.esf.org/

[2] Final report. http://www.lx.it.pt/cost273.

[3] L. M. Correia (Ed.), *Mobile Broadband Multimedia Networks (Techniques, Models and Tools for 4G)*, Elsevier, 2006, 600 p.

[4] Memorandum of understanding and Technical Annex, http://www.lx.it.pt/cost273.

[5] M. Coinchon, A.-P. Salovaara, J.-F., Wagen, “The impact of radio propagation predictions on urban UMTS planning”, *COST TD(01)041*, http://www.lx.it.pt/cost273.

[6] WaveSight. http://www.wavecall.com

[7] I. Sirkova, “Overview of COST 273 Part II: Parabolic equation method application”, *XLI ICEST*, Sofia, Bulgaria, 2006.

[8] J.-M. Conrat, P. Pajusco, “A Versatile Propagation Channel Simulator for MIMO Link Level Simulation”, *COST 273 TD(03)120,* http://www.lx.it.pt/cost273.

[9] H. Hofstetter, G. Steinbock, “A Geometry based Stochastic Channel Model for MIMO Systems”, *ITG Workshop on Smart Antennas*, Munich, Germany, March 2004.

Overview of COST 273 Part II: Parabolic equation method application

Irina D. Sirkova[1]

Abstract **– This paper summarizes the application of the parabolic equation method in COST 273 activities. The parabolic equation method has been applied to solve microwave propagation prediction problems over irregular terrain and in forest environment.**

Keywords **– Microwave propagation prediction, parabolic equation, Digital Terrain Elevation Data, forest.**

I. INTRODUCTION

The Parabolic Equation (PE) method is one of the deterministic propagation methods used in COST 273 activities. It is based on a paraxial approximation to the wave equation. The PE is a full-wave method accounting simultaneously and accurately for the wave diffraction, refraction and scattering propagation mechanisms. The advantage of the PE in comparison to the wave equation is that, combined with a numerical technique, it can be easily marched in range provided the field is known on an initial plane and adequate boundary conditions on the scattering objects and at the outer boundaries of the integration domain are given. This has turned the PE method into one of the most widely used propagation prediction techniques for large classes of wave propagation problems.

In fulfilment of the IE-BAS commitment to the COST 273 initiative, that was the investigation of microwave propagation in strongly varying environment, the 2D version of this method has been applied firstly to the study of the influence of tropospheric ducting phenomenon on microwave propagation, [1], [2]. Indeed, ducting layers of different nature are present significant percentage of the time all over the world, especially in coastal zones, and hence being able to compute the propagation in those circumstances is important. The following investigations using PE method have been made: study of the surface-based ducts influence in short ranges (up to 3 km) [1], study of the path loss changes provoked by changes in duct parameters including range dependent ducting (corresponding to the refractivity profiles along a mixed land-sea path) [2], assessment of the combined effect of terrain and ducting on path losses at UMTS frequencies [3]. These Temporary Documents (TDs) have been summarized in [4].

This paper is based essentially on TDs [5] and [6]. The first TD reports results on the combination of the PE method with digital terrain data used when path loss is calculated in an open area. The second deals with wave propagation over a hilly terrain containing a forest edge in its end. In the next Section these TDs PE method is very briefly summarized.

[1]Irina D. Sirkova is with the Institute of electronics – BAS, 72 blvd. "Tzarigradsko chaussee", 1784 Sofia, Bulgaria, E-mail: irina@ie.bas.bg

II. PE METHOD DESCRIPTION

The standard 2D narrow-angle forward propagating parabolic equation is given by Eq. (1), [7]:

$$\frac{\partial u(x,z)}{\partial x} = \frac{i}{2k}\frac{\partial^2 u(x,z)}{\partial z^2} + \frac{ik}{2}\left(n^2(x,z)-1\right)u(x,z), \qquad (1)$$

where azimuthal symmetry is assumed, k is the free-space wave number, n is the refractive index of the troposphere, $u(x,z)$ is the reduced function related to an electromagnetic field component Ψ as: $\Psi(x,z) = u(x,z)exp(ikx)/(kx)^{1/2}$. To solve Eq. (1) different numerical techniques are applied. For elevated terrain study in [5] we made essentially use of the split-step Fourier technique, [7]. This technique searches the solution of the Fourier transform of Eq. (1) where the transform is performed while treating the term ($n^2(x,z)$-1) as constant:

$$U(x,p) = \Im\{u(x,z)\} \equiv \int_{-\infty}^{\infty} u(x,z)e^{-ipz}dz, \qquad (2)$$

$$U(x+\Delta x,p) = U(x,p)e^{i(-p^2/2k + kM10^{-6})\Delta x}, \qquad (3)$$

where Eq. (2) defines the transform with the transform variable p referred to as the vertical wave number and the solution of the transformed equation at $x+\Delta x$ in terms of the solution at x is given by Eq. (3). Then the inverse transform is applied to Eq. (3) to obtain $u(x+\Delta x,z)$:

$$u(x+\Delta x,z) = e^{ikM(z)10^{-6}\Delta x}\Im^{-1}\left\{e^{-j(p^2/2k)\Delta x}\Im\{u(x,z)\}\right\}. \qquad (4)$$

Equations (3)-(4) have been obtained under the assumption that the refractive index varies only with height and under the Earth-flattening concept, [8], that is, $n(z)$ in Eq. (1) is replaced by the modified refractive index $m(z)$ ($m(z)=n(z)+z/a_e$, with a_e - the Earth radius; $M=(m-1)10^6$ is the modified refractivity).

The narrow-angle PE given by Eq. (1) is very accurate at propagation angles within $\pm15^0$ of the preferred direction of propagation, [9]. Equation (1) holds for both (horizontal and vertical) polarizations, the difference between them being contained in the boundary conditions at the Earth surface. Perfect conductivity of the ground has been assumed in this paper. The initial field required to start the calculation procedure is provided by horizontally polarized Gaussian beam source with pattern factor:

$$F(\theta) = \exp\left[\frac{\ln(0.707)(\theta-\theta_s)^2}{(\theta_0/2)^2}\right], \tag{5}$$

where θ_0 and θ_s are the half power beamwidth and the antenna elevation angle.

To model the vertical profile $M(z)$ of the modified refractivity the standard troposphere is used. The standard troposphere is modeled by a modified refractivity gradient that increases monotonically at a rate of 0.118 M-units per m. The results are presented in the form of path loss (PL, in dB):

$$PL = 20\log\left(\frac{4\pi r}{\lambda}\right) - PF, \tag{6}$$

where λ is the free-space wavelength, r is the distance between the corresponding points and PF is the pattern propagation factor as defined in [8].

The authors of TD [6] have applied the finite difference (FD) numerical technique to solve Eq. (1). This technique may be found in [9] and will not be given here.

III. Results and Discussion

Relying on the accuracy of the PE method and its ability to provide quantitative field assessment, TD [5] compares results of path loss obtained using PE in combination with two Digital Terrain Elevation Data (DTED) collections for terrain description. The examples are for a specific region in the plateau of Sofia. The two terrain data collections we made use of are those of the DTED Level 0 provided by USA National Imagery and Mapping Agency (NIMA). These are standardized digital datasets, the one providing the minimal heights, the other containing the maximal heights, taken in 30 arc second square areas (nominally one kilometre). These data are publicly available and give near world wide coverage. The used polarization is horizontal. Antenna half power beamwidth θ_0=2^0 is used with different antenna tilt. The other parameters of the problem concerning path loss calculation are: frequency f=2 GHz, transmitter height h_t=30 m and receiver situated in the first two meters above the ground.

Fig. 1 shows the region of Sofia plateau with the surrounding mountains, the red line indicating one of the studied paths. Below are seen the terrain elevation profiles for this path as obtained from the DTED Level 0 maximal and minimal heights datasets, respectively, with the two Fresnel zones indicated on the upper profile. The difference between the two elevation profiles is not only in the absolute height values but, especially, in the shadowed regions (the red parts of the curves). Figs. 2-4 show path loss for θ_s=-3^0, θ_s=-2^0 and θ_s=-1^0, respectively, obtained from the two DTED datasets. As it is seen from these figures, in the vicinity of the path loss shadow peaks obtained from the DTED maximal heights dataset there is drastic difference between the two path loss

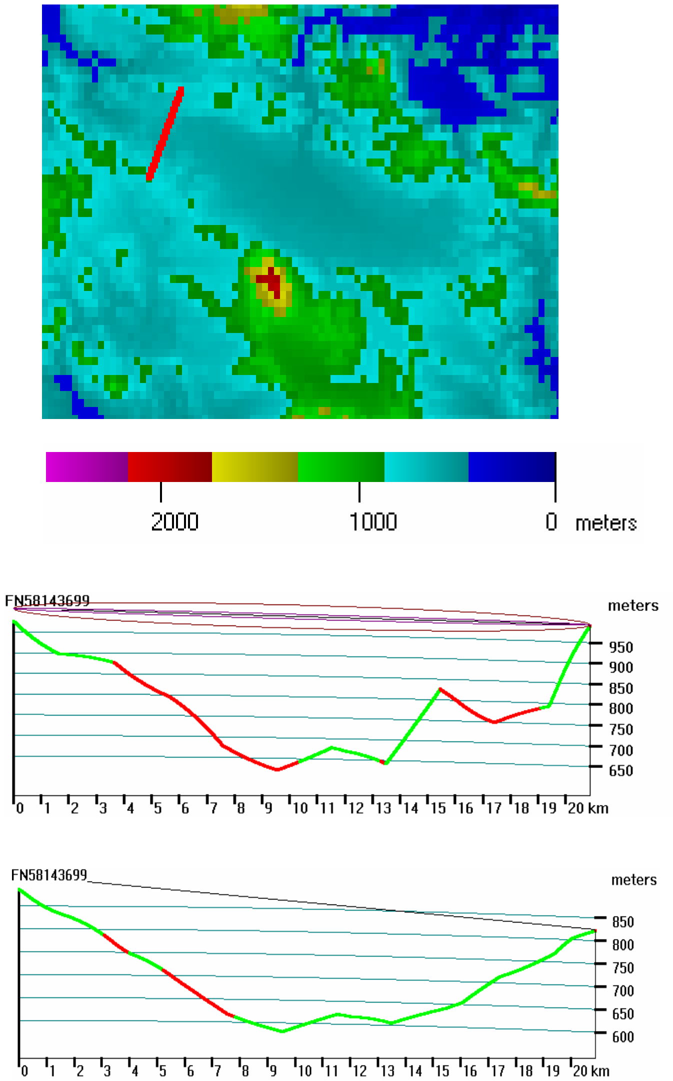


Fig. 1. Region of Sofia with the studied path loss indicated and the two terrain profiles for it obtained from DTED maximal (upper) and minimal (lower) heights dataset

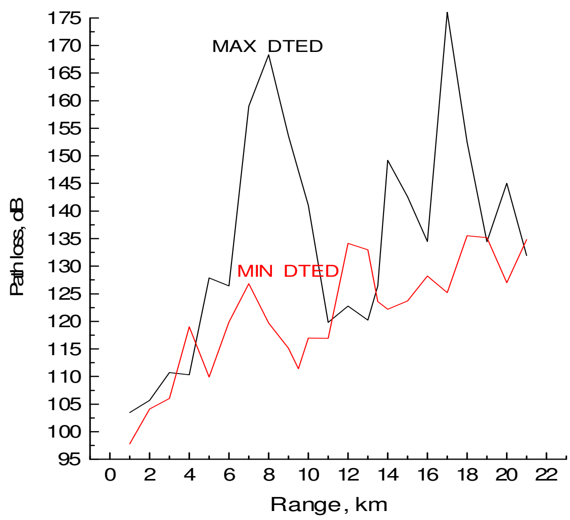


Fig. 2. Antenna tilt $\theta_s = -3^0$

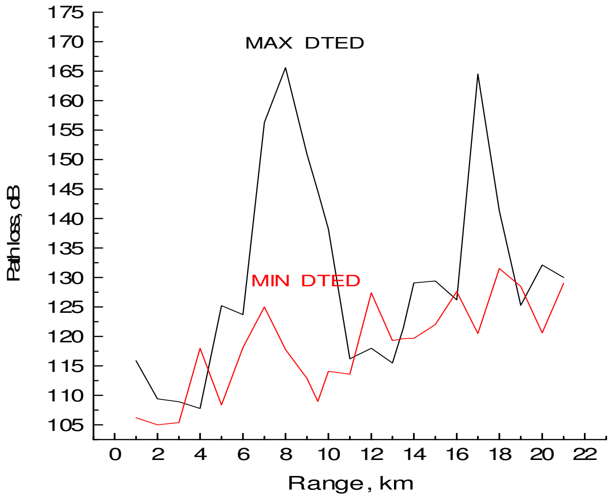


Fig. 3. Antenna tilt $\theta_s = -2^0$

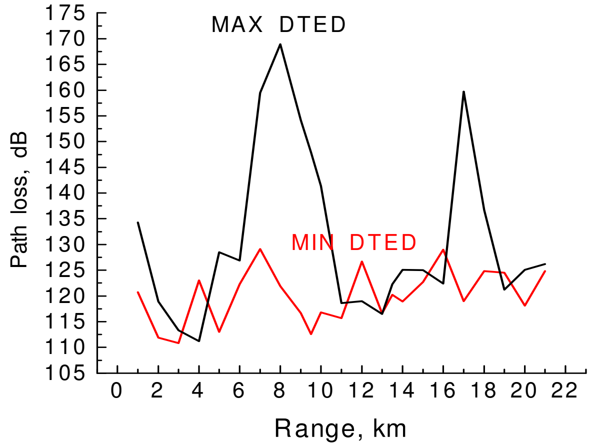


Fig. 4. Antenna tilt $\theta_s = -1^0$

curves. This difference can not be overcome with changes in the antenna tilt. It is not possible to obtain such quantitative results for path loss in the shadowed regions with empirical or simple deterministic propagation methods. The PE method provides the possibility to make an accurate assessment/prediction of the field strength in an early stage, when the coverage is calculated, thus giving information of power needed and possible interference.

In [6] the 2D PE method combined with FD numerical scheme is applied to model the propagation over irregular terrain partly covered by fir forest. The PE based results are compared to the Geometric Theory of Diffraction (GTD) technique and to measurements. The geometry and parameters of the problem are shown on Figs. 5 and 6. Measurements were made with the antennas located in the open area at three different distances behind the forest: 9 m, 51 m, and 109 m, see Fig. 5. For each of these antenna positions, the antenna heights were varied between 6 and 25 m in steps of about 2 m. When applying PE method, Leontovitch boundary condition for the ground was used and the forest was modeled as a dielectric slab characterized by height (18 m) and complex dielectric constant, i.e., relative permittivity ε_r (1.004) and conductivity σ (180 μS/m.). For the GTD model, the forest is modeled as two wedges with the interior angle 90°, see also Fig. 6. Both the wedges and the ground are modeled as perfectly conducting. This means that, for the PE model, some waves will penetrate through the forest while for the GTD model, the forest is a nonpenetrable object. These differences in applying the two models may partially explain the differences in transmission loss obtained through them and shown on Figs. 7 and 8.

The results for horizontally polarised links at 1.3555 and 1.5995 GHz obtained in TD [6] (two examples of them reported on Figs. 7 and 8) show that transmission loss obtained by GTD approximation is in general up to 20 dB worse that the measured cases for low receiver heights. The 2D PE solution agrees in general within ± 5 dB with the measured values. These results indicate also that the main contribution to the field near the forest is transmitted through the forest rather than diffracted over it. However, for larger distances from the forest edge the diffracted field gives the main contribution and the PE method and GTD give similar results.

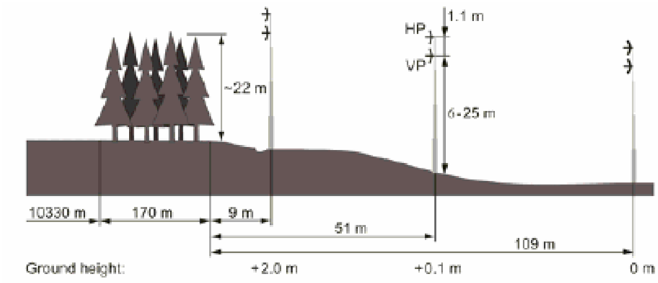


Fig. 5. The receiver antenna locations. (Fig. taken from [6])

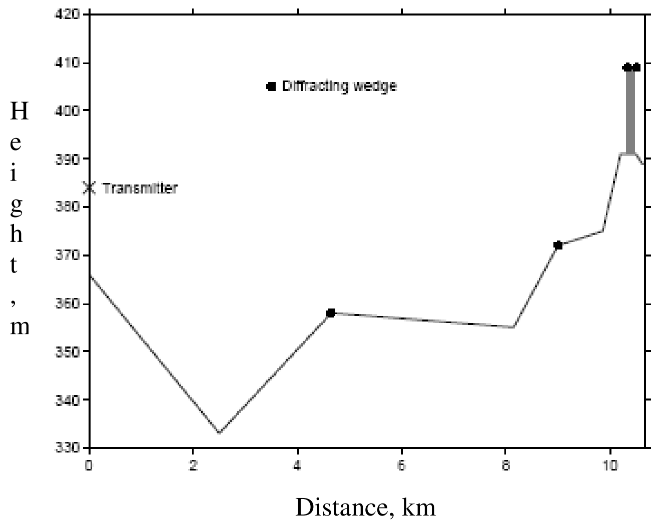


Fig. 6. Terrain height profile used for the PE and GTD calculations. The dots show the choice of the diffracting wedges in the GTD model. The forest is the gray area at the end of the path. (Fig. taken from [6])

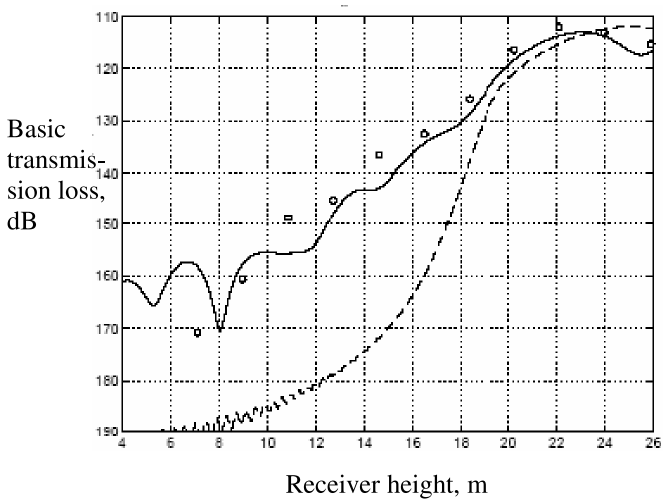


Fig. 7. Basic transmission loss vs receiver height for the frequency 1355.5 MHz and at the distance 9 m from the forest edge. The solid line shows calculations with the PE method, the dashed line is for GTD, the measurements are marked with o. (Fig. taken from [6])

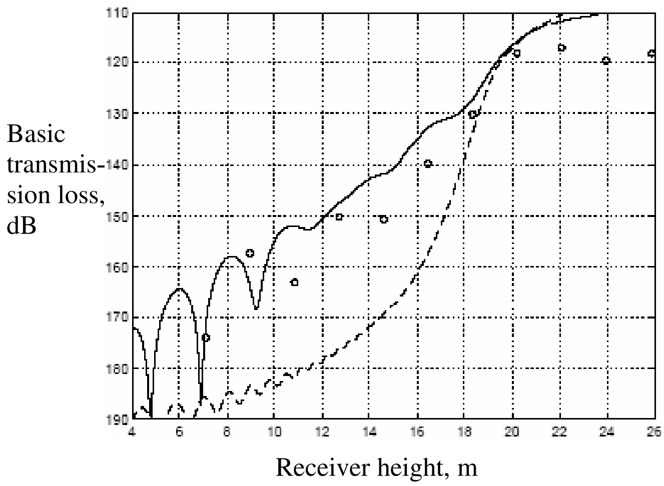


Fig. 8. The same as in Fig. 7 except that the frequency is 1599.5 MHz. (Fig. taken from [6])

It is to note also that another participant in COST 273, the AWE Communications GmbH, which is a spin off from the Stuttgart University, has already implemented the PE method among the other widely used propagation prediction methods in its planning tool, WinProp, [10]. The PE method is used especially for accurate propagation assessment in rural and suburban areas.

IV. Conclusion

This paper presents results obtained in COST 273 with the parabolic equation method applied to the prediction of the microwave propagation in complicated environments. Demonstrated is the ability of the PE method to provide accurate quantitative field assessment and its applicability to some difficult to solve problems as propagation through forest.

Acknowledgements

The author is grateful to the NIMA of USA for providing the DTED Level 0.

References

[1] I. Sirkova, M. Mikhalev, “Influence of Tropospheric Ducting on Microwave Propagation in Short Distances”, *COST 273 TD (02)086*, http://www.lx.it.pt/cost273.
[2] I. Sirkova, M. Mikhalev, “On the Influence of Tropospheric Duct Parameters on Microwave Radio Path Design”, *COST 273 TD(02)152*, http://www.lx.it.pt/cost273.
[3] I. Sirkova, M. Mikhalev, “PE Based Study of Path Loss Changes Due to Ducting Layers: Hilly Terrain Application”, *COST 273 TD (05)006*, http://www.lx.it.pt/cost273.
[4] I. Sirkova, M. Mikhalev, “Parabolic equation based study of ducting effects on microwave propagation”, *J. of Microw. Opt. Technol. Lett.*, Sept 2004, vol. 42, no. 5, pp. 390-394.
[5] I. Sirkova, M. Mikhalev, “Digital Terrain Elevation Data Combined with the PE Method: a Sofia Region Study”, *COST 273 TD(05)005,* http://www.lx.it.pt/cost273.
[6] A. Waern, P. Holm, “Wave propagation over a forest edge parabolic equation modelling vs. GTD modeling”, *COST 273 TD(03)169*, http://www.lx.it.pt/cost273.
[7] A. E. Barrios, “A Terrain Parabolic Equation Model for Propagation in the Troposphere,” *IEEE Trans.*, vol. AP-42, pp. 90–98, 1994.
[8] D. E. Kerr (Ed.), *Propagation of short radio waves*, Los Altos, CA, Peninsula Publishing, 1988.
[9] Levy, M., *Parabolic equation methods for electromagnetic wave propagation*, IEE electromagnetic waves series 45, UK, 2000.
[10] WinProp, http://www.awe-communications.com/.

Electromagnetic Field Calculation in the Time-Domain at Points outside the TLM Workspace

Nebojša Dončov, Bratislav Milovanović

Abstract – **In this paper, an approach for electromagnetic field calculation in the time-domain at points outside the defined numerical workspace is presented. This approach is primarily intended for electromagnetic compatibility problems analysis and allows for fast and accurate field calculation without resorting to discrete mesh which encloses these points. It is based on integration over Huygens surface surrounding all electromagnetic structures within the workspace. Approach is implemented in TLM numerical algorithm by means of a conveyor belt used for the introduction of the time delay in the near-to-near transform. Accuracy and efficiency of the approach are illustrated on the appropriate examples.**

Keywords – **electromagnetic compatibility, TLM, Huygens surface, time delay, output outside workspace**

I. INTRODUCTION

Electromagnetic compatibility (EMC) is the branch of science and engineering concerned with the design and operation of equipment in a manner which makes them immune to certain amounts of electromagnetic (EM) interference, while at the same time keeping equipment-generated interference within specific limits [1]. The scope of EMC is thus very wide as it encompasses virtually all equipment powered by electrical supplies. In recent years, the rapid increase in the use of radio communications, digital systems, fast processors and the introduction of new design practices have brought EMC to the forefront of advanced design.

How practical modern systems exceed in complexity anything that can be solved analytically or using approximate techniques, numerical computer-based models are the only alternative to study trends in EMC design and to understand the behaviour of systems. Differential numerical techniques in the time-domain, such as Finite-Difference Time-Domain (FD-TD) [2] and Transmission-Line Matrix (TLM) [3], are well established for solving a number of EMC problems over a wide frequency range. These methods, due to their characteristics and the development of powerful computer stations, offer a significant extension of the range of EMC problems that can be tackled.

However, there are still numerous practical EMC problems where these techniques, even with the use of computer stations of remarkable memory and run-time performances, are incapable to allow for their fast and correct modelling. The examples are geometrically small but electrically important features (such as thin wires, slots, air-vents, etc) in an otherwise large modelling space. For their description, extremely fine mesh is required which can result in a prohibitively large number of cells and large number of time steps. In recent years, enhancements to TLM technique in the form of so called sub-cell or compact models have been developed to allow for an efficient simulation of these structures [4-6]. These models take into account the EM presence of fine details without resorting to extremely fine mesh around them.

Level of equipment-generated interference is one of the key concerns in EMC design. Thus a complete analysis of EMC problems requires calculation of EM field response at points far away from the EM structures generating this response. Again, large number of cells is required to cover the distance between equipment and outputs, which can exceed the dimensions of EM structures by one order of magnitude or more. This leads to time consuming and inefficient EMC simulation even with the use of developed compact models.

The technique of EM field calculation at points located at large distances from the radiating EM structures, without defining discrete numerical mesh up to these points, is developed and presented in this paper. It is based on Love's equivalence principle [7,8] and requires definition of so-called Huygens closed surface within the workspace surrounding all EM structures of analyzed EMC problem. Integration over the Huygens surface, divided into elemental patches whose cross-section is determined by the resolution of applied numerical mesh, at every time step allows for calculating the time-domain EM field at points outside the numerical workspace. The time delay in near-to-near transform is introduced by means of a conveyor belt. Proposed scheme is implemented in TLM numerical algorithm and verified on appropriate examples. Compared to the conventional simulation, huge saving in simulation run-time can be achieved.

II. LOVE'S EQUIVALENCE PRINCIPLE

The surface equivalence theorem is a principle by which actual sources, such as an antenna and transmitter, are replaced by equivalent sources. The fictitious sources are said to be equivalent within a region because they produce within that region the same field as the actual sources. According to this principle, the fields outside an imaginary closed surface are obtained by placing, over the closed surface, suitable electric and magnetic current densities that satisfy the boundary conditions. The current densities are selected so that the fields inside the closed surface are zero and outside are equal to the radiation produced by the actual sources. Thus this technique can be used to obtain the fields radiated outside

Nebojša Dončov and Bratislav Milovanović are with the Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia and Montenegro, E-mail: [doncov, bata]@elfak.ni.ac.yu

a closed surface by sources enclosed within it. The formulation is exact but requires integration over the closed surface.

The surface equivalent theorem can be illustrated by considering an actual radiating source, which is represented electrically by current densities J_1 and M_1, as shown in Fig.1. The source radiates fields E_1 and H_1 everywhere. In order to develop a method that will yield the fields outside a closed surface as actual sources, a closed surface S has to be chosen (dashed lines in Fig.1a) which encloses the current densities J_1 and M_1. The volume within S is denoted by V_1 and outside S by V_2. The primary task is to replace the original problem shown in Fig.1a with an equivalent that will yield the same field E_1 and H_1 outside S (within V_2)

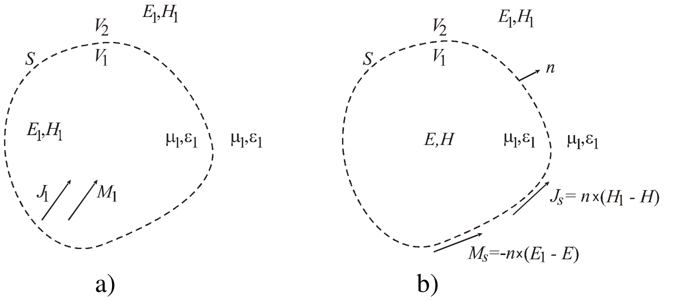


Fig. 1. a) Actual and b) equivalent problem models

An equivalent problem to Fig.1a is shown in Fig1.b. The original sources J_1 and M_1 are removed, and one can assume that there exist fields E and H inside S and fields E_1 and H_1 outside S. For these fields to exist within and outside S, they must satisfy the boundary conditions on the tangential electric and magnetic field components. Thus on the imaginary surface S the following equivalent sources must exist:

$$J_s = n\times(H_1 - H) \tag{1}$$

$$M_s = -n\times(E_1 - E) \tag{2}$$

which radiate into an unbounded space (same medium everywhere). The current densities J_s and M_s are said to be equivalent only within V_2, because they will produce the original fields (E_1 and H_1) only outside S. A fields (E and H) different from the original (E_1 and H_1) will result within V_1. Since the fields (E and H) within S, which is not the region of interest, can be anything, it can be assumed that they are zero. Then the equivalent current densities are equal to:

$$J_s = n\times(H_1 - H)\big|_{H=0} = n\times H_1 \tag{3}$$

$$M_s = -n\times(E_1 - E)\big|_{E=0} = -n\times E_1 \tag{4}$$

This form of the field equivalence principle is known as Love's equivalence principle [7,8]. Since the current densities J_s and M_s radiate in an unbounded medium that is, have the medium (μ_1,ε_1) everywhere, they can be used in conjunction with the Eqs.(5) and (6) to find the fields everywhere:

$$E_1 = -j\frac{1}{\omega\mu_1\varepsilon_1}\nabla(\nabla\cdot A) - j\omega A - \frac{1}{\varepsilon_1}\nabla\times F \tag{5}$$

$$H_1 = -j\frac{1}{\omega\mu_1\varepsilon_1}\nabla(\nabla\cdot F) - j\omega F + \frac{1}{\mu_1}\nabla\times A \tag{6}$$

where A and F are vector magnatic and electric potential, respectively, that can be expressed through electric and magnetic surface currents as [8]:

$$A = \frac{\mu_1}{4\pi}\int_S J_s \frac{e^{-jkR}}{R}dS',\ F = \frac{\varepsilon_1}{4\pi}\int_S M_s \frac{e^{-jkR}}{R}dS' \tag{7}$$

III. Near-to-near Transform

The contribution of an elemental patch of the Huygens surface to EM field at a point removed from it by the vector R can be found in the time-domain from Eqs.(5-7) as:

$$\Delta E(t+R/c) = \frac{\Delta S}{4\pi R\cdot R}\left(\Delta E_{rad}(t) + \Delta E_{ind}(t) + \Delta E_{stat}(t)\right) \tag{8}$$

$$\Delta H(t+R/c) = \frac{\Delta S}{4\pi R\cdot R}\left(\Delta H_{rad}(t) + \Delta H_{ind}(t) + \Delta H_{stat}(t)\right) \tag{9}$$

where ΔS is the area of the elemental patch determined by TLM node. As it can be seen the signal from each surface patch consists of a differential radiation term (ΔE_{rad} and ΔH_{rad}), a direct induction term (ΔE_{ind} and ΔH_{ind}) and an integral static field term (ΔE_{stat} and ΔH_{stat}) that can be expressed in rectangular coordinate system as:

$$\Delta E_{rad}(t) = \frac{|R|}{c}\frac{d}{dt}\left(Z_0((J_s\cdot R_{ort})\cdot R_{ort} - J_s) - M_s\times R_{ort}\right) \tag{10}$$

$$\Delta E_{ind}(t) = Z_0(3(J_s\cdot R_{ort})\cdot R_{ort} - J_s) - M_s\times R_{ort} \tag{11}$$

$$\Delta E_{stat}(t) = \frac{c}{|R|}\int Z_0(3(J_s\cdot R_{ort})\cdot R_{ort} - J_s)dt \tag{12}$$

$$\Delta H_{rad}(t) = \frac{|R|}{c}\frac{d}{dt}\left(Y_0((M_s\cdot R_{ort})\cdot R_{ort} - M_s) + J_s\times R_{ort}\right) \tag{13}$$

$$\Delta H_{ind}(t) = Y_0(3(M_s\cdot R_{ort})\cdot R_{ort} - M_s) + J_s\times R_{ort} \tag{14}$$

$$\Delta H_{stat}(t) = \frac{c}{R}\int Y_0(3(M_s\cdot R_{ort})\cdot R_{ort} - M_s)dt \tag{15}$$

where Z_0 is an intrinsic impedance of free space ($Y_0 = 1/Z_0$), R_{ort} *is* unit vector ($R_{ort} = R/|R|$) and J_s and M_s are the equivalent electric and magnetic current densities, respectively.

The time delay R/c in the near-to-near transform is introduced by means of a conveyor belt shown in Fig.2. Each output point is at the delivery end of its own conveyor which carries a time-domain signal towards it at the speed of light. In order to introduce a delay R/c into a signal, the signal is dropped onto the conveyor at the distance R upstream of the output point.

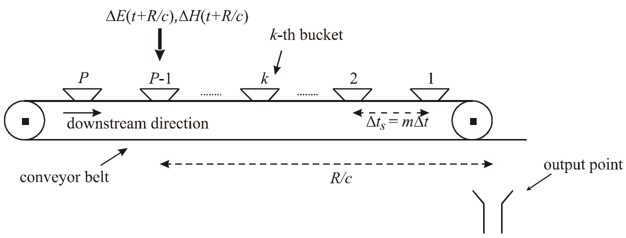


Fig. 2. Conveyor belt for one output point

The conveyors are not continuous but consist of buckets spaced one time sample, Δt_s, apart. Time sample depends on highest frequency of interest for EM response analysis and in general can be expressed as integer number, m, of used time step in TLM simulation, Δt ($\Delta t_s = m\Delta t$). If a signal is dropped onto the conveyor belt at time, $i\Delta t$, (i=1,2,...,I_{max}, where I_{max} is maximum number of iteration of simulation in the time-domain) when a bucket is not passing the loading point in question, then the signal falls between buckets and is lost. The conveyor belt and buckets on it advance with step Δt towards output point (downstream direction). The time delay R/c is introduced in the way that signal corresponding to the appropriate elemental patch of Huygens surface at distance R from the output point is dropped onto appropriate place at conveyor belt ($\Delta E(t+R/c)$ and $\Delta H(t+R/c)$, given by Eqs.(8-15)). The total length of a conveyor is determined by the maximum required delay, i.e. by the distance between the output point and the most remote patch of the Huygens surface and it can be calculated as:

$$P_{\max} = 1 + \mathrm{int}\left(R_{\max} / c / \Delta t\right) \qquad (16)$$

a) Differential radiation term

The differential term is generated by dropping an up-down pulse onto the conveyor. Specifically, if a signal of $dv(t)/dt$ is required, then at each time step a value of $+v/dt$ is dropped onto the conveyor at one point, and a value of $-v/dt$ is dropped at a point one time step upstream of the first. At most one of these values will actually land in a bucket. The up-pulse at one time step will always land at the same place on the conveyor as the down-pulse at the previous time step (the conveyor having advanced) so that bucket will catch: $(+v(t+\Delta t)-v(t))/\Delta t$. Because the desired delay may not be an exact half-integral number of time steps, the loading points for the up- and down- pulses may be misaligned with the conveyor. To handle this, each of the up- and down- pulses may be shared between two adjacent locations on the conveyor, so that the entire up-down pulse is spread over a three time steps length (at least 2/3 of this falling between buckets). If the desired delay requires a loading point, n time steps from the delivery end of the conveyor, and if N=nint(n) is the nearest integer value, then weightings are:

$$\text{drop}[N\text{-}1] = \left(N + 1/2 - n\right)v(t)/\Delta t$$
$$\text{drop}[N] = 2\left(n - N\right)v(t)/\Delta t$$
$$\text{drop}[N\text{+}1] = \left(N - 1/2 - n\right)v(t)/\Delta t$$

b) Direct term

The value for the direct induction term is simply dropped onto the conveyor at the desired point. However, since this point may not a non-integral number of time steps from the output point, the value may need to be shared between two adjacent locations on the conveyor. If the desired delay requires a loading point n time steps from the delivery end of the conveyor, and $N1$ and $N2$ are consecutive integers bracketing n, then weightings are:

$$\text{drop}[N1] = \left(N2 - n\right)v(t)$$
$$\text{drop}[N2] = \left(n - N1\right)v(t)$$

At most one of these values will actually land in a bucket.

c) Integral term

The integration for the static-field term is achieved by dropping the integrand (times Δt) onto all locations on the conveyor upstream of the desired delay point (most of these being between buckets). To achieve the effect of a conveyor that extends upstream indefinitely, the most-upstream bucket is initialized with the contents of the previously most-upstream bucket every time the conveyor advances by a whole time sample. Care must be taken that the most-upstream bucket is sufficiently far upstream to avoid contamination from the differential or direct terms emanating from the remotest part of the Huygens surface.

If the desired delay is n time steps, and n is not a half-integral number, then a weighting of

$$\text{drop}[N] = \left(N + 1/2 - n\right)v\left(t\right)\Delta t$$

may be applied to the nearest location N=nint(n) (should a bucket be passing this point at the time), with the full value $v(t)\Delta t$ being applied to all upstream locations.

IV. Numerical Analysis

The accuracy of scheme for calculating the time-domain EM field at points outside the TLM workspace is verified on the simple example of dipole antenna radiating in free space. Dipole antenna is represented as straight wire conductor of radius r=0.5 cm and length l=28 cm. Compact wire model [4] is used for wire modelling. Real voltage generator, 1V and 50 Ω, is placed at the centre of the wire. EM field time-domain response at two output points placed around dipole antenna is calculated by one TLM simulation and two different approaches: a) enclosing the output points by numerical mesh of TLM nodes and using conventional TLM algorithm of scattering and connection [2]; and b) applying integration over a Huygens surface enclosing the dipole antenna. After using discrete Fourier transformation, the EM field at these points in the frequency-domain is shown in Fig. 3. It can be noticed a good agreement between the results obtained by these two approaches.

The efficiency of proposed scheme is illustrated on the realistic EMC example in the form of rectangular shielding enclosure (Fig.4). The enclosure was constructed of five pieces of 0.635 cm thick aluminum, and one plate of 0.05 cm thick aluminum (face containing the slot). The inside dimensions of the enclosure were 22 cm × 14 cm × 30 cm. The feed probe was represented as a wire conductor of 0.16 cm diameter terminated at the bottom of the cavity by 47 Ω resistor. Simple voltage source, 1mV with 50 Ω resistance, was incorporated into the wire at the top of the cavity. Slot length and width were 12 cm and 0.1 cm, respectively. Compact wire [4] and slot model [5] were used to model wire conductor and slot structure in order to avoid the need of TLM mesh with extremely high resolution. The choice of geometry, excitation and output was governed by experimental arrangements used in [9].

Huygens virtual surface, completely surrounding the TLM model of the enclosure, is located at the distance of one TLM node from the enclosure. The external boundaries of TLM mesh are placed around the enclosure at the distance greater than 30% of the largest dimension of the enclosure, which is still far away from the required output point. The results for far zone electric field at 3 m away from the face of enclosure containing slot, obtained by using the proposed scheme, are compared with the experimental results [9] and shown in Fig.5. An excellent agreement between these two results can be observed. In addition huge saving in the computer resources, compared to conventional TLM, is achieved with accuracy acceptable for most EMC applications.

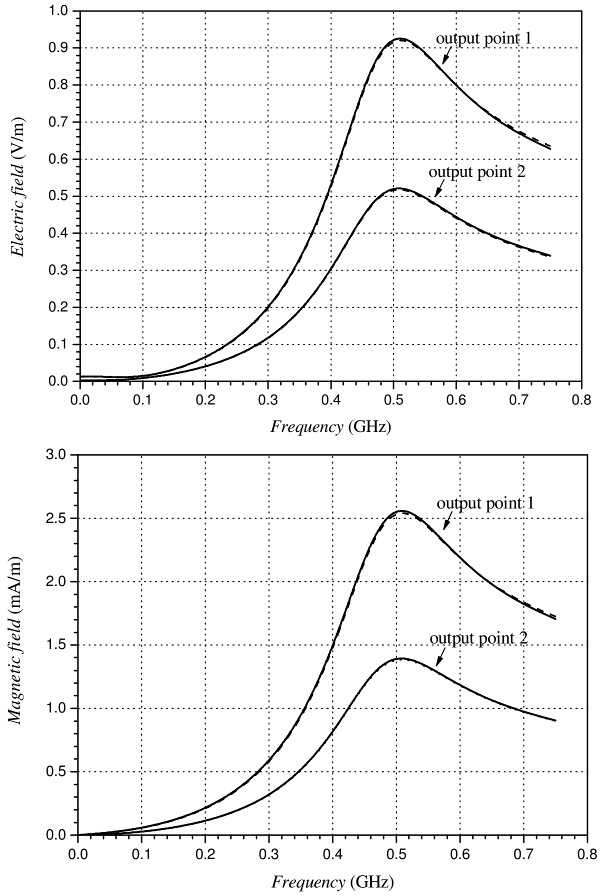


Fig. 3. EM field at output points obtained by: a) conventional TLM (dotted line), b) proposed scheme (solid line)

V. CONCLUSION

This paper describes an efficient scheme for calculating the time-domain EM field at points outside the TLM workspace. It is primarily indented for EMC applications. The proposed scheme is based on integration over a Huygens surface within the defined workspace using the calculated equivalent electric and magnetic surface currents. The effect of time delay in signal propagation from elemental patches of Huygens surface to output points is realized through appropriate conveyor belts. The scheme is general and it can be implemented into other differential numerical techniques in the time-domain such as FD-TD method. The efficiency of this approach is illustrated on the appropriate examples.

Because the integration over the Huygens surface involves a significant amount of computation, it may be worthwhile attempting to reduce the integration frequency from once every time step to once every time sample. Even if the computational burden for a single output point proves acceptable, it can be easily be made unacceptable by specifying sufficiently many output points. The modification of integration frequency that incurs the cost of extra storage, but has an advantage that this cost is fixed and it does not increase with the number of output points will be presented in future papers.

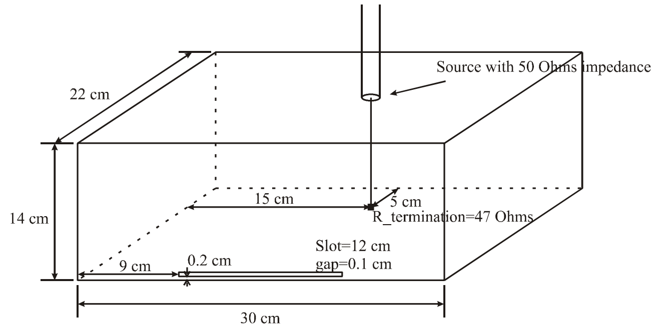


Fig. 4. Rectangular test enclosure

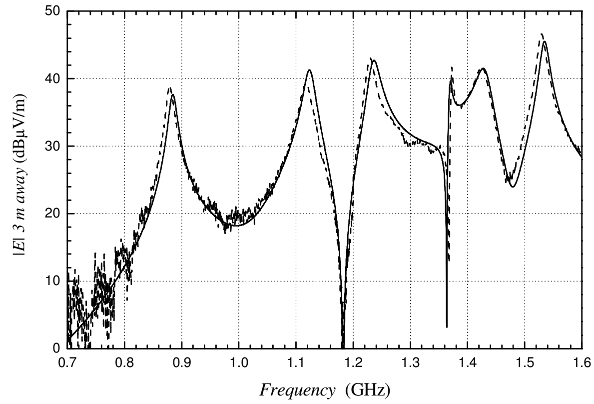


Fig. 5. Radiated electric field at 3 m away from the enclosure: a) measurements (dotted line), b) proposed scheme (solid line)

REFERENCES

[1] C. Christopoulos, *Principles and Techniques of Electromagnetic Compatibility*, CRC Press, 2000.

[2] C. Christopoulos, *The Transmission-Line Modelling Method*, Series on Electromagnetic Wave Theory, IEEE/OUP Press, 1995.

[3] K.S. Kunz, R.J. Luebbers, *The Finite Difference Time Domain for Electromagnetics*, CRC Press, 1993.

[4] A.J. Wlodarczyk, V. Trenkić, R. Scaramuzza, "*A Fully Integrated Multiconductor Model for TLM*", IEEE Transactions on Microwave Theory and Techniques, Vol.46, No.12, 1998, pp.2431-2437.

[5] V. Trenkić, R. Scaramuzza, "*Modelling of Arbitrary Slot Structures using Transmission Line Matrix (TLM) Method*", Proceedings of International Zurich Symposium on Electromagnetic Compatibility, Zurich, Switzerland, 2001, pp.393-396.

[6] N. Dončov, A.J. Wlodarczyk, V. Trenkić, R. Scaramuzza, "*Compact TLM Model for Air-vents*", Electronics Letters, Vol.38, No.16, 2002, pp.887-889.

[7] E.C. Jordan, K.G. Balmain, *Electromagnetic Waves and Radiating Systems*, Prentice-Hall Inc., 1968.

[8] C.A. Balanis, *Advanced Engineering Electromagnetics*, John Wiley & Sons, New York, 1989.

[9] M. Li, J. Nuebel, J.L. Drewniak, R.E. DuBroff, T.H. Hubing, T.P. Van Doren, "*EMI from Cavity Modes of Shielding Enclosures – FDTD Modelling and Measurements*", IEEE Transactions on Electromagnetic Compatibility, Vol.42, No.1, 2000, pp.29-38.

TLM Modelling and Analyse of Power Divider Using Linear Electric Probes Coupling Inside Cylindrical Cavity

Jugoslav Joković, Bratislav Milovanović

Abstract – **In this paper, power divider using linear electric probes coupling inside cylindrical cavity is analysed by using 3D TLM software in order to study effects of the feed probe and power dividing probes for determining the maximum power transfer for practical applications. TLM nunerical results of reflection coeficient and standing wave ratio (SWR) of feed probe and dividing probes for various probe length are compared with results obtained using Method of Moments and experimantal ones. In comparison with Method of Moments when probes are considered to be with negligible thickness, TLM results including of modelled probes with real radius shows better agreement with experimental ones.**

Keywords – **Power Divider, Cavity, Electric Probe, TLM Method, TLM Wire Node**

I. Introduction

Communications systems such as television, radio an mobile phone have been extensively and continuously used for distribution over wide of the service area [1]. The antenna that provides unidirectional beam pattern by using sectoral cylindrical cavity-backed slot antenna excited by probe is attractive candidate for usage [2]. When combining each element to be circular array, omni directional pattern are achieved.

TLM (Transmission Line Modelling) method is a general, electromagnetically based numerical method that has been applied successfully to the wide range of problems. For example, it has been applied to the modelling of metallic cavities [3]. By using a real feed probe for establishing desired field distribution in the modeled cavity, with some improvements in TLM method, it is possible to model a probe inside the cavity using TLM wire node [4] and to investigate the influence of the real feed to the resonant frequencies in the cavity [3, 5].

This paper focuses on the TLM modeling and analyses of characteristics for power divider using probes coupling inside cylindrical cavity [6]. The structure of the power divider is composed of conducting cylindrical cavity of the radius r and the height c. The feed probe is aligned along z direction that is located at the center of the cavity and the length probe is l_f as shown in the Fig. 1.

In this way, it is possible to excite modes having z-component of the electrical field in the cavity (TM modes). The length of the power dividing probes are l_{p1}, l_{p2}, l_{p3} and l_{p4}, respectively. They are aligned along r direction and located at the distance $c/2$ from the bottom of the cavity and the angle between the probe is 90 degrees apart.

Jugoslav Joković and Bratislav Milovanović are with the Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Nis, Serbia and Montenegro, E-mail: [jugoslav, bata] @elfak.ni.ac.yu

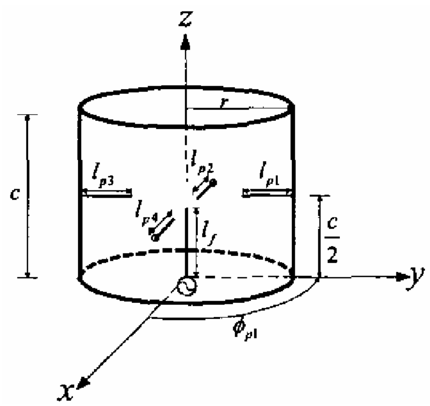


Fig. 1. The structure of power divider using linear electric probes inside cylindrical cavity

In order to investigate the possibilities and effectiveness of TLM method for modelling of power divider with complex geometries that have significant application in modern communication systems, obtained TLM numerical results of characteristics of power divider for various lengths of feed probe and power dividing probes are compared with results obtained using Method of Moments to solve the same problem and experimental results as well [6].

II. TLM Modelling Procedure

In the TLM time-domain method, an electromagnetic field strength in three dimensions, for a specified mode of oscillation in a cylindrical metallic cavity, is modelled by filling the field space with a network of link lines and exciting a particular field component [7]. Electromagnetic properties of mediums in the cavity are modelled by using a network of interconnected nodes (Fig. 2), a typical structure being the symmetrical condensed node (SCN) [8]. Each node describes a portion of the medium shaped like a cuboid or a slice of cake depending on the applied (rectangular/cylindrical) coordinate system (grid).

TLM wire node is based on SCN with one small modification in the form of additional link and stub lines interposed over the exiting network to account for increase of capacitance and inductance of the medium caused by wire presence [4]. This wire network is usually placed into the centre of the TLM nodes to allow modelling of complex wire structures, e.g. wire junctions and bends (Fig.3). The single column of TLM nodes, through which wire conductor passes, can be used to approximately form the fictitious cylinder which represents capacitance and inductance of wire per unit length. Its effective diameter, different for capacitance and inductance, can be expressed as a product of factors empirically obtained by using known characteristics of TLM network and the mean dimensions of the node cross-section in the direction of wire running.

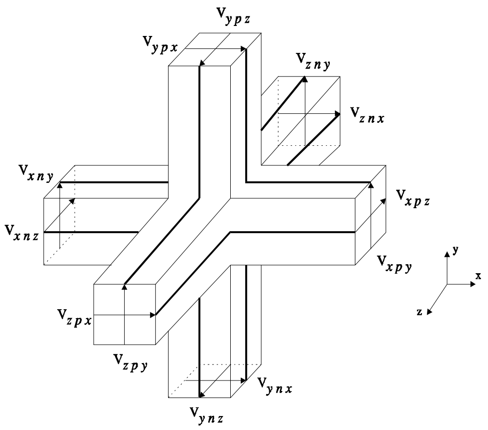


Fig. 2. Symmetrical condensed node

Requirement that the equivalent radius of fictitious cylinder is constant along nodes column can be easily met in a rectangular grid. However, in the cylindrical grid for wire conductor in the radial direction, mean cross-section dimensions of TLM nodes, through which wire passes, are changeable making difficult to preserve distributed capacitance and inductance of wire per unit length. Because of that, a rectangular grid has been chosen for modelling of cylindrical cavity analysed in this paper. At the same time, the numerical errors introduced by describing boundary surfaces of the modelling cavity in a step-wise fashion are reduced applying the TLM mesh higher resolution around cavity walls.

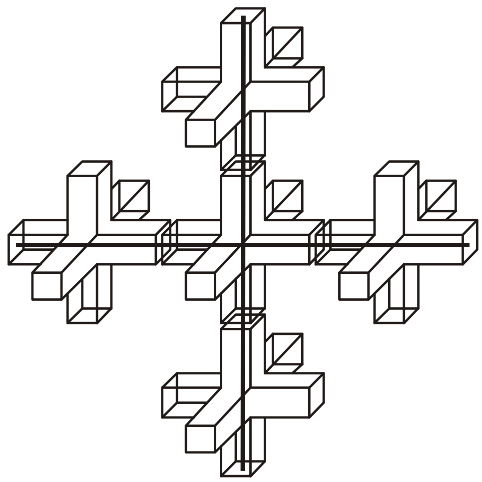
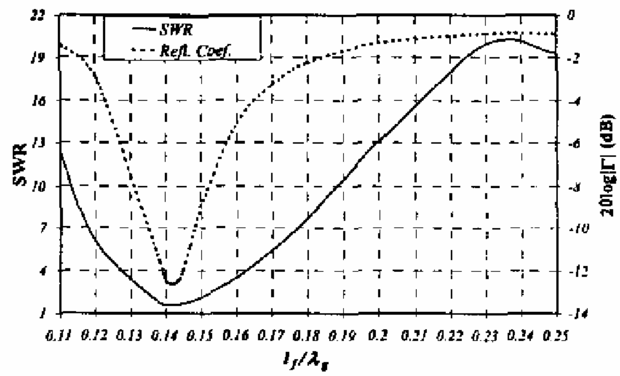


Fig. 3. Wire network embedded within the TLM nodes

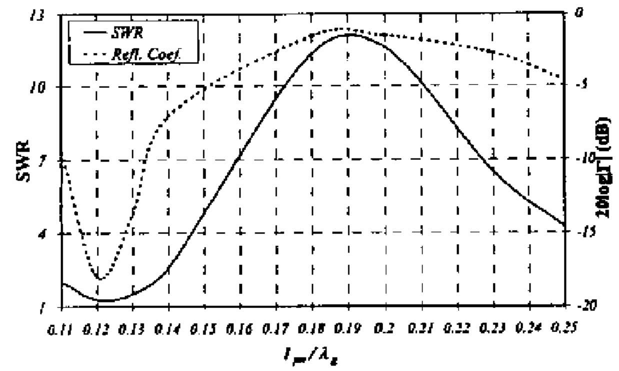
III. Method of Moments Results

In the case of Method of Moments applying for study of power divider, analysis is beginning with the integral equations of electric field intensities that can be formulated by applying Field Equivalence Principle together with boundary conditions at all the probes. The dyadic Green function that is the impulse response is derived and Method of Moments is used to solve for unknown current densities in the integral equation [6].

Numerical results based on Method of Moments of reflection coefficient and standing wave ratio (SWR) as a function of probe lengths is presented in the Fig.4. First, the length of feed probe is varied from $0.11\lambda_g$ to $0.25\lambda_g$, at the operating frequency of 2.4GHz. From Fig.4.a. the optimum SWR of 1.62 is realised when the length of feed probe is $0.14\lambda_g$, corresponding the level of reflection coefficient of -12.5dB. Furthermore, for this optimum value of length of feed probe, the length of dividing probes is also varied from $0.11\lambda_g$ to $0.25\lambda_g$. The best of SWR of 1.28 is occurred when the length of probes is $0.12\lambda_g$, the level of level of reflection coefficient of -18dB (Fig.4.b). These optimum values of probe lengths are used as the design parameter. In all calculations based on the applying of Meted of Moments the linear electric probes are considered to be perfect electric conductor and the thickness is negligible [6].



a)

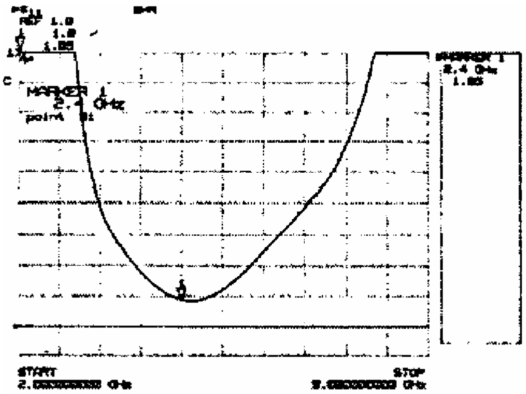


b)

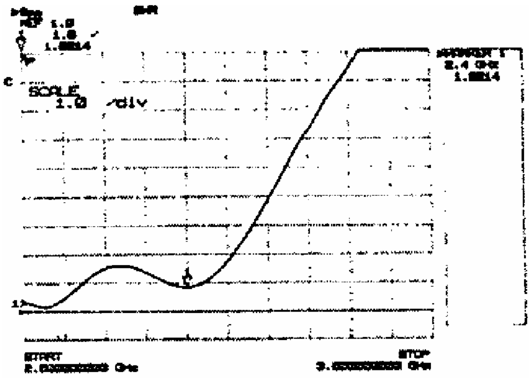
Fig. 4. SWR and reflection coefficient obtained using Method of Moments for: a) feed probe b) dividing probe

IV. Experiment

The prototype of power divider is designed based on TM mode excitation in conducting cylindrical cavity and the length of probes are chosen according to optimum condition as specified is previous section [6]. The Fig 5. illustrates measured SWR of feed probe and dividing probe. Optimum value of SWR on the operating frequency is 1.85 for feed probe, and 1.82 for dividing probe.



a)



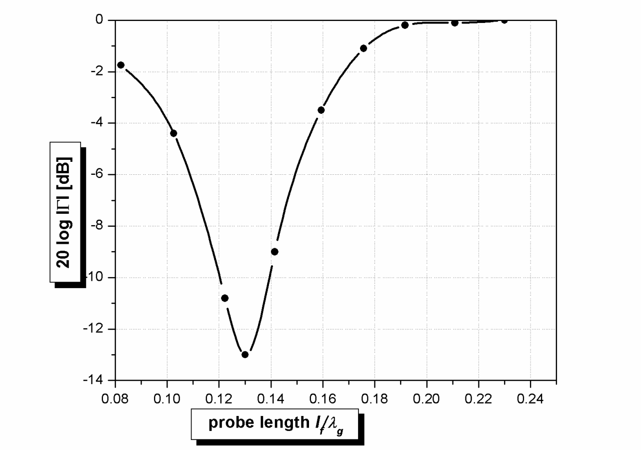
b)

Fig. 5. Measured SWR for:
a) feed probe b) dividing probe

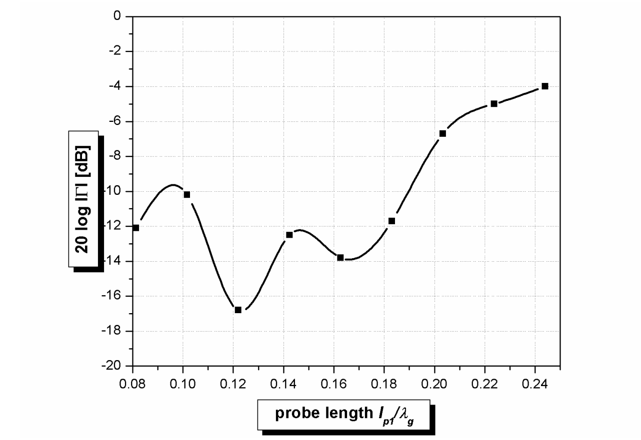
V. TLM NUMERICAL ANALYSIS

The power divider configuration shown in Fig 1 are analysed using 3D TLM software in order to investigate possibilities of TLM method for modelling of power divider structure with excitation in the form of wire probe loaded in the cavity and four dividing probes. Dimensions of the modelled cavity are $2r$=9.25cm and c=7.25cm. For cavity modelling, a non-uniform rectangular TLM mesh was used. The feed and dividing probes are modeled through TLM wire node. The radius of probes are chosen to be r_w =0.5mm, while length is varied to follow the results based on Method of Moments described in previous section.

The numerical results, which illustrate the effect of the feed probe and its length to the reflection coefficient, are presented in the Fig.6.a. The feed probe is used as a receiving probe as well, for resonant frequency detection from the reflection coefficient. From Fig.6.a the level of reflection coefficient of -13dB is realised when the length of feed probe is $0.13\lambda_g$. Corresponding SWR for the level of reflection coefficient of -13dB is 1.58. These results are slightly different from numerical ones based on Method of Moments.



a)



b)

Fig. 6. SWR and reflection coefficient obtained using TLM Method for: a) feed probe b) dividing probe

Furthermore, as in the case of Method of Moments, for this optimum value of length of feed probe, the length of dividing probes is also varied in the same range. The numerical results of power dividing probes are similar for all the probes since the structure of conducting cylindrical cavity is symmetrical. Results, which illustrate the effect of the dividing probe p_1 and its length to the reflection coefficient, are presented in the Fig.6.b. The best level of reflection coefficient of -16.8dB is occurred when the length of probes is $0.12\lambda_g$. Corresponding value of SWR is 1.34.

VI. CONCLUSION

This paper proposes an analysis of characteristics of power divider using linear electric probes inside conducting cylindrical cavity. The power divider configuration with one feed and four dividing probes are analysed by using 3D TLM approach including the wire node for probes modelling. 3D TLM software is applied to study effects of the feed probe and power dividing probes for determining the maximum power transfer for practical applications. The TLM nunerical results of reflection coeficient and standing wave ratio (SWR) of feed probe and dividing probes for various probe length are compared with results obtained using Method of Moments and good agreement has ben achieved. Comparing numerical results with experimantal ones, it has been shown that the application of the TLM method, in comparison with Method of Moments, is more convenient for power divider modelling because alows modelling of wire structures with real thickness.

REFERENCES

[1] H. Holma, A. Toskala, "WCDMA for UMTS", *John Wiley& Sons*, New York, 2000.

[2] R. Wongsan, C.Phongcharoenpannich, M. Krairiksh, "Analysis of input impedance of a sectorial cylindrical caviyu-backed slot antenna fed by probe", *Proceeding of International Symposium on Antenna and Propagation*, Fukuoka, pp. 613-616, 2000.

[3] B. Milovanović, N. Dončov and J. Joković, "Real Excitation Modeling in a Loaded Cylindrical Metallic Cavity Using 3-D TLM Method," *Proceedings of the 5th IEEE TELSIKS 2001 Conference*, Niš, Yugoslavia, pp. 213-216, 2001

[4] V.Trenkić, A.J.Wlodarczyk and R.A.Scaramuzza, "Modelling of Coupling between Transient Electromagnetic Field and Complex Wire Structures", *International Journal of Numerical Modelling: Electronic Networks*, Devices and Fields, Vol.12, No.4, pp.257-273, 1999

[5] B. Milovanović, A. Marinčić, N. Dončov, V. Marković, J. Joković, and A. Atanasković, "Analysis of real feed probe influence to the resonant frequencies and field distribution in the cylindrical metallic cavity using 3D TLM method", *Proceedings of the 6th IEEE TELSIKS 2003 Conference*, Nis, Serbia and Montenegro, pp. 223-228. 2003.

[6] S. Amnartpluk, C. Phongcharoenpanich, S. Kosulvit, M. Krairiksh, "A power divider using linear electric probes coupling inside conducting cylindrical cavity", *Circuits and Systems, Proceedings of the 2003 International Symposium on ISCAS '03*, Volume 3, pp. 419-422, 25-28 May 2003.

[7] C. Christopoulos, "The Transmission-Line Modelling Method", *IEE/OUP Press*, 1995.

.[8] V. Trenkić, "The Development and Characterization of Advanced Nodes for TLM Method" *Ph.D. Thesis*, University of Nottingham, 1995.

Reflection Coefficient Equations for Tapered Transmission Lines

Zlata Cvetković[1], Slavoljub Aleksić, Bojana Nikolić

Abstract – **An approximation theory, based on the theory of small reflections, to predict the reflection coefficient response as a function of the impedance taper is applied in the paper. Obtained result is applied to a few common types of tapers. Transmission lines with exponential, triangular and Hermite taper are considered. All results are plotted using program package Mathematica 3.0. The presented results should be useful in solving reflection coefficient problems.**

Keywords – **Reflection coefficient, Voltage standing wave ratio (VSWR), Exponential, triangular and Hermite taper.**

I. Introduction

The analysis of nonuniform transmission lines has been a subject of interest of many authors. Uniform transmission lines can be used as impedance transformers depending on the frequency and length of the line [1]. The nonuniform lines have the advantage of wide-band impedance matching when they are used as impedance transformers and larger rejection bandwidths when they are used as filters [1, 2]. Some results for nonuniform exponential loss transmission line used as impedance transformer are presented in papers [3-5]. The paper [6] gives a solution in closed-form of the equation for value of arbitrary complex impedance transformed through a length of lossless, nonuniform transmission line with exponential, cosine-squared and parabolic taper.

In this paper we will derive an approximation theory, based on the theory of small reflections, to determine the reflection coefficient response as a function of the impedance taper. Obtained result is applied to transmission lines with exponential, triangular and Hermite taper [7]. For all examples the voltage standing wave ratio is plotted using program package Mathematica 3.0.

II. Reflection coefficient equation for nonuniform transmission line

The traditional way of determining RF impedance was to measure voltage standing wave ratio (VSWR) using an RF detector, a length of slotted transmission line and a VSWR meter. VSWR is defined as the maximum value of the RF envelope over the minimum value of the RF envelope. As the probe detector was moved along the transmission line, the relative position and values of the peaks and valleys were noted on the meter. From these measurements, impedance could be derived. The procedure was repeated at different frequencies. Modern network analyzers measure the incident and reflected waves directly during a frequency sweep, and impedance results can be displayed in any number of formats.

Reflection loss is away to express the reflection coefficient in logarithmic terms (dB). The reflection coefficient is the ratio of the reflected signal voltage level to the incident signal voltage level. Reflection loss is the number of decibels that the reflected signal is below the incident signal. Reflection loss is always expressed as a positive number and varies between infinity for a load at the characteristic impedance and 0 dB for an open or short circuit.

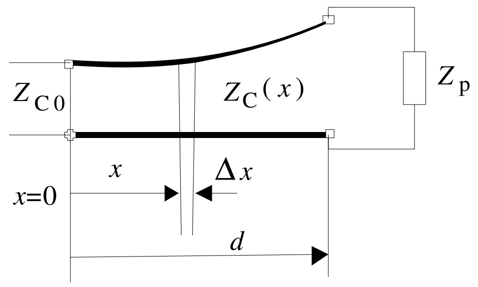


Fig. 1. A tapered transmission line matching section.

Let the transmission line shown in Fig. 1 be considered. The continuously tapered line can be modelled by large number of incremental sections of length Δx. One of these sections, connected at, has a characteristic impedance of $Z_C(x)+\Delta Z_C$ and one before has a characteristic impedance of $Z_C(x)$, as it is shown in Fig. 2. These impedance values are conveniently normalized by Z_{C0}. Then the incremental reflection coefficient from the step at distance x is given by

$$\Delta\Gamma = \frac{(Z_C(x)+\Delta Z_C)-Z_C(x)}{(Z_C(x)+\Delta Z_C)+Z_C(x)} \approx \frac{\Delta Z_C}{2 Z_C}. \qquad (1)$$

In the limit as $\Delta x \to 0$, it can be written as

$$\mathrm{d}\Gamma = \frac{d Z_C}{2 Z_C} = \frac{1}{2}\frac{\mathrm{d}(\ln Z_C(x))}{\mathrm{d}x}\mathrm{d}x \qquad (2)$$

since

$$\frac{\mathrm{d}(\ln f(x))}{\mathrm{d}x} = \frac{1}{f(x)}\frac{\mathrm{d}(f(x))}{\mathrm{d}x}.$$

---

[1] Zlata Z. Cvetković is with the Faculty of Electronic Engineering, University of Nis, Aleksandra Medvedeva 14, 18 000 Niš, Serbia and Montenegro, E-mail: {zlata, as, bojananik}@elfak.ni.ac.yu

The corresponding incremental reflection coefficient at the input end can be written as follows

$$\mathrm{d}\Gamma_{in} \approx e^{-j2\beta x}\,\mathrm{d}\Gamma . \quad (3)$$

By using formula (2) of small reflections, the total reflection coefficient at the input end of the tapered section can be determined by summing all the partial reflections with their appropriate phase angles

$$\Gamma_{in} = \int_0^d \mathrm{d}\Gamma_{in} = \frac{1}{2}\int_0^d e^{-\mathrm{j}2\beta x} \frac{\mathrm{d}}{\mathrm{d}x}\left(\ln\frac{Z_C(x)}{Z_{C0}}\right)\mathrm{d}x = \left|\Gamma_{in}\right| e^{-\mathrm{j}\varphi} . \quad (4)$$

So if $Z_C(x)$ is known, reflection coefficient at $x = 0$ can be found as a function of frequency. On the other hand, if Γ_{in} is specified, then in principle $Z_C(x)$ can be determined.

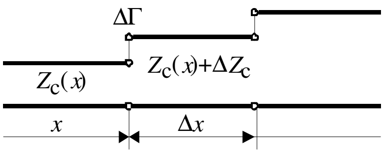


Fig. 2. Model for an incremental step change in impedance of the tapered line.

When a transmission line is terminated with an impedance, Z_P, that is not equal to the characteristic impedance of the transmission line, Z_C, not all of the incident power is absorbed by the termination. One part of the power is reflected back so that phase addition and subtraction of the incident and reflected waves creates a voltage standing wave pattern on the transmission line. The ratio of the maximum to minimum voltage is known as the voltage standing wave ratio (VSWR) and successive maxima and minima are spaced by 180°.

If the equation for reflection coefficient is solved for the VSWR, it is found that

$$\mathrm{VSWR} = \frac{1+\left|\Gamma_{in}\right|}{1-\left|\Gamma_{in}\right|} . \quad (5)$$

The reflection or return loss is related through the following equation

$$\text{Reflection loss} = 20\ \log_{10}\left|\Gamma_{in}\right|\ \mathrm{dB} . \quad (6)$$

Reflection loss is a measure in dB of the ratio of power in the incident wave to that in the reflected wave, and as defined above always has a positive value.

Also of considerable interest is the mismatch or insertion loss. This is a measure of how much the transmitted power is attenuated due to reflection. It is given by the following equation:

$$\text{Insertion loss} = -10\ \log_{10}\left(1-\left|\Gamma_{in}\right|^2\right)\mathrm{dB} . \quad (7)$$

Transmission line attenuation improves the VSWR of a load or antenna. Therefore, if you are interested in determining the performance of antennas, the VSWR should always be measured at the antenna connector itself rather than at the output of the transmitter. Transmission lines should have their insertion loss (attenuation) measured in lieu of VSWR, but VSWR measurements of transmission lines are still important because connection problems usually show up as VSWR spikes.

III. Tapered transmission lines

In this paper, three types of nonuniform transmission lines with exponential, triangular and Hermite function taper are considered.

A. Exponential taper

Along an exponential taper, the impedance is changing exponentially with distance,

$$Z_C(x) = Z_{C0}\,\mathrm{e}^{kx}, \text{ for } 0 < x < d \quad (8)$$

as indicated in Fig. 3. At $x = 0$ we have $Z_C(0) = Z_{C0}$. At $x = d$ we wish to have $Z_C(d) = Z_{Cd} = Z_{C0}e^{kd}$, what determines the constant

$$k = \frac{1}{d}\ln M , \quad (9)$$

where k is a taper coefficient and

$$M = \frac{Z_{Cd}}{Z_{C0}} \quad (10)$$

is a taper ratio. Z_{C0} and Z_{Cd} are the characteristic impedances of the transmission line at the left (source) and right (load) sides, respectively.

From (4), the total reflection coefficient at the input end is found as

$$\Gamma_{in} = \frac{1}{2}\int_0^d e^{-\mathrm{j}2\beta x}\frac{\mathrm{d}}{\mathrm{d}x}\left(\ln e^{kx}\right)\mathrm{d}x =$$

$$= \frac{1}{2}\ln\frac{Z_{Cd}}{Z_{C0}}\, e^{-\mathrm{j}\beta d}\,\frac{\sin\beta d}{\beta d} = \left|\Gamma_{in}\right| e^{-\mathrm{j}\varphi} . \quad (11)$$

B. *Triangular taper*

Characteristic impedance along triangular taper changes as

$$Z_C(x) = Z_{C0}\,\mathrm{e}^{2\left(\frac{x}{d}\right)^2 \ln\frac{Z_{Cd}}{Z_{C0}}}, \text{ for } 0 < x < d/2$$

$$Z_C(x) = Z_{C0}\,\mathrm{e}^{\left[4\frac{x}{d} - 2\left(\frac{x}{d}\right)^2 - 1\right]\ln\frac{Z_{Cd}}{Z_{C0}}}, \text{ for } d/2 < x < d , \quad (12)$$

as it is presented in Fig. 3.

Evaluating reflection coefficient from (4) gives

$$\Gamma_{in} = \frac{1}{2} \ln M \, e^{-j\beta d} \left[\frac{\sin \beta d}{\beta d} \right]^2 = \left| \Gamma_{in} \right| e^{-j\varphi} \qquad (13)$$

C. Hermite taper

In the case of Hermite taper, impedance of the transmission line varies with distance x as follow

$$Z_C(x) = Z_{C0} \, \mathrm{e}^{(k\,x)^2} \text{, for } 0 < x < d \qquad (14)$$

The resulting reflection coefficient response is given by (4)

$$\Gamma_{in} = \ln M \, \frac{e^{-j2\beta d} \left(1 + j2\beta d - e^{j2\beta d} \right)}{(2\beta d)^2} = \left| \Gamma_{in} \right| e^{-j\varphi}. \qquad (15)$$

Fig. 3 shows the impedance variations for the exponential, triangular and Hermite tapers of the same taper ratio, $M = 3$.

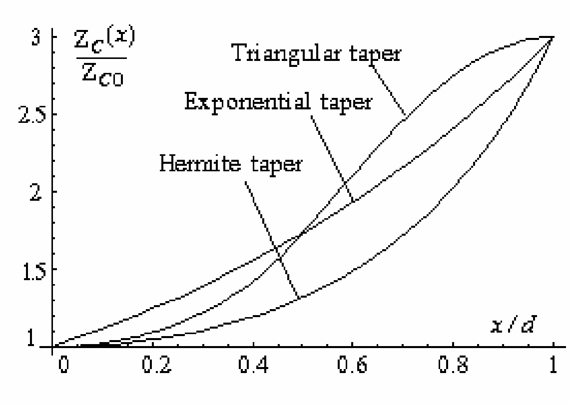


Fig. 3. Impedance variations for the exponential, triangular and Hermite tapers of the same taper ratio, $M = 3$.

IV. Numerical results

According to the analysis presented above for exponential, triangular and Hermite transmission line, different calculations are done.

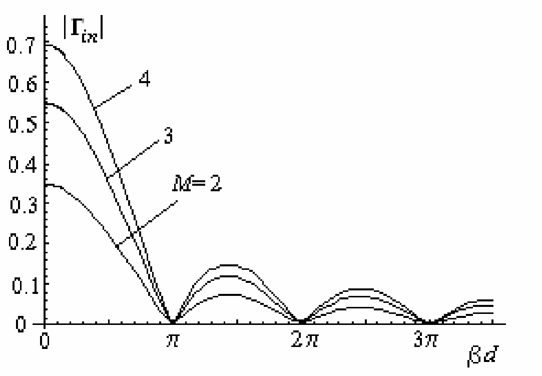


Fig. 4. Resulting reflection coefficient magnitude response on exponential line for different taper ratio.

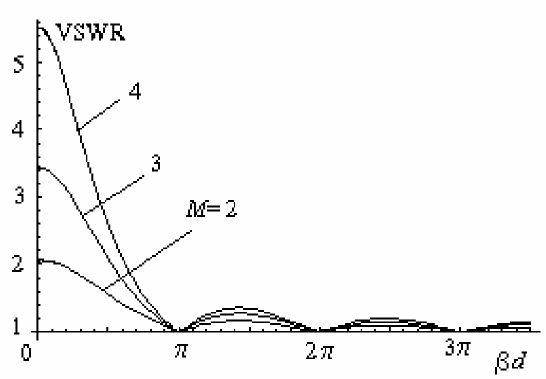


Fig. 5. Resulting VSWR magnitude response on exponential line for different taper ratio.

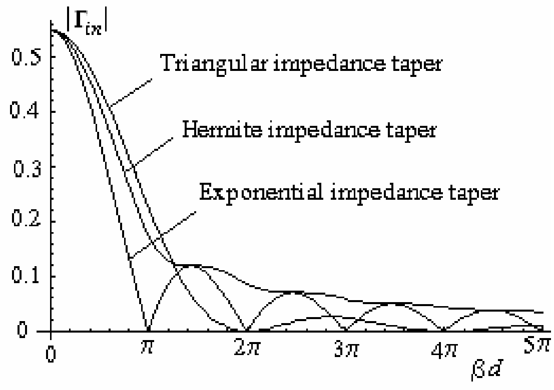


Fig. 6. Resulting reflection coefficient magnitude versus frequency for the exponential, triangular and Hermite taper.

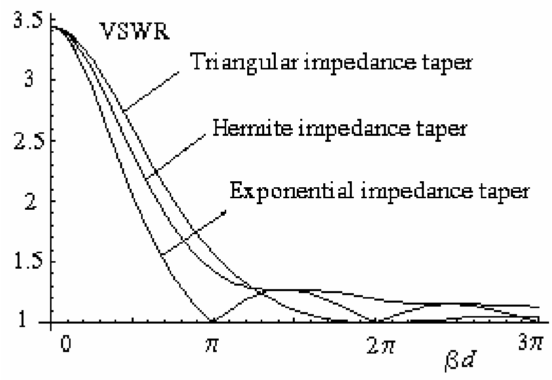


Fig. 7. Resulting reflection coefficient magnitude versus frequency for the exponential, triangular and Hermite taper.

Fig. 4 shows resulting reflection coefficient magnitude response on exponential line for different taper ratio $M = 2, 3, 4$. For the same taper ratios on exponential line, Fig. 5 presents VSWR magnitude response.

In order to compare obtained results, we consider an exponential, triangular and Hermite taper transmission line, which are used at the same taper ratio $M = 3$, as impedance transformers from $Z_{C0} = 100\,\Omega$ to $Z_{Cd} = 300\,\Omega$.

Resulting reflection coefficient magnitude versus frequency for the exponential, triangular and Hermite transmission line taper is shown in Fig. 6.

In Fig. 7, the voltage standing wave ratio (VSWR) of these nonuniform transmission lines is shown.

All the figures are plotted using program package Mathematica 3.0.

V. CONCLUSION

The modern network analyzer system sweeps very large frequency bandwidths and measures the incident power, P_i, and the reflected power, P_r . Because of the considerable computing power in the network analyzer, the reflection loss is calculated from the equation given previously, and displayed in real time. Optionally, the VSWR can also be calculated from the reflection loss and displayed real time.

In this paper, by using small reflections, the total reflection coefficient at the input end of the tapered section is determined by summing all the partial reflections up these incremental reflections with their appropriate phase angles. For exponential transmission line, resulting reflection coefficient and VSWR magnitude response for different taper ratio are presented. In order to compare obtained results for exponential, triangular and Hermite transmission line, resulting reflection coefficient and VSWR are shown for the same taper ratio $M = Z_{Cd} / Z_{C0} = 3$.

Obtained results are plotted using program package Mathematica 3.0.

REFERENCES

[1] Devendra K. M., *Radio-Frequency and Microwave Communication Circuits: Analysis and Design,* John Wiley & Sons, Inc. 2001.

[2] Ghausi, S. M., and Kelly, J. J., *Introduction to Distributed-Parameter Networks*, Holt, Rinehart and Winston, Inc., New York, 1968.

[3] Cvetkovic Z.Z., Aleksic S.R., Javor V.: "Lossy Exponential Transmission Line as Impedance Transformer," *Proceedings of papers*, *XIII International Symposium on Theoretical Electrical Engineering PSC 2005*, pp. 167-170, Timisoara, Romania, November 2005.

[4] Cvetkovic Z.Z. Petkovic R.A."Characteristics of Loss Exponential Line Network", *Proc. of* 4^{th} *Electronic Devices and Systems Conference EDS'96,* pp. 34-38, Brno, Czech Republic, 1996.

[5] Cvetkovic Z.Z., "Eksponencijalni vod kao sirokopojasni transformator impedanse ", *II Telekomunikacioni forum TELFOR'94,* str. 183-186. Beograd, 1994.

[6] Ahmed M. J.: *Impedance Transformation Equations for Exponential, Cosine−Squared, and Parabolic Tapered Line Transmission Lines*, IEEE Trans. on Microwave Theory and Tech., Vol. MTT-29, No. 1, pp. 67-68, 1981.

[7] Kobayashi K., Nemoto Y., Sato R., “Time Equivalent Representation of Nonuniform Transmission lines Based on the Extended Kuroda’s Identity”, IEEE Trans. on Microwave Theory and Tech ., Vol. 30, No.2, pp. 140-146, 1982.

Efficient Neural Model of Microwave Patch Antennas

Zoran Stanković, Bratislav Milovanović, Marija Milijić

Abtract – **Modeling of patch antennas using multilayer perceptron neural network model (MLP) is presented in this paper. To achieve high accuracy, two neural networks are used for calculating resonant frequency f_r and minimum value of s_{11} parameter (s_{11min}). Both neural networks have the same four input parameters: patch antenna length *L*, patch antenna width *W*, depth of patch antenna slot *l* and width of patch antenna slot *s*, but they have different output: first neural network gives value of resonant frequency f_r and the second gives value of s_{11min} parameter.**

Keywords – **Neural network, patch antenna, modeling**

I. INTRODUCTION

The intense development of up-to-date wireless communication systems has led to widespread using of patch antennas [1]. Patch antennas have a lot of advantages then other traditional types of antennas (efficacy, compact physically realization and mechanical reliability). Requests concerning quality, performance and realization quickness of wireless communication systems dictate that tools for patch antenna modeling must be fast and failsafe.

A patch antenna is a narrowband, wide-beam antenna fabricated by etching the antenna element pattern in metal trace bonded to an insulating substrate. Because such antennas have a very low profile, are mechanically rugged and can be conformable, they are often mounted on the exterior of aircraft and spacecraft, or are incorporated into mobile radio communications devices. A typical patch antenna is shown in Figure 1. A patch antenna consists of a radiating patch of any planar geometry (e.g. circle, square, ellipse, ring and rectangle) on one side of a dielectric material substrate backed by a ground plane on the other side [2]. The patch is generally made of conducting material such as copper or gold [3]. The radiating patch and the feed lines are usually photo etched on the dielectric substrate. Generally, a substrate with a low dielectric constant (ε_r) is used (typically ~ 2.5), but in loss critical applications, Alumina (ε_r =10) must be used. Patch antennas are also relatively inexpensive to manufacture and design because of the simple 2-dimensional physical geometry. Their properties allow patch antennas to be used in many areas of communications (in wireless applications, as antenna in mobile phone, pager, etc., and in satellite communications) [4].

The authors are with the Faculty of Electronic Engineering, Aleksandra Medvedeva 14, Niš, Serbia and Montenegro,
E-mail: zoran@elfak.ni.ac.yu, batam@pogled.net, marijam@elfak.ni.ac.yu,

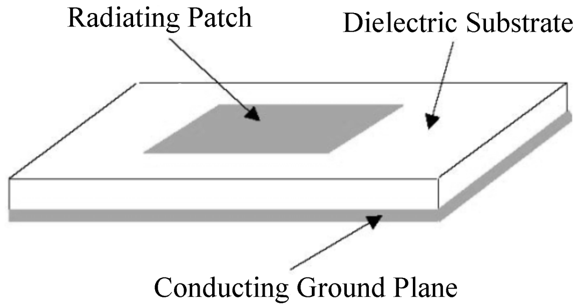


Fig. 1. Typical patch antenna

The most known method used for modeling of patch antennas is electromagnetic simulation. Although it gives models with high accuracy, it has some disadvantages. Its basic disadvantage is that electromagnetic simulation has high demands concerning the hardware resources necessary for its software implementation. The software implementation itself might be very complicated and faced with many difficulties. Also the time needed for numerical calculation when using an electromagnetic simulation could be unacceptably long.

Good alternative for overcoming all these problems is modeling of patch antennas using an artificial neural network model. Neural network model in these cases can be fast and accurate as detailed EM simulation method [5,6]. To improve accuracy of modeling of patch antennas using neural network model, two neural networks are used: one for modeling patch antenna resonant frequency f_r and second for modeling patch antenna s_{11min} parameter.

II. NEURAL MODEL OF PATCH ANTENNA

Multilayer perceptron (MLP) neural network is high-parallel and high-adaptive feed-forward structure that is consisted of mutually connected neurons with nonlinear activation functions in hidden layers [7,8,9]. Researching of MLP application in wireless communications has showed that this network is able to approximate highly nonlinear functions with satisfactory accuracy and high level of generalization. Using this structure there is no need of knowledge for the explicit functional connection between the output and input parameters. Architecture of the MLP neural model for modeling of patch antennas is presented in Fig.2. It consists of two neural networks: one for modeling patch antenna resonant frequency f_r and second for modeling patch s_{11min} parameter. Both neural networks consist of neurons grouped into the following layers: input layer, output layer and one or more hidden layers. The number of neurons in the hidden layers can be variable. The general symbol of this type of MLP neural model is MLPH-N_1-N_2-...-N_H, where H is the number of hidden layers, and N_i is the number of neurons in the i-th hidden layer.

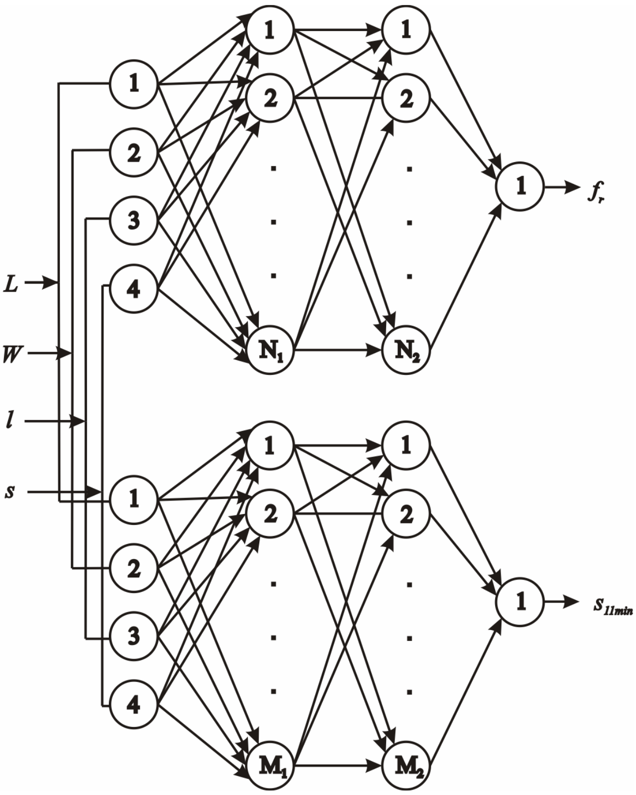


Fig. 2. MLP neural model for modeling of patch antennas

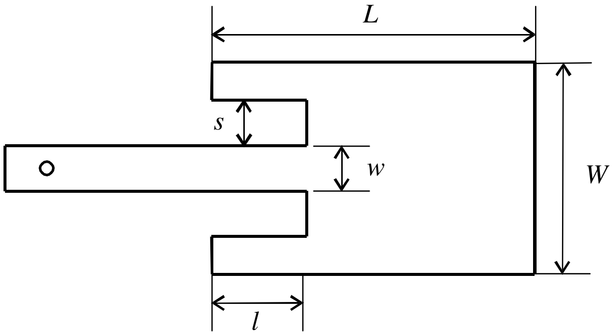


Fig. 3. Slotted patch antenna

Each neural network has been trained separately. Their numbers of neurons in hidden layers and their transfer functions in hidden and output layers have been changeably to achieve a goal of minimum error.

First neural network has been trained to calculate value of resonant frequency f_r. Input vector $\mathbf{x}=[L,W,l,s]^T$ is presented to the input layer and fed through the network that then yields the output $\mathbf{y}=[f_r]$. Therefore, the neural model is given by $\mathbf{y}=y(\mathbf{x},W)$ where W is a set of connection weight matrices among neurons [7,8,9]. The vector of l-th hidden layer outputs is:

$$\mathbf{y}_l=F(\mathbf{w}_l\mathbf{y}_{l-1}+\mathbf{b}_l) \quad (1)$$

where $\mathbf{y}_l$ is a $N_l \times 1$ vector of l-th hidden layer outputs, $\mathbf{y}_{l-1}$ is a $N_{l-1} \times 1$ vector of (l-1)-th hidden layer outputs, $\mathbf{w}_l$ is a $N_l \times N_{l-1}$ connection weight matrix among (l-1)-th and l-th hidden layer neurons, and $\mathbf{b}_l$ is a vector containing biases of l-th hidden layer neurons. In the above notation $\mathbf{y}_0$ represents outputs of the buffered input layer $\mathbf{y}_0 = \mathbf{x}$. F is the transfer function of hidden layer neurons and F was changeable during training process to obtain network with minimum error. All neurons from the last hidden layer H are connected with the neuron of the output layer. Since the transfer function of output layer is linear, the output of the network is:

$$f_r=\mathbf{w}_o\mathbf{y}_H$$

where $\mathbf{w}_o$ is a $1 \times N_H$ connection weight matrix among the H-th hidden layer neurons and output layer neuron (Fig. 2).

Table 1. Test results of neural network for patch antenna f_r modeling

Layer-Transfer function and number of neurons			WCE	ACE	p^{rom}
Hidden		Output			
1	2				
Tansig-9	Tansig-9	Purelin-1	9.88%	0.88%	0.98
Logsig-10	Logsig-5	Purelin-1	14.5%	0.96%	0.99
Logsig-9	Tansig-9	Purelin-1	19.11%	0.86%	0.99
Losig-10	Tansig-8	Satlins-1	19.21%	0.89	0.99
Tansig-15	Logsig-15	Purelin-1	23.1%	0.9%	0.98

Table 1. Test results of neural network for patch antenna s_{11min} parameter modeling

Layer-Transfer function and number of neurons			WCE	ACE	p^{rom}
Hidden		Output			
1	2				
Tansig-15	Tansig-10	Satlin-1	11.75%	3.45%	0.96
Logsig-10	Logsig-5	Satlin-1	19.9%	3.72%	0.96
Tansig-14	Tansig-10	Poslin-1	17.32%	4.54%	0.95
Logsig-9	Logsig-9	Poslin-1	24.6%	3.38%	0.96
Tansig-10	Tansig-8	Purelin-1	21.91%	3.68%	0.96

Output from second neural network is s_{11min} parameter of patch antenna. This network has the same input as the first network. During training it must approximate function $\mathbf{y}=y(\mathbf{x},W)$, where $\mathbf{x}=[L,W,s,l]^T$ is model input, $\mathbf{y}=[s_{11min}]^T$ model output and W is a set of connection weight matrices among neurons. Also, the second neural network is feed-forward and it has the same principles and formulas of connections among neurons and layers. Its transfer functions were changeable during training process, too. The output of the second network is:

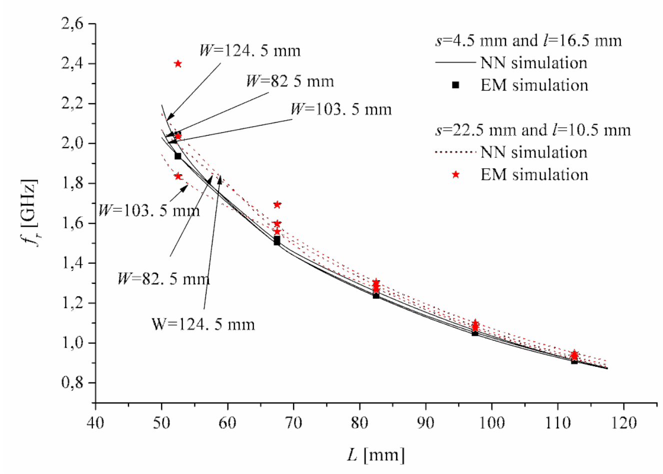


Fig. 4. f_r in the function of parameter L

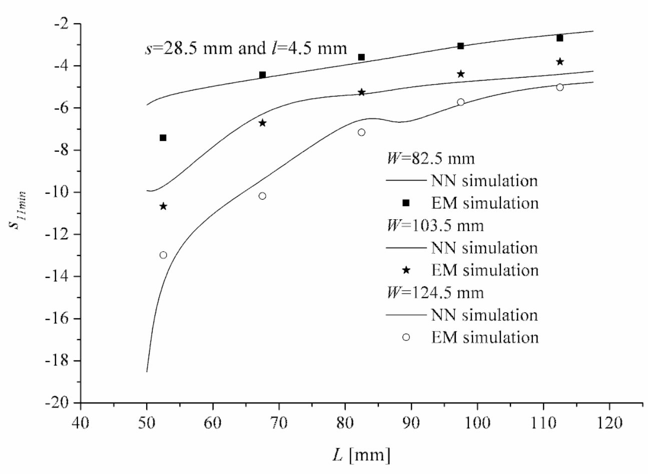


Fig. 6. s_{11min} in the function of parameter L

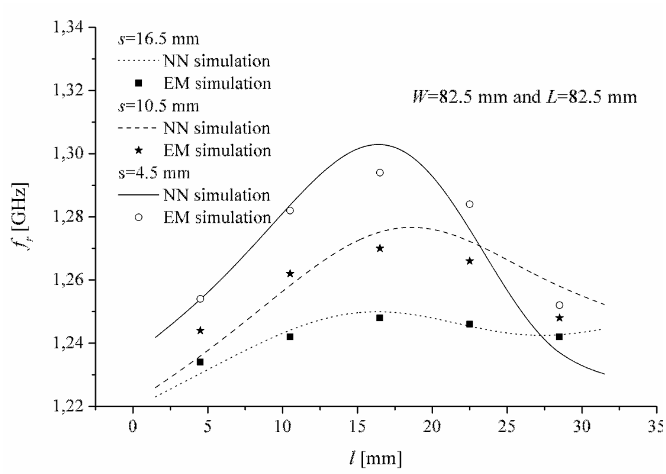


Fig. 5. f_r in the function of parameter l

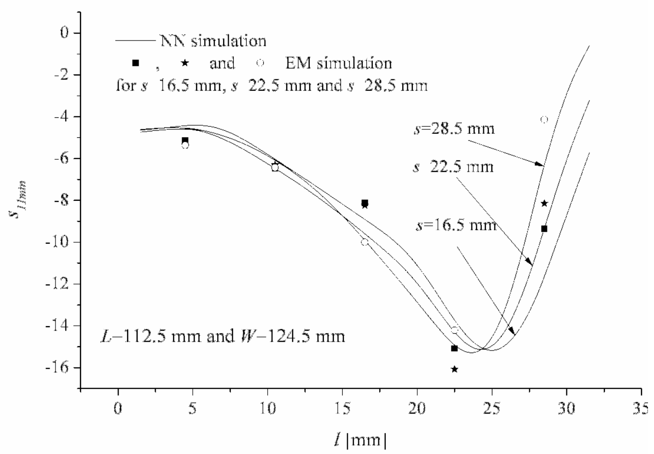


Fig. 7. s_{11min} in the function of parameter l

$$s_{11min}=\mathbf{w}_o\mathbf{y}_H$$

where $\mathbf{w}_o$ is a $1 \times N_H$ connection weight matrix among the H-th hidden layer neurons and output layer neuron (Fig.2).

This neural model is used for slotted patch antenna, which is shown in Figure 3. Antenna is constructed of substrate with fallowing features: relative dielectric constant $\varepsilon_r = 2.17$, thickness $h = 0.508$ mm, conductor thickness $T = 0.017$ mm. Electromagnetic simulation of slotted patch antenna using HFSS 9.0 software, for constant parameter w =0.5 mm and for variable parameter L, W, l and s, has given values of resonant frequency f_r and s_{11min} parameter. Electromagnetic simulation results have been used as input for training neural model whose output is values of parameters f_r and s_{11min} parameter.

III. Test Results

Each neural network is tested to check obtained generalization level. The test of both neural networks has been done using uniform test group consisted of samples which have not been used in training. To quantify models' accuracy an average test error (*ACE* [%]), worst-case error (*WCE* [%]), and Pearson Product-Moment correlation coefficient (r^{PPM}) between the referent and the modeled data were calculated. In Table 1 and Table 2., test results for the both neural networks are presented. In both cases, five best models are shown.

IV. Simulation Results

Neural models M4-9-9 (transfer functions: Tansig, Tansig, Purelin) and M4-15-10 (transfer functions: Tansig, Tansig, Satlin) have the best testing results and they are used for checking obtained generalization level. First neural models M4-9-9 has been trained for modeling patch antenna resonant frequency f_r and it is used for representation resonant frequency f_r as function of antenna parameters L and l and for comparison these results with results obtained by EM simulation (Figure 4. and Figure 5.). These figures show the quality of neural models M4-9-9 training which is very satisfying. Also, they represent how resonant frequency f_r depends on antenna parameters L and l (when other parameters are constant).

Also, neural models M4-15-10, trained for modeling patch antenna s_{11min} parameter, is used for verifying its genera-

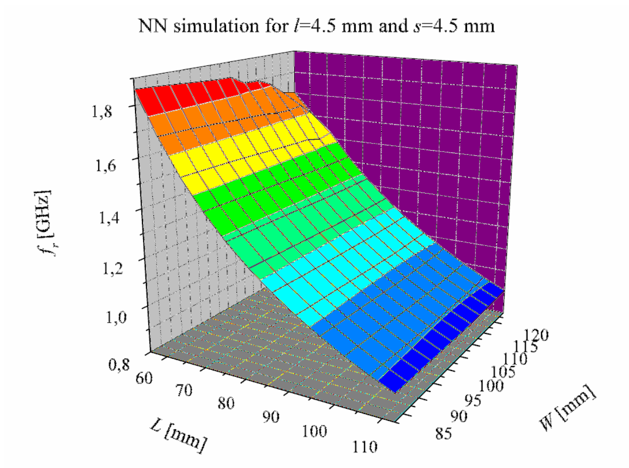


Fig. 8. f_r in the function of parameters L and W

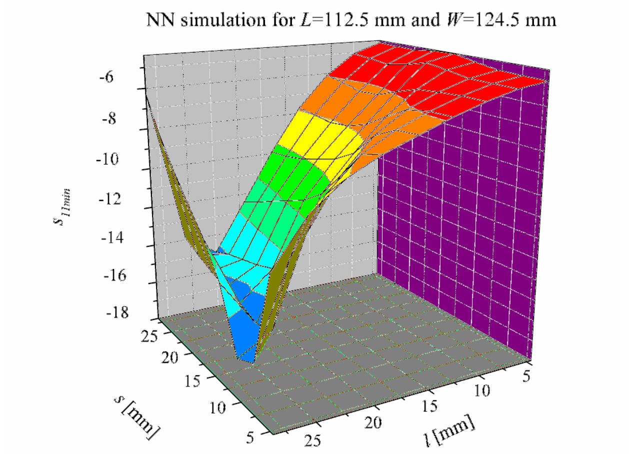


Fig. 9. s_{11min} in the function of parameters l and s

lization level. Parameter s_{11min} is shown as function of antenna parameters L and l and these results have been compared with values obtained by EM simulation (Figure 6. and Figure 7.). As these figures show, neural models M4-15-10 is correct and it can be used for modeling patch antenna s_{11min} parameter very correctly.

Both neural networks have satisfying precision, but they improve patch antenna modeling with very great speed of work. In Figure 8., the dependence f_r on parameter L and W, when l and s parameters are constant, is shown. This dependence is shown using 315 values of f_r obtained by NN simulation for less then 1 seconds. If we use EM simulation to obtain the same number of f_r values, we will do it for 2 hours. Also, Figure 9. represents s_{11min} as function of parameters s and l when W and L are constant values. If we use EM simulation to obtain 441 values of s_{11min} parameter, we will finish it after 3 hours of simulation. NN simulation obtains 441 needed values of s_{11min} parameter for 1 second. For these reasons, NN simulation is better alternative in applications where simulation has to be finished in certain period of time.

V. Conclusion

Software including EM simulation is the most used method for modeling of patch antenna. Its accuracy is good, although it has many limitations (the simulation process is complicated and slow, there is the problem of hardware resources necessary for its software implementation). Good alternative is to use neural network modeling. MLP model is easy to be developed, has high simulation speed and satisfying accuracy.

Further usage of this neural model of patch antennas will be as base of software for modeling patch antennas. It will represent software part for numerical calculation and software interface will be part which will respond on user requests.

References

[1] www.odyseus.nildram.co.uk/Systems_ And_Devices_ Files/Patch_Antenna.pdf

[2] www.odyseus.nildram.co.uk/Systems_ And_Devices_Files/Patch_Antenna.pdf

[3] etd.lib.fsu.edu/theses/available/etd-04102004-143656/unrestricted/Chapter3.pdf

[4] Lal Chand Godara “Handbook of Antennas in Wireless Communications”, Chapter 3, CRC Press, 2002

[5] “Modeling of Patch Antennas Using Neural Networks”, Bratislav Milovanović, Marija Milijić, Aleksandar Atanasković, Zoran Stanković, TELSIK 2005 Konferencija, Niš, 2005

[6] “Modelovanje mikrotalasnih patch antena pomoću neuronskih mreža”, Marija Milijić, Zoran Stanković, Aleksandar Atanasković, ETRAN 2005, Budva, 2005

[7] Q. J. Zhang, K. C. Gupta, Neural Networks for RF and Microwave Design, Artech House,2000.

[8] S. Haykin, Neural Networks, New York, IEEE, 1994.

[9] Wang, F., and Q. J. Zhang, Knowledge Based Neural Networks for Microwave Design, IEEE Trans. Microwave Theory and Techniques, Vol. 45, 1997, pp. 2333-2343.

29 June – 1 July 2006, Sofia, Bulgaria

A Unified Neural Network for DC and RF Modeling of AlGaAs HBT's

Vera Marković, Aleksandar Stošić

Abstract – **The advantages of heterojunction bipolar transistors (HBTs) make them very promising for modern RF communication systems and there is a need for their valid description by means of a model. A procedure for HBT DC and RF modeling based on a unified neural network approach is presented in this paper. The proposed model is characterized by high accuracy and efficiency commonly requested for today's CAD techniques.**

Keywords – **HBT, neural networks, modeling**

I. Introduction

Due to the continuously increasing performance of digital wireless communication systems, the performances of active microwave devices have undergone a tremendous improvement in recent years as well. The range of modern microwave transistor available for the microwave wireless communication systems is wide and includes MESFETs, HEMTs, PHEMTs, HBTs, MOSFETs, etc. The choice of the RF transistors depends on the performance required for the selected wireless application as well as from the commercial accessibility, price, availability of CAD models and so on.

Heterojunction Bipolar Transistors (HBT) have become very promising devices for different applications at the microwave and millimeter-wave frequencies [1], [2]. They are used for power amplifiers as well as for low noise amplifiers in mobile communication systems. This device technology is considered as very convenient for RF front-end circuits in next-generation wireless communications.

Due to the increasing application of HBT's in microwave circuits and having a need for efficient design of these circuits, a valid description of these devices by means of a model is required. A shift in the design of microwave components can be observed: in addition to the electrical characteristics, other issues such as reduced time to market, yield optimization, manufactured-oriented design, tolerance analysis, etc., are becoming increasingly important.

During the last decade a tremendous work has been done for developing physical and empirical HBT models [3],[4]. Despite this fact, we still do not have a standard, fast and enough accurate model for HBTs. In most case, the Gummel-Poon model is insufficient for today's bipolar transistors. A lot of DC models and RF models can be find in the literature. The advanced transistor models could characterize the transistor operation in a large bias and frequency range at the cost of more complicated extraction methods and measurement efforts due to a large number of unknowns of the transistor equivalent circuit. However, it is not convenient to perform statistical CAD which requires, for instance, hundreds of analysis, by using these approaches.

Last years, from the aspect of efficiency, accuracy and simplicity, neural network approach has been considered to be a good solution for microwave device modeling [5]. They can handle severe nonlinearities that are present in the majority of practical problems. They are especially useful in situations where a classical model-based or parametric approach to information processing is difficult to formulate.

Once developed neural model provides fast response for different input vectors that in principle can cover the whole operating range. A very important property of neural networks is generalisation capability [6], which provides sufficiently accurate response for different vectors not included in the training set, without additional computational efforts or new measurements.

In this paper, the application of neural network approach for modeling DC and RF performances of AlGaAs/GaAs HBTs is presented. In this way, an efficient prediction of transistor's characteristics over the wide frequency and bias condition ranges can be enabled.

II. Modeling of HBT's by using Neural Network Approach

Fig 1. shows an overall neural network configuration that provides DC and *S*-parameters of an HBT at the output for any frequency and bias point within the transistor's operating range, presented at the input.

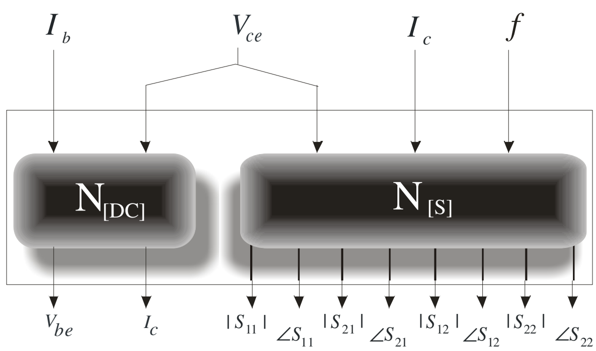


Fig 1. Neural network for DC and RF modeling of HBT's.

Neural network configuration presented in Fig. 1 is composed of two sub-networks, one for DC modeling, denoted by $N_{[DC]}$, and the other for *S*-parameter modeling, denoted by $N_{[S]}$. Both sub-networks are MLP (***M**ulti-**L**ayer **P**erceptron Network*)-type neural networks.

Vera Marković and Aleksandar Stošić are with the Faculty of Electronic Engineering, A. Medvedeva 14, 18000 Niš, Serbia & Montenegro, E-mails: vera@elfak.ni.ac.yu, stosha@elfak.ni.ac.yu

Neural sub-network $N_{[DC]}$ provides DC collector current I_c and DC base-emitter voltage V_{be} for any DC collector-emitter voltage V_{ce} and DC base current I_b within the transistor's operating range presented at the input of neural model. With respect to that, there are two neurons in the input layer corresponding to V_{ce} and I_b, and two neurons in the output layer corresponding to V_{be} and I_c.

Therefore, the proposed neural model can predict two DC characteristics: 1) $I_c(V_{ce}, I_b)$ - DC collector current I_c in terms of DC collector-emitter voltage V_{ce} and DC base current I_b, and 2) $V_{be}(V_{ce}, I_b)$ - DC base-emitter voltages V_{be} in terms of DC collector-emitter voltage V_{ce} and DC base current I_b. The DC data needed for obtaining a training set for the first sub-network have been measured for AlGaAs HBT's with common emitter configuration.

The second neural sub-network denoted by $N_{[S]}$ enables an accurate prediction of magnitudes and angles of four *S*-parameters over the whole frequency range and for any DC collector-emitter voltages V_{ce} and DC collector current I_c within the operating bias range. Hence, there are three neurons in the input layer of the second sub-network $N_{[S]}$ corresponding to V_{ce}, I_c, and frequency *f*, and eight neurons in the output layer corresponding to the magnitudes and angles of *S*-parameters.

In both cases two hidden layers have been chosen because in this way slightly better have been obtained than with a structure containing only one hidden layer. The numbers of neurons in hidden layers have been selected on the basis of testing several networks with different numbers of hidden neurons.

The neural sub-networks have been trained using a back-propagation algorithm that is commonly considered as quite adequate for this purpose. In order to compare the accuracy of the model, the average test error (ATE [%]), the worst-case error (WCE [%]), and the *Pearson Product-Moment* correlation coefficient (r) between the measured and simulated data [5] have been calculated. The correlation coefficient indicates how well the modeled values match the referent values, i.e. a value near 1 indicates an excellent predictive ability.

The test procedure has been performed not only for the data from the training set, but also for the data that are not used in the training process, with the aim of checking the generalization capability of developed neural networks.

The data for training and test sets that we used in modeling procedure had been obtained by the collaboration with a microwave laboratory at Northeastern University, Boston, USA, where HBT DC and *S*-parameter measurements were performed. The DC and *S*-parameter were directly measured on wafer AlGaAs HBT's denoted by HBT40020-002-8.

The measured data for V_{be} and I_c refer to collector-emitter voltages V_{ce} within -0.5V÷6V range and to base currents I_b of 50, 130, 210, 290 and 370 [μA]. The overall V_{ce} range was divided into two sub-ranges as follows: first sub-range (-0.5÷1)V with 0.05V step, and second sub-range (1÷6)V with 0.5V step. Therefore, the operating DC collector-emitter voltages V_{ce} range was covered with 41 discrete V_{ce} points and the operating DC base currents I_b range was covered with 5 discrete points. The V_{be} data and I_c data refer to 205 points and the total number of these DC data used in training and test procedure for the selected HBT transistor was 410. From this number, 328 data were used for the training and the rest of 82 data was used for a test set with the aim to check the generalization capability of the neural network.

Neural networks with a different number of hidden neurons varying between 2 and 10 neurons, have been trained. With the aim to additional improve the accuracy of ANN noise model triple successive training each neural network was performed. Therefore the effective number of trained neural networks was $3 \times 90 = 270$. The number of training epochs of each network was limited to a maximum of 180. The average time needed for the training process on a Pentium 4 with processor declared on 2500+ and 512MB RAM was 35 minutes. However, once trained, the network provides an instantaneous response for different input vectors.

The total number of *S*-parameters data used for training and test procedure for the selected HBT transistor was 5880. The data refer to the frequency range (0.05÷40) GHz. This frequency range was divided into four sub-ranges as follows: first sub-range (0.05÷0.5) GHz with 0.05 GHz step, second sub-range (0.5÷1) GHz with 0.1 GHz step, third sub-range (1÷10) with 1 GHz step, and fourth sub-range (10÷40) GHz with 2 GHz step. Therefore, operating frequency range was covered with 35 discrete frequency points. *S*-parameters have been measured for different combinations of DC collector-emitor voltages and base currents in the whole frequency range. DC collector-emitter bias had the fallowing values: 1V, 3V, 4V, and DC collector current had the following values [mA]: 0.5, 1.11, 2.03, 4.26, 9.01, 20.23, 29.99. Therefore, the measurements have been performed at 735 operating points and eight *S*-parameter data (magnitudes and angles) correspond to each point: $|S_{11}|$, $\angle S_{11}$, $|S_{12}|$, $\angle S_{12}$, $|S_{21}|$, $\angle S_{21}$, $|S_{22}|$, and $\angle S_{22}$. Training set was obtained by extracting 595 data points from the measurement data. Therefore the training set contained 4760 *S*-parameters data.

With the aim to avoid the errors caused by a rapid change of some *S*-parameters angle characteristics between the values -180° and +180°, a conversion of the angle range from this range to the range (0÷360)° has been applied.

By using these measured data, several neural networks with different number of hidden neurons (between 9 and 16) have been trained in the similar way as above.

In order to check the generalization capability of neural sub-network $N_{[S]}$, a test set is generated from the rest of data points containing 1120 *S*-parameters data. The bias points that have not been included in the training set had the following values:

1) $V_{ce} = 3V$, $I_c = 1.11mA$; 2) $V_{ce} = 1V$, $I_c = 4.26mA$;
3) $V_{ce} = 3V$, $I_c = 8.43mA$; 4) $V_{ce} = 4V$, $I_c = 4.09mA$

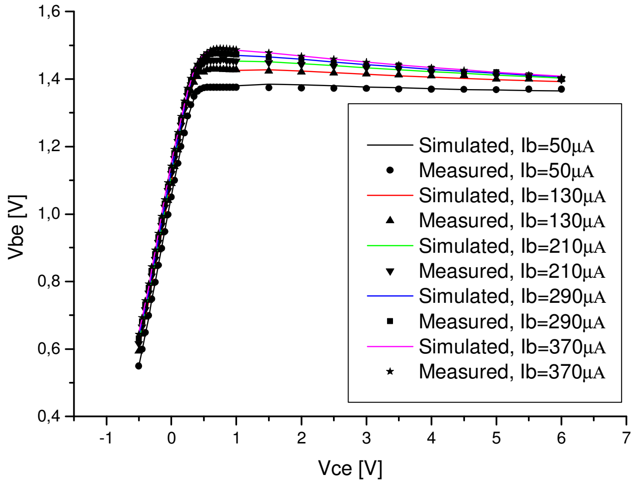


Fig. 2. The simulated (continual curve) and measured (symbols) DC characteristics $V_{be}(V_{ce}, I_b)$

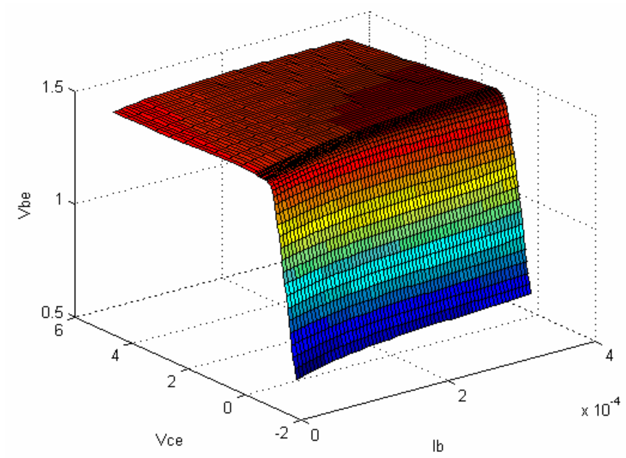


Fig. 4. Three-dimensional $V_{be}(V_{ce}, I_b)$ DC characteristics

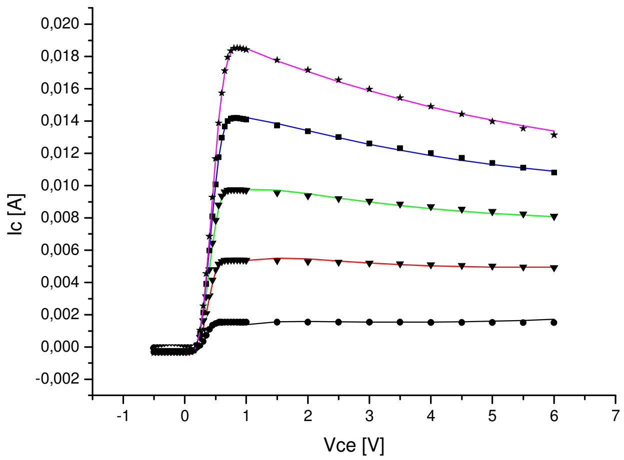


Fig. 3. The simulated (continual curve) and measured (symbols) DC characteristics $I_c(V_{ce}, I_b)$

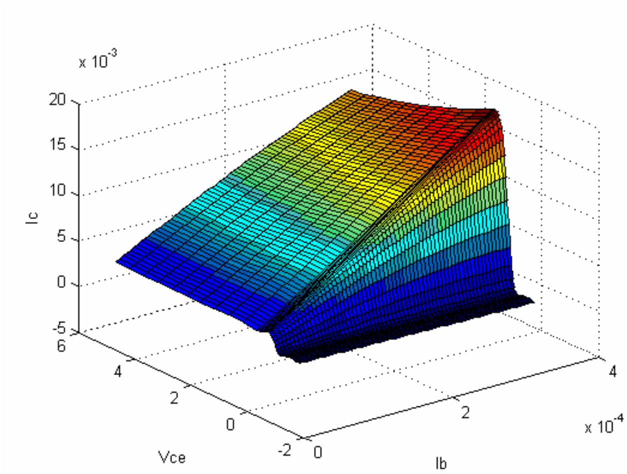


Fig. 5. Three-dimensional $I_c(V_{ce}, I_b)$ DC characteristics

III. Modeling Results

After the training process, neural models have been applied to get DC outputs as well as scattering parameter values for various input data different from the ones used for training. The results have been compared and on the basis of above-mentioned criteria, the best model has been selected. The best results give the first sub-network $N_{[DC]}$ marked by 1M4_4_9 and the second sub-network $N_{[S]}$ marked by 1M4_15_14. The number 1 denotes first of three successive training for the selected neural network. The number 4 shows that the neural network has four layers. Numbers 4 and 9 denote the number of neurons in the first and second hidden layer, respectively.

In Table 1, as an illustration of the accuracy of the selected model, test statistics for DC characteristics for training and test data is presented. It could be seen that the value of ATE for the training set data is less than 0,442%, and for the test set data is less than 0,602%. The value of WCE for the training set data is less than 1,660%, and for the test set data is less than 1,784%. The correlation coefficient *r*, in all cases, is greater than 0.99. These results show that the selected neural sub-network $N_{[DC]}$ gives results of great accuracy and the excellent predictive ability.

TABLE 1 Error statistics for DC characteristics

		ATE[%]:	WCE[%]:	r:
Training data	V_{be}	0.347	1.198	0.9999
	I_c	0.442	1.660	0.9998
Test data	V_{be}	0.476	1.197	0.9999
	I_c	0.602	1.784	0.9999

The simulated DC characteristics obtained by the selected neural model, compared with measured data, are presented in following figures: DC base-emitter voltages V_{be} and collector currents I_c versus DC collector-emitter voltages V_{ce} at five discrete DC base currents I_b points are shown in Fig. 2 and Fig. 3, respectively. It is important to note that the DC base-emitter voltages V_{be} and collector currents I_c are simulated for $I_b = 290\mu A$, a value not included in the training set. Very

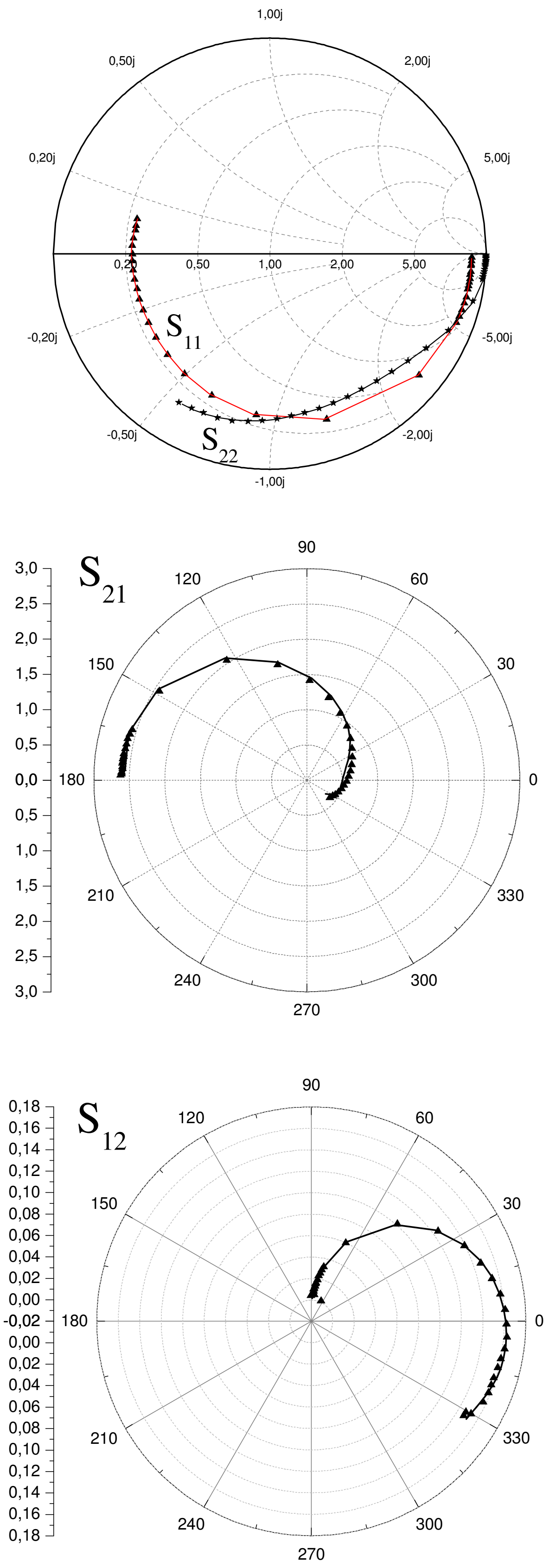


Fig. 6. The simulated (continual curve) and measured (symbols) *S*-parameters at the bias point $V = 3V$, $I = 1.11mA$

good agreement between simulated and measured characteristics can be observed in all cases, which means that the developed neural model has a good generalisation capability.

It is known that in some cases neural network over-learning could be happen. As a consequence, the prediction for the input values used for the training can be excellent (meaning very small ATE and WCE and correlation coefficient very close to one), but for some other inputs the network can give very bad and unexpected results. With the aim to checking the model validity additionally, neural network responses for practically continuous changes (small steps of change) of I_b and V_{ce}, have been generated and plotted. Figures 4 and 5 show the three-dimensional plots of DC base-emitter voltages V_{be} and DC collector currents I_c, respectively, as a function of continuous changes I_b and V_{ce}. The forms of three-dimensional surfaces confirm a very good prediction for the input values outside of the training set..

Fig. 6 shows the magnitudes and angles of *S*-parameters versus frequency, obtained by using the selected neural model, at a bias point not included in the training set. For the comparison purpose, measured data are shown in the same figure. It can be seen that the developed neural model can predict device *S*-parameters with a very good accuracy.

IV. Conclusion

A new, ANN-based unified approach can be used successfully for modeling the DC and *S*-parameters of HBTs. For developing a neural model only a number of measured data is needed. That gives an advantage to ANN approach in comparison with other modeling approaches, especially when the physical operating mechanisms of the device are too complex or not well known, which occurs often when some novel active devices for modern communication systems have to be considered. Developed neural models are characterized by high accuracy together with the efficiency and simplicity and therefore are convenient for CAD purposes.

References

[1] N.L.Wang, "Transistor Technologies for RFICs in Wireless Applications", *Microwave Journal*, pp. 98-110, February, 1998.

[2] P.F.Chen et al., "Application of GaInP/GaAs DHBTs to Power Amplifiers for Wireless Communications", *IEEE Trans. Microwave Theory Tech.*, vol. MTT-47, pp. 1433-1437, August 1999.

[3] R.Plana, L.Escotte, "Noise Properties of Microwave Heterojunction Bipolar Transistors", *MIEL'97 Proceedings,* pp. 215-222, Yugoslavia, 1997.

[4] A.Garlapati, S.Prasad, "Large Signal Characterization of Heterojunction Bipolar Transistors", *Proceedings of MICROCOLL'99*, pp. 181-189, Hungary, 1999.

[5] Q. J. Zhang, K. C. Gupta, Neural Networks for RF and Microwave Design, Artech House, 2000.

[6] S. Haykin, Neural networks, New York, IEEE, 1994.

ANNs in Bias Dependent Scalable Modeling of HEMT S-Parameters

Zlatica Marinković, Olivera Pronić, Vera Marković

Abstract – **An efficient procedure for modelling of scattering parameters for a class of microwave FETs manufactured in the same technology is presented in this paper. It is based on multilayer perceptron artificial neural network (ANN) that produces scattering parameters at its outputs for device gate width, biases and frequency presented at its inputs. After the ANN training, the scattering parameters' prediction under different operating conditions for any device from the class requires only calculation of the ANN response, without changes in the ANN structure. Numerical examples for *S* parameters modeling for one specific series of pHEMT devices are shown.**

Keywords – **Artificial neural network, scattering parameters, pHEMT**

I. Introduction

In the last decade some attempts have been made to model devices and systems in the area of microwaves by using artificial neural networks, [1]-[7]. Unlike complex time-consuming electromagnetic models, once developed neural models give responses almost instantaneously because response providing is based on performing basic mathematical operations and calculating elementary mathematic functions (such as an exponential or hyperbolic tangent function). ANNs have the capability of approximating any nonlinear function and the ability to learn from experimental data. Therefore, it is possible to develop neural model from source-response data points without the knowledge about the physical characteristics of the problem to be solved.

The most important feature of neural models is their generalization capability, i.e. the capability to provide the correct response even for the input values not presented during the training process. In that way, the developed models can be used for a reliable prediction over a wide range of input parameters.

The characterization of microwave FETs (MESFET, HEMT) includes knowledge about device signal and noise parameters that are frequency-, temperature- and bias-dependent. Since measurements of these parameters, especially of noise parameters, are complex and time-consuming procedures, device models are usually used for device *S*- and noise parameters' prediction in the microwave circuit design.

According to the recent research, neural networks seem to be good alternative to conventional transistor modeling. They enable transistor signal and noise modeling versus biases and/or temperature and frequency ([2], [3], [5]-[7]) in the whole device operating range. However, these models are valid for the considered device only.

In this paper, a problem of bias-dependent *S*- parameters prediction for different-gate-width microwave FET transistors is considered. The modeling procedure proposed here is an extension of the method presented earlier in [5]-[7]. The ANN inputs are not only biases and frequency, but also device gate-width, therefore the proposed ANN model is valid for a whole class of devices made in the same technology.

The paper is organized as follows. In Section II neural networks are shortly introduced. The transistor S- parameters modeling procedure based on ANNs is stated in Section III. Some modeling examples and the main results are presented in Section IV. Finally, in Section V, the main conclusions are reported.

II. Multilayer Neural Networks

A standard multilayer perceptron (MLP) neural network is shown in Fig.1. [1].

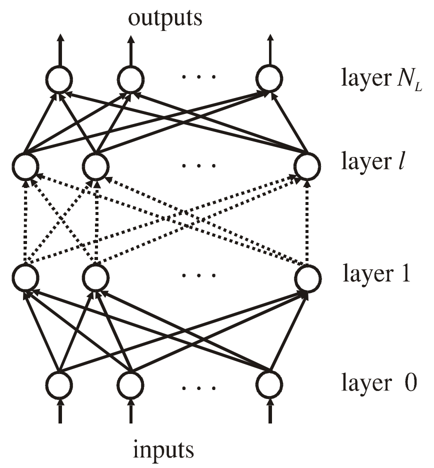


Fig. 1. MLP neural network

This network consists of an input layer (layer 0), an output layer (layer N_L), as well as several hidden layers.

Input vectors are presented to the input layer and fed through the network that then yields the output vector. The l-th layer output is:

$$\mathbf{Y}_l = F(\mathbf{W}_l \mathbf{Y}_{l-1} + \mathbf{B}_l) \quad (1)$$

where $\mathbf{Y}_l$ and $\mathbf{Y}_{l-1}$ are outputs of l-th and (l-1)-th layer, respectively, $\mathbf{W}_l$ is a weight matrix between (l-1)-th and l-th layer and $\mathbf{B}_l$ is a bias matrix between (l-1)-th and l-th layer.

Authors are with the Faculty of Electronic Engineering,
Aleksandra Medvedeva 14, 18 000 Niš, SERBIA & MONTENEGRO
e-mail: [zlatica, oljap, vera]@elfak.ni.ac.yu

The function F is an activation function of each neuron and, in our case, is linear for input and output layer and sigmoid for hidden layers:

$$F(u) = 1/(1+e^{-u}). \tag{2}$$

The neural network learns relationship among sets of input-output data (training sets) that are characteristics of the device under consideration. First, input vectors are presented to the input neurons and output vectors are computed. These output vectors are then compared with desired values and errors are computed. Error derivatives are then calculated and summed up for each weight and bias until whole training set has been presented to the network. These error derivatives are then used to update the weights and biases for neurons in the model. The training process proceeds until errors are lower than the prescribed values or until the maximum number of epochs (epoch is the whole training set processing) is reached. Once trained, the network provides fast response for various input vectors (even for those not included in the training set) without additional optimizations.

III. Transistor S- Parameters Modeling Based on ANNs

A problem of bias dependent transistor scattering paremeters' prediction for different gate widths is considered.

Recently, bias-dependent scattering parameters' modeling of microwave FETs' by using of neural networks have been proposed [6], [7]. This model is an MLP neural network that gives values of the S- parameters for given bias conditions (dc drain-to-source voltage, V_{ds}, and dc drain-to-source current, I_{ds}) and frequency, f, at its input neurons. The model is valid for the modeled device only. In order to extend its validity to a class of transistors, the transistor gate width, W, is proposed to be an additional input into the neural model. Therefore, the model proposed here (Fig 2) is an MLP neural network with four neurons in the input layer corresponding to:

- gate width W,
- dc drain-to-source voltage V_{ds},
- dc drain-to-source current I_{ds} and
- frequency f.

The output layer consists of eight neurons corresponding to magnitudes and angles of scattering parameters.

Number of the hidden layers can be one or two. The network is trained using S- parameters' data referring to several devices of different gate widths that are made in the same technology. For each device it is necessary to acquire data for certain number of bias points in the operating frequency range. Generally, neural networks with different number of hidden neurons are trained, tested and after their comparison, the network with the best testing results is chosen to be the bias dependent neural noise model for the class of the modeled transistors.

In order to quantify accuracy of the model, average test error (ATE [%]), worst-case error (WCE [%]), and correlation coefficient, r, between the referent and the modeled data are calculated, [1].

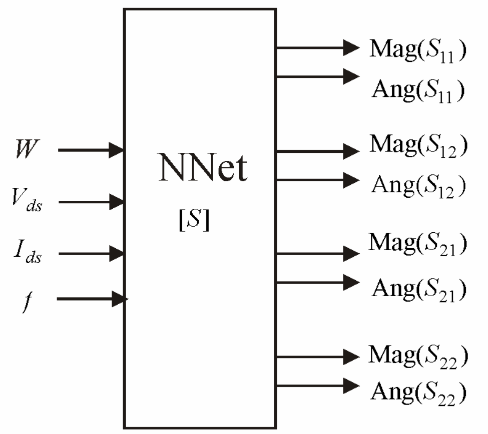


Fig. 2. Proposed ANN model

The Pearson Product-Moment correlation coefficient r is defined by:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \tag{3}$$

where x_i is referent value, y_i is the neural network computed value, $\bar{x}$ is the referent sample mean, and $\bar{y}$ is the neural network sample mean. The correlation coefficient indicates how well the modeled values match the referent values. A correlation coefficient near one indicates an excellent predictive ability, while a coefficient near zero indicates poor predictive ability.

IV. Numerical Results

In this Section, some numerical modeling results are given. Three *Hewlett Packard's* pHEMT devices, ATF3x143 series, were modeled:

- ATF35143 (gate width 400µm),
- ATF34143 (gate width 800µm) and
- ATF33143 (gate width 1600µm).

The S- parameters used as the training data were taken from manufacturer WEB site [8]. These data refer to certain number of bias points. The S- parameters are available in (0.5-18) GHz frequency range for each bias point. The available set of the bias points was divided in two subsets, one used for the training of the neural networks (training set) and the other, smaller one, used for the evaluation of the network generalization capabilities (test set).

Neural networks with four inputs, eight outputs and different number of hidden neurons were trained using the same training set. All of the trained networks were tested on the training set and on the test set. The best results for the S parameters modeling, for both - training and test sets, were obtained by a network with seven neurons in the first and four neurons in the second hidden layer.

During the training process, ANNs are trained to satisfy a priori given accuracy. In order to illustrate quality of learning the training data, a scatter plot of neural model output vs. reference values for the magnitude and angle of S_{12} parameter, referring to the training biases, is shown in Fig. 3 and Fig. 4, respectively. Data points very close to a straight line along the diagonal axis indicate good predictive accuracy.

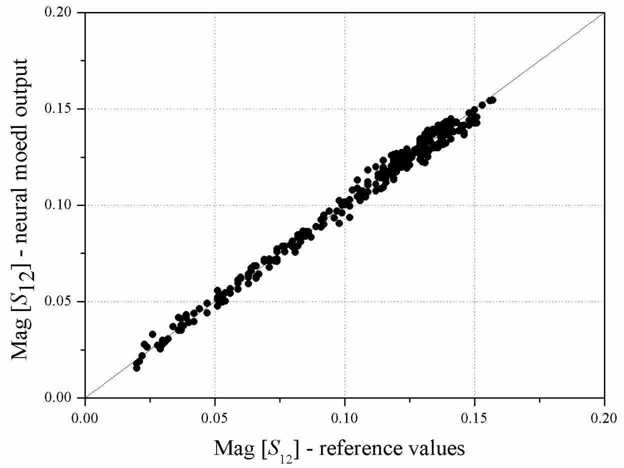


Fig. 3. Magnitude of S_{12} – a scatter plot of neural model output vs. reference values (for the training biases)

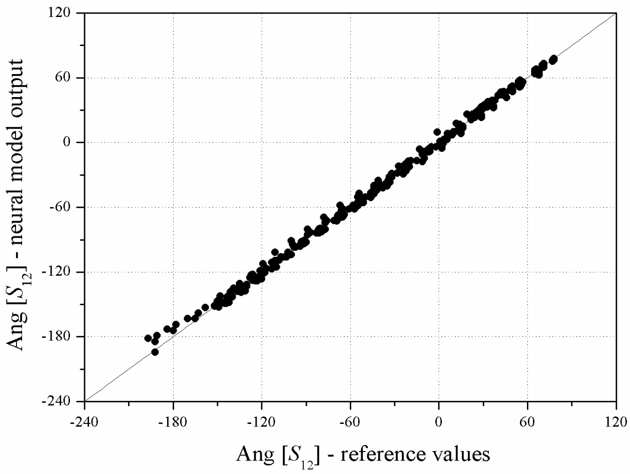


Fig. 4. Angle of S_{12} – a scatter plot of neural model output vs. reference values (for the training biases)

Further, in Tables I and II there are results of the test process for *S*-parameters for the training values as well as for the test values not included in the training set . In Table I there are statistic data for the bias points used in the training process. It can be seen that ATE is lower than 1.9% and WCE is lower than 8.2%, showing that the ANN learnt training data very well. Considering these results and values of correlation coefficient *r* that are very close to one, it is obviously that very good modeling has been achieved, either for input values used as the training data or for those presented to the ANN for the first time during the test process.

Table I. Testing results for the bias points used in the training process

	ATE[%]	WCE[%]	r
$\lvert S_{11}\rvert$	1.8132	7.0316	0.99628
$Ang(S_{11})$	1.2605	4.6728	0.99859
$\lvert S_{21}\rvert$	1.1543	7.0788	0.99708
$Ang(S_{21})$	0.9563	3.0278	0.99912
$\lvert S_{12}\rvert$	1.9544	6.8694	0.99526
$Ang(S_{12})$	0.9373	5.5646	0.99891
$\lvert S_{22}\rvert$	1.4566	8.1647	0.99728
$Ang(S_{22})$	1.7406	7.7594	0.99699

Table II. Testing results for the bias points not used in the training process ATF35143 (2V, 15mA)

	ATE[%]	WCE[%]	r
$\lvert S_{11}\rvert$	1.4878	3.1918	0.99845
$Ang(S_{11})$	1.0633	2.0011	0.99952
$\lvert S_{21}\rvert$	3.2862	10.3900	0.99520
$Ang(S_{21})$	1.1073	2.30264	0.99984
$\lvert S_{12}\rvert$	3.7799	7.00619	0.99828
$Ang(S_{12})$	1.6385	2.76736	0.99973
$\lvert S_{22}\rvert$	3.5683	9.68265	0.99131
$Ang(S_{22})$	3.2738	6.52157	0.99859

As a further illustration, in Fig. 5, there are frequency dependencies of *S*- parameters for two devices referring to bias points not used in the training process. The ANN outputs are denoted by solid line and reference (measured) data by symbols. From a fact that the predicted values match very well to the referent ones, it can be confirmed that very good generalization has been achieved.

V. CONCLUSION

A bias dependent signal model of microwave MESFETs / HEMTs scaling with the device gate width is proposed in this paper. It is a multilayer perceptron neural network trained with the aim to learn *S* parameters' dependence on gate width, bias conditions and frequency. The network is trained using measured values of *S* parameters for several different gate width devices produced in the same technology. Once the network is trained its structure remains unchanged. After the training, the *S* parameters determination is done without additional optimizations.

The proposed neural model is able to predict *S*- parameters with a good accuracy for any given bias and frequency point from the device operating range, either for those used for the network training or for those presented to the network for the first time, as it is illustrated by numerical examples.

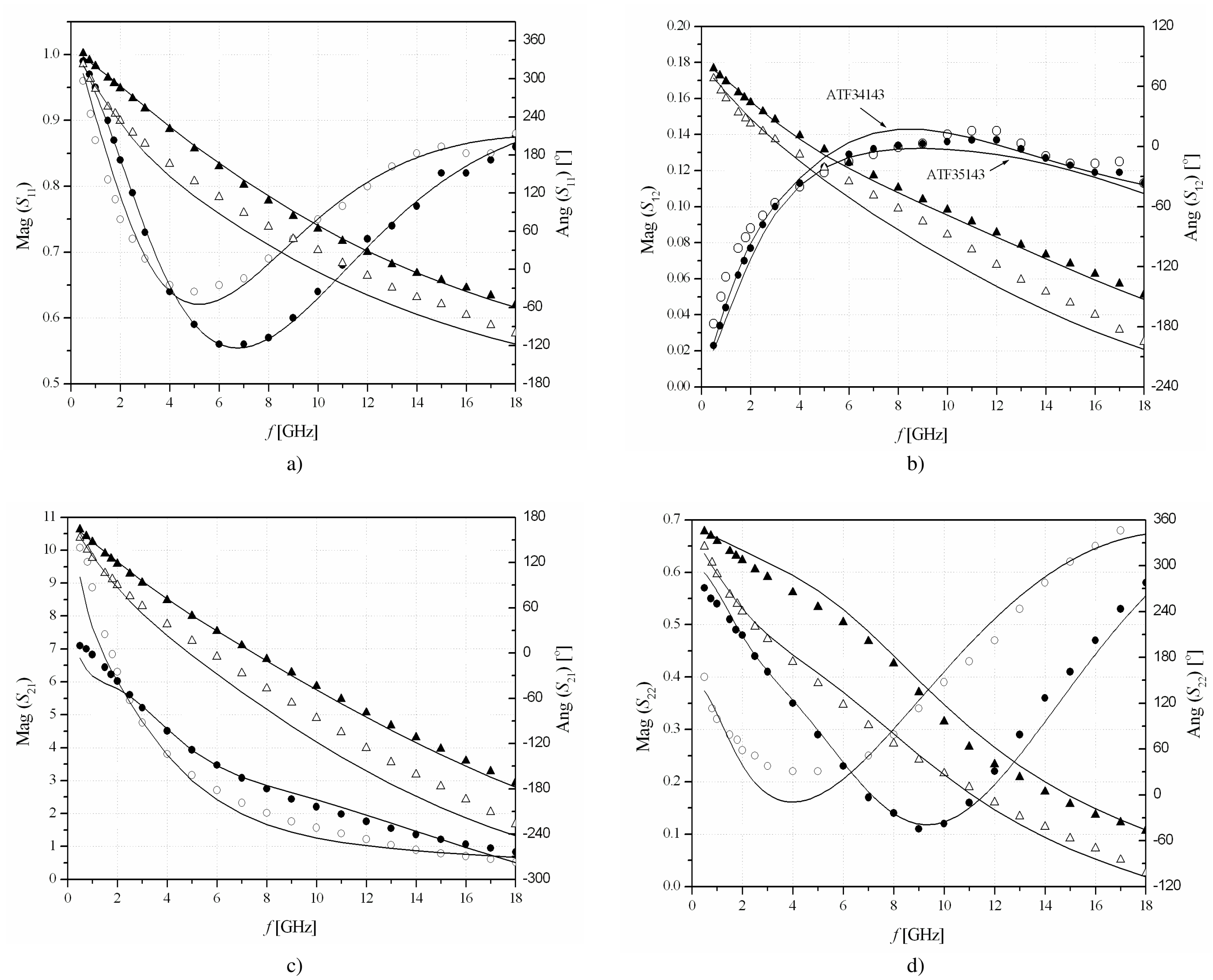


Fig. 5. Magnitudes (circles) and angles (triangles) of scattering parameters for bias points not used for the training process (black symbols – ATF35143 (2V, 15mA) ; white symbols - ATF34143 (3V, 20mA); solid lines – neural model output) a)S_{11} b) S_{12} c) S_{21} d) S_{22}

REFERENCES

[1] Q. J. Zhang, K. C. Gupta, *Neural Networks for RF and Microwave Design,* Artech House, 2000

[2] P. M. Watson, K. C. Gupta, "EM-ANN Models for Microstrip Vias and Intreconnects in Dataset Circuits", *IEEE Trans., Microwave Theory Tech.,* Vol. 44, No-12, 1996, pp. 2495-2503.d

[3] Vijaya K. Devbahaktuni, C. Xi, and Q. J. Zhang, "neural network approach to the modeling of heterojunction bipolar transistors from *S*-parameter data", *Proc. of 28th European Microwave Conf.*, Amsterdam, pp. 306-311, Oct. 1998.

[4] M. Vai, S. Prasad, Neural networks in microwave circuit design – beyond black box models, *Int. J. RF and Microwave Computer-Aided Eng.*, Special issue on Applications of Artificial Neural networks to RF and Microwave Design, 1999, pp. 187-197.

[5] V.Marković, Z.Marinković, "HEMT Noise Neural Model Based on Bias Conditions", *The Int. Journal for Computation and Mathematics in Electrical and Electronic Engineering - COMPEL*, Vol. 23 No.2, 2004, pp.426-435

[6] F.Gunes, H.Torpi, F.Gurgen, "Multidimensional signal-noise neural network model", *Circuits, Devices and Systems, IEE Proceedings,* Vol.145, Iss.2, Apr 1998, pp. 111-117

[7] Z.Marinković, V.Marković, "Prediction of HEMT's Scattering and Noise Parameters Using Neural Networks", *Microwave Review*, Vol. 8, No.1, 2002, pp. 28-31

[8] http://www.semiconductor.agilent.com

29 June – 1 July 2006, Sofia, Bulgaria

2D Electrical Circuit Analysis by Gaussian Procedures

Miodrag V. Gmitrović and Biljana P. Stošić

Abstract – **In this paper two efficient procedures for solving *2D* electrical circuits based on Gaussian elimination procedure are proposed. In order to reconsider the efficiency of the proposed procedures they are compared with two known procedures and the Equivalent Thevenin Sources (*ETS*) method. The voltage vector and impedance matrix of the *ETS* can be calculated by the Gaussian backward elimination procedure.**

Keywords – ***2D* electrical circuit, Gaussian eliminations**

I. INTRODUCTION

Two-dimensional (*2D*) electrical circuit consisting of elements with lumped parameters can be used for modeling of different physical processes. Some of these processes are propagation in microwave transmission lines [1-5], propagation in connections in microelectronic circuits (*RC* transmission lines) [6-7], superconductivity problems [8], diffusion problems in nuclear reactors [9] and etc. A complex *2D* electrical circuit can be modeled in different ways and solved in either time or frequency domain. The main problem is solving the system of linear equations $\mathbf{YU} = \mathbf{I}$, where $\mathbf{I}$ is a source vector, $\mathbf{U}$ is a vector of unknown voltages and $\mathbf{Y}$ is an admittance matrix of the *2D* electrical circuit. It is a very important to solve that equation system by using as less as possible arithmetic operations. There are some different procedures for solving such kind of systems and one of the most used is standard Gaussian elimination procedure (*GSE*) [9-11].

In the previously published papers [1,2], the Equivalent Thevenin Source (*ETS*) method is given. The analysis of *2D* circuit is based on decomposition of the complex circuit structure into cascade-connected ladder subnetworks with *2L* ports. For each ladder subnetwork corresponding transmission matrices are counted and all previously subnetworks are then substituted by its *L ETS*. They represent now the excitation of the next ladder subnetwork. The input and output voltages can be found by successive application of this procedure.

A *2D* electrical circuit can be represented as cascade connection of $2L$ port networks by diakoptics procedure [9], as shown in this paper. Admittance matrix $\mathbf{Y}$ is formed by appropriate choice of unknown voltages. The matrix is a tridiagonal matrix with non-zero submatrices on the main diagonal and on the first diagonals above and below the main diagonal and zero submatrices elsewhere. In order to solve the equation system with this tridiagonal matrix, two efficient procedures based on *GSE* procedure are proposed. The first procedure is Gaussian backward elimination (*GBE*) and the second one is Gaussian direct elimination (*GDE*). The voltage vector $\mathbf{U}_T$ and the impedance matrix $\mathbf{Z}_T$ of *ETS* for whole *2D* electrical circuit can be obtained by *GBE* procedure. In order to reconsider the efficiency of the suggested procedures (*GBE* and *GDE*) they are compared with *GSE* procedure, *ETS* method and the procedure that involves direct inversion of admittance matrix $\mathbf{Y}$ (*INV*).

Miodrag V. Gmitrović and Biljana P. Stošić are with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia and Montenegro.

E-mails: gmitrovic@elfak.ni.ac.yu and bilja@elfak.ni.ac.yu

II. GAUSSIAN PROCEDURES

In Fig.1 a *2D* electrical circuit with losses is shown. The circuit can be excited by several real voltage sources and terminated by several real loads.

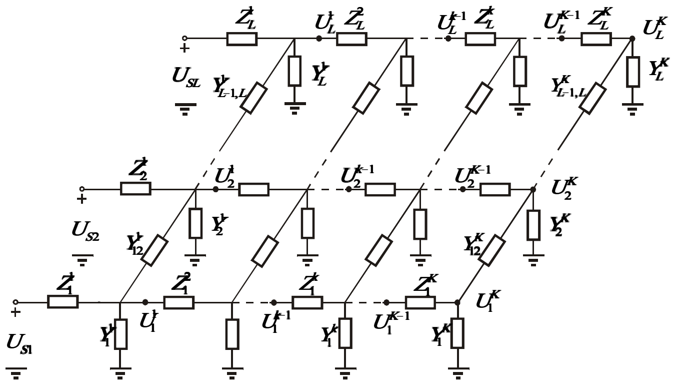


Fig. 1. A *2D* electrical circuit.

A *2D* electrical circuit with known voltage sources at the input ports and all immittances is observed. The goal is to count the node voltages at the output ports.

If the current sources at the input ports are known $I_{Sl} = U_{Sl} / Z_l^1$ than the unknown node voltages U_l^k, $k = 1,2,..., K$, $l = 1,2,..., L$, can be counted from the matrix equation system

$$\mathbf{Y}_K \cdot \mathbf{U}_K = \mathbf{I}_K . \tag{1}$$

In the previous equation system, the current vector is

$$\mathbf{I}_K = \begin{bmatrix} \mathbf{I}^1 & \mathbf{I}^2 & \ldots & \mathbf{I}^K \end{bmatrix}^T = \begin{bmatrix} \mathbf{I}_S & \mathbf{0} & \ldots & \mathbf{0} \end{bmatrix}^T , \tag{2}$$

where

$$\mathbf{I}^1 = \mathbf{I}_S = \begin{bmatrix} I_{S1} & I_{S2} & \ldots & I_{SL} \end{bmatrix}^T \tag{3}$$

and $\boldsymbol{0}$ is a zero vector.

The voltage vector is

$$\mathbf{U}_K = \begin{bmatrix} \mathbf{U}^1 & \mathbf{U}^2 & \ldots & \mathbf{U}^K \end{bmatrix}^T \tag{4}$$

and the admittance matrix is

$$\mathbf{Y}_K = \begin{bmatrix} \mathbf{Y}_{11}^1 & \mathbf{Y}_{12}^1 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{Y}_{21}^2 & \mathbf{Y}_{22}^2 & \mathbf{Y}_{23}^2 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Y}_{32}^3 & \mathbf{Y}_{33}^3 & \mathbf{Y}_{34}^3 & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{Y}_{K-1,K-1}^{K-1} & \mathbf{Y}_{K-1,K}^{K-1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{Y}_{K,K-1}^{K} & \mathbf{Y}_{K,K}^{K} \end{bmatrix} \quad (5)$$

where $\boldsymbol{0}$ is a zero matrix.

In Fig. 2, a l^{th} node of the k^{th} ladder subnetwork is depicted. The immittances Z_l^k, $l=1,2,...,L$, $Y_{l,l+1}^k$, $l=1,2,...,L-1$, correspond to the serial connections of resistors and inductors, and the admittances Y_l^k, $l=1,2,...,L$, $k=1,2,...,K$ correspond to the shunt connections of conductances and capacitors [1-3].

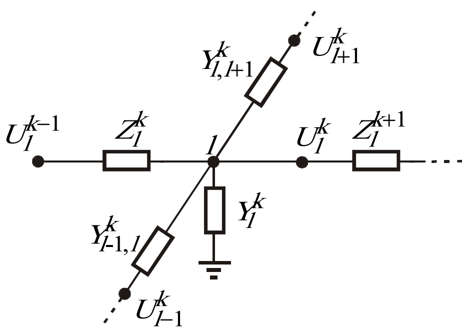


Fig. 2. The l^{th} node of the k^{th} ladder subnetwork.

The admittance matrix $\mathbf{Y}_K$ is a tridiagonal matrix. Also, their submatrices $\mathbf{Y}_{k,k}^k$, $k=1,2,...,K$, are tridiagonal matrices. The elements of these submatrices on the main diagonal are $1/Z_l^k + Y_l^k + 1/Z_l^{k+1} + Y_{l-1,l}^k + Y_{l,l+1}^k$, $l=1,2,...,L$ and $Y_{0,1}^k = Y_{L,L+1}^k = 0$. The elements on the first diagonal above the main diagonal are equal to those ones on the first diagonal below the main diagonal. Those elements are $-Y_{l,l+1}^k$, $l=1,2,...,L-1$. The other elements in these submatrices are zeros.

The submatrices of the matrix $\mathbf{Y}_K$ on the first diagonals above and below the main diagonal are equal $\mathbf{Y}_{k-1,k}^{k-1} = \mathbf{Y}_{k,k-1}^k$, $k=2,3,...,K$. These submatrices are diagonal matrices with the elements on the main diagonal $-1/Z_l^{k+1}$, $l=1,2,...,L$, $k=1,2,...,K-1$ and zeros elsewhere.

The system of matrix equations (1) can be solved by *GSE* [8-10] and in that case the admittance matrix $\mathbf{Y}_K$ is treated as full matrix. The matrix $\mathbf{Y}_K$ (5) has a lot of zero submatrices, and because of that standard Gaussian procedure can be modified in order to use only non-zero submatrices. Two new algorithms that are more efficient for solving matrix equation system (1) are proposed in further text. The matrix equation system (1) can be observed as a set of K equation subsystems. Each subsystem has L linear equations.

2.1. Gaussian backward elimination procedure (GBE)

The *GBE* procedure is a successive solving of the matrix equation system (1) starting from the last K^{th} linear subsystem, which corresponds to the voltage vector $\mathbf{U}^K$, and concluding with the first linear equation subsystem corresponds to the voltage vector $\mathbf{U}^1$. The electrical circuit from Fig.1, for taken voltage vectors $\mathbf{U}^k$, $k=1,2,\dots,K$, can be treated as cascade connection of $2L$ port subnetworks, Fig.3. A tridiagonal admittance matrix $\mathbf{Y}_K$ given by equation (5) can be reduced to a twodiagonal matrix by using *GBE* procedure [9-10]

$$^K\mathbf{Y}_K = \begin{bmatrix} ^1\mathbf{Y}_{11}^1 & \mathbf{Y}_{12}^1 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & ^2\mathbf{Y}_{22}^2 & \mathbf{Y}_{23}^2 & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & ^3\mathbf{Y}_{33}^3 & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & ^{K-1}\mathbf{Y}_{K-1,K-1}^{K-1} & \mathbf{Y}_{K-1,K}^{K-1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & ^K\mathbf{Y}_{K,K}^{K} \end{bmatrix}. \quad (6)$$

The signs in the left upper corner indicate the iteration number. This is valid also for all relations in the further text.

The submatrices above the main diagonal are equal to those ones in the primary matrix (5). The submatrices on the main diagonal are counted by using next expressions

$$^1\mathbf{Y}_{11}^1 = \mathbf{Y}_{11}^1, \quad (7)$$

$$\mathbf{M}_{k,k-1} = \mathbf{Y}_{k,k-1}^k \cdot \left(^{k-1}\mathbf{Y}_{k-1,k-1}^{k-1}\right)^{-1}, \quad (8)$$

$$^k\mathbf{Y}_{k,k}^k = \mathbf{Y}_{k,k}^k - \mathbf{M}_{k,k-1} \cdot \mathbf{Y}_{k-1,k}^{k-1}, \quad (9)$$

where $k=2,3,...,K$.

The current vectors can be counted by the relations

$$^1\mathbf{I}^1 = \mathbf{I}^1, \quad (10)$$

$$^k\mathbf{I}^k = ^{k-1}\mathbf{I}^k - \mathbf{M}_{k,k-1} \cdot ^{k-1}\mathbf{I}^{k-1}, \quad (11)$$

where $k=2,3,...,K$. Having on mind, the current vector $\boldsymbol{I}_K$ given by relation (2), it can be concluded that the current vector $^{k-1}\mathbf{I}^k \equiv \boldsymbol{0}$ for $k=2,3,...,K$. So, the relation (11) can be written as

$$^k\mathbf{I}^k = -\mathbf{M}_{k,k-1} \cdot ^{k-1}\mathbf{I}^{k-1}. \quad (12)$$

Starting by output ports of the *2D* electrical circuit, the voltage vectors can be counted by the next relations

$$\mathbf{U}^K = \left(^K\mathbf{Y}_{K,K}^K\right)^{-1} \cdot ^K\mathbf{I}^K, \quad (13)$$

$$\mathbf{U}^{k-1} = \left({}^{k-1}\mathbf{Y}^{k-1}_{k-1,k-1}\right)^{-1} \cdot \left({}^{k-1}\mathbf{I}^{k-1} - \mathbf{Y}^{k-1}_{k-1,k} \cdot \mathbf{U}^{k}\right), \quad (14)$$

where $k = K, K-1, ..., 3, 2$.

A *2D* electrical circuit, Fig. 1, can be represented as cascade connection of $2L$ port subnetworks as shown in Fig. 3.

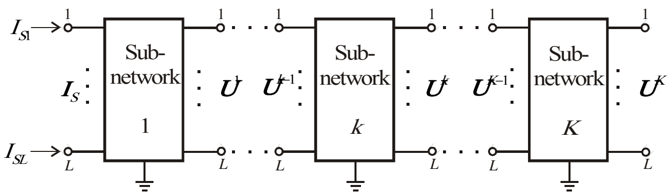


Fig. 3. Cascade connection of subnetworks with *2L* ports.

Cascade connection of *K* ladder subnetworks shown in Fig. 3, can be presented by *ETS* [2] as shown in Fig. 4.

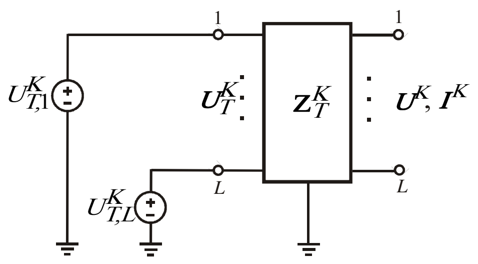


Fig. 4. *ETS* of voltage $\mathbf{U}_T^K$ and impedance $\mathbf{Z}_T^K$.

From the Fig.4 can be concluded that the voltage vector at the output ports is

$$\mathbf{U}^K = \mathbf{U}_T^K - \mathbf{Z}_T^K \cdot \mathbf{I}^K, \quad (15)$$

where $\mathbf{U}_T^K$ is the voltage vector and $\mathbf{Z}_T^K$ is the impedance matrix of the *ETS*. The impedance matrix of the *ETS*

$$\mathbf{Z}_T^K = ({}^K\mathbf{Y}^K_{K,K})^{-1} \quad (16)$$

is counted using the submatrix ${}^K\mathbf{Y}^K_{K,K}$ from the matrix ${}^K\mathbf{Y}_K$ (6). The voltage vector $\mathbf{U}_T^K$ for $\mathbf{I}^K = \mathbf{0}$ is equal to the voltage vector $\mathbf{U}^K$ obtained by relation (13). The calculated voltage vector $\mathbf{U}_T^K$ and impedance matrix $\mathbf{Z}_T^K$ can be used as inputs for the next ladder subnetwork in cascade connection. In this way can be solved cascade connection of ladder subnetworks with different number of ports. Such type of subnetworks can be used for successful analysis of different discontinuities in microwave transmission lines.

2.2. Gaussian direct elimination procedure (GDE)

The *GDE* procedure is a successive solving of the matrix equation system (1) starting from the first linear equation subsystem corresponds to the voltage vector $\mathbf{U}^1$ and concluding with the last K^{th} linear subsystem which corresponds to the voltage vector $\mathbf{U}^K$. In other words, *GDE* procedure is used for calculating the voltages from the matrix system (1) starting from the voltage vector $\mathbf{U}^1$. That vector presents the voltages on the output ports of the first subnetwork in cascade connection, Fig. 3. By the *GDE* procedure, the matrix $\mathbf{Y}_K$ is reduced to twodiagonal matrix of shape

$${}^K\mathbf{Y}_K = \begin{bmatrix} {}^K\mathbf{Y}^1_{11} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{Y}^2_{21} & {}^{K-1}\mathbf{Y}^2_{22} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Y}^3_{32} & {}^{K-2}\mathbf{Y}^3_{33} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & {}^2\mathbf{Y}^{K-1}_{K-1,K-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{Y}^K_{K,K-1} & {}^1\mathbf{Y}^K_{K,K} \end{bmatrix} \quad (17)$$

The matrices below the main diagonal are equal to those ones in the primary matrix (5). The matrices on the main diagonal are counted by next expressions

$${}^1\mathbf{Y}^K_{K,K} = \mathbf{Y}^K_{K,K}, \quad (18)$$

$$\mathbf{M}_{k,k+1} = \mathbf{Y}^k_{k,k+1} \cdot \left({}^{K-k}\mathbf{Y}^{k+1}_{k+1,k+1}\right)^{-1}, \quad (19)$$

$${}^{K-k+1}\mathbf{Y}^k_{k,k} = \mathbf{Y}^k_{k,k} - \mathbf{M}_{k,k+1} \cdot \mathbf{Y}^{k+1}_{k+1,k}, \quad (20)$$

where $k = K-1, K-2, ..., 2, 1$.

The voltages are here counted starting from the voltage vector on input ports by using the next relations

$$\mathbf{U}^1 = \left({}^K\mathbf{Y}^1_{11}\right)^{-1} \cdot \mathbf{I}^1, \quad (21)$$

$$\mathbf{U}^k = -\left({}^{K-k+1}\mathbf{Y}^k_{k,k}\right)^{-1} \cdot \mathbf{Y}^k_{k,k-1} \cdot \mathbf{U}^{k-1}, \quad (22)$$

where $k = 2, 3, ..., K$.

The *GDE* procedure in regards of the *GBE* procedure does not perform changes in current vector and the matrix equation system (1) is solved with less number of arithmetic operations.

III. Efficiency Analysis

In order to reconsider the efficiency of the suggested procedures (*GBE* and *GDE*), *2D* electrical circuits of various complexities are solved. The equivalent circuit composed as cascade-connected multi-port subnetworks shown in Fig.3 is analysed. The number of input ports is $L = 10$ and the number of cascade-connected subnetworks is $K = 1, 2, ..., 200$. In this case, the largest equation system has 2000 unknown voltages. The efficiency-testing program is done in MATLAB [12] on PC 2.4 *GHz*.

Graphs of time needed for solving the equation system versus the number of networks in cascade connection K, are depicted in Fig.5. The time needed for the forming of admittance matrix Y_K of the *2D* circuit is included in the time needed for solving the equation system for all procedures. The MATLAB built-in functions are used for standard Gaussian elimination procedure (*GSE*), i.e. $\mathbf{U} = \mathbf{Y} \backslash \mathbf{I}$, and matrix inversion procedure (*INV*), i.e. $\mathbf{U} = \mathrm{inv}(\mathbf{Y}) \cdot \mathbf{I}$. The expressions given in the papers [1-3] are used for the *ETS* method (*ETS*).

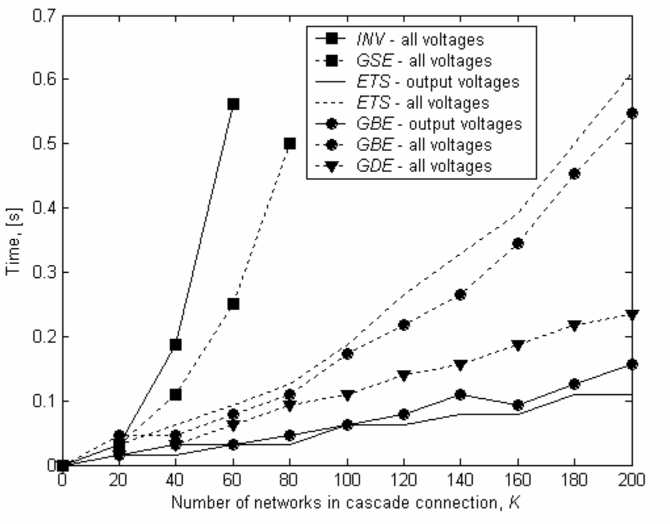


Fig. 5. Curve efficiency graph.

It can be inferred that the matrix inversion procedure requires the longest time needed for solving the equation system. In that case, a full admittance matrix $\mathbf{Y}_K$ is solved. The full admittance matrix is also solved by Gaussian standard procedure, but it is more efficient than the matrix inversion procedure. In MATLAB, the computations involving the backslash operator ($\backslash$) require less computer time, less memory and have better error detection properties than the other one which forms the direct inverse (inv).

In the case when voltages in all nodes $\mathbf{U}_K$ are to be calculated, the most efficient method of *INV*, *GSE*, *ETS*, *GBE* and *GDE*, is the suggested procedure *GDE* since it doesn't require the calculation of the current vector $\mathbf{I}_K$. In the case when only output voltages $\mathbf{U}^K$ are to be calculated, the suggested procedure *GBE* and the *ETS* method require almost the same time for solving *2D* circuit. The *ETS* method doesn't require the calculation of the current vectors.

It is important to say that *GDE* procedure requires calculation of all node voltages in order to get the voltage vector $\mathbf{U}^K$ at the output ports. The essential advantage of the suggested *GBE* procedure is direct calculation of the output voltage vector without calculation of voltages in all circuit nodes. The other advantage of the *GBE* procedure is direct calculation of the *ETS* elements, i.e. the voltage vector $\mathbf{U}_T^K$ and the impedance matrix $\mathbf{Z}_T^K$.

IV. CONCLUSION

Two efficient procedures for solving matrix equation systems (1), which admittance matrix is a tridiagonal matrix, are proposed in this paper. Such admittance matrix is found for the *2D* electrical circuits, Fig. 1, represented as cascade-connected ladder $2L$ port subnetworks with voltages assigned as shown in Fig. 3.

It is shown that in the case when all voltages are to be calculated, the most efficient procedure of all is the *GDE* procedure. The most common case in practice is the calculation of the output voltages $\mathbf{U}^K$ only. In that case, the most efficient procedures are the *GBE* procedure and the *ETS* method [2]. The *GBE* procedure can be used for calculating the voltage vector $\mathbf{U}_T^K$ and the impedance matrix $\mathbf{Z}_T^K$ of the known *2D* electrical circuit. The suggested *GBE* procedure can be used successfully for solving complex microwave circuits containing transmission lines.

REFERENCES

[1] B. P. Stojanović and M. V. Gmitrović, "Modelling and Analysis of Cascade-Connected Planar Transmission Lines", *XII International Symposium on Theoretical Electrical Engineering - ISTET'03*, Poland, Warsaw, July 6-9, 2003, Volume II, pp. 473-476.

[2] B. P. Sto{i}, *Frequency Analysis of Planar Microwave Circuits by Thevenin Source Method,* Master thesis, Faculty of Electronic Engineering, University of Ni{, Ni{, February 2004, (in Serbian).

[3] B. P. Stošić and M. V. Gmitrović, "Frequency Analysis of Cascade-Connected Planar Transmission Lines by *ETS* Method", *Microwave Review*, No. 1, Vol. 10, June 2004, Serbia and Montenegro, pp. 20-29.

[4] W. K. Gwarek, "Analysis of Arbitrarily Shaped Two-Dimensional Microwave Circuits by Finite-Difference Time-Domain Method", *IEEE Transactions on Microwave Theory and Techniques,* Vol.36, No.4, April, 1988, pp. 738-744.

[5] L. N. Merugu and V. F. Fusco, "Concurent Network Diakoptics for Electromagnetic Field Problems", *IEEE Transactions on Microwave Theory and Techniques,* Vol.45, No.7, July, 1997, pp. 1086-1092.

[6] W. - K. Chen, *The Circuits and Filters Handbook*, CRC Press, 1995.

[7] M. V. Gmitrović, "Efikasan postupak za izračunavanje odziva u *RC* mrežama oblika stabla", *XXXI Jugoslovenska konferencija ETAN*, str. III.283-III.289, Bled, 1987 (in Serbian).

[8] C. E. Tong, L. Chen and R. Blundell, "Theory of Distributed Mixing and Amplification in a Superconducting Quasi-Particle Nonlinear Transmission Line", *IEEE Transactions on Microwave Theory and Techniques,* Vol.41, No.4, April, 1993, pp. 708-716.

[9] G. Kron, *Issledovanie sloænih sistem po åastàm – Diakoptika*, Nauka, Moskva, 1972.

[10] V. B. Litovski, *Projektovanje elektronskih kola*, DGIP "Nova Jugoslavija" - Vranje, Niš, 2000 (in Serbian).

[11] H. Kremer, *Numerical Analysis of Linear Networks and Systems*, Artech House, London, 1987.

[12] MATLAB – The Language of Technical Computing, Version 6.5.0.180913a, Release 13, June 18, 2002, The Math Works Inc. 1984-2002.

Metamaterial in Finline Configuration

Antoniya Georgieva[1]

Abstract – **This paper presents the results of experimental study of finlines, periodically loaded by split-ring resonators (SRR) and wires. These are the common elements used to construct metamaterials – artificial materials having simultaneously negative effective magnetic permeability and dielectric permittivity. The transmission measurements of SRR-and-wire-loaded finline, carried out around 3 GHz, show clearly a pass-band emergence in the region of non-transmission of the SRR-loaded and wires-loaded finlines. This behaviour, predicted by theory, is considered an indirect evidence of simultaneously negative effective permittivity and permeability of the structure.**

Keywords – **metamaterial, split-ring resonator, finline**

I. Introduction

Metamaterials are a novel kind of artificial materials that exhibit negative effective magnetic permeability and dielectric permittivity. The idea that a negative macroscopic parameters media would support wave propagation belongs to V.Veselago, who studied it theoretically in 1968 [1]. As such media are not readily found in nature, the subject was not further developed until recently, when Pendry [2] suggested that negative effective permeability in a narrow frequency band could be achieved through artificial magnetism of resonant diamagnetic particles known as split-ring resonators (SRR). Subsequently, Shelby [3] used a combination of SRR array and straight thin wires array, the latter claimed to have negative effective permittivity up to some frequency, to demonstrate for the first time a passband in the frequency region where $\varepsilon_{r,\ eff}$ and $\mu_{r,\ eff}$ are simultaneously negative. Other experimental studies followed employing mainly free-space or parallel-plate waveguide metamaterial configuretions. Although these experiments are highly valuable for understanding the electromagnetic phenomena in metamaterials, the structures used are three-dimensional and of considerable size, and are not feasible for practical microwave implementations. Therefore, a number of subsequent studies focus on microstrip and waveguide metamaterial implementations of filters and other devices.

In this paper, we employ the main principles of metamaterial construction to study a transmission line, loaded with SRR and wires, that exhibits a transmission passband due to negative effective permittivity and permeability. The transmission line itself is an unilateral finline – a slotline inside a rectangular waveguide. The main advantages of this configuration are:

- The field is concentrated inside the slot providing strong excitation of the SRRs;
- The line is shielded by the waveguide so there are no radiation losses;
- The structure is inherently two-dimensional and therefore can include SMD elements such as capacitors, inductors, diodes, etc.;
- Small size and easy fabrication.

II. Metamaterial-Loaded Finline

The finline is actually a slotline placed in the E-plane in the middle of a standard rectangular waveguide with dimensions of 34x72 mm. The frequency range of waveguide operation is 2.60 ÷ 4.16 GHz. The finline is etched on a standard PCB substrate (FR4) with thickness of 1.5 mm and relative permittivity $\varepsilon_r = 4.4$. This kind of substrate is chosen for its low price and availability considerations, but, as this material is not intended for microwave applications because of its high loss tangent, it degrades to some extent the finline performance. In principle, high quality microwave materials (for example, Rogers Duroid 5880™) should be used.

The finline dimensions are:

- Gap between the strips: 3 mm;
- Full length (with taper sections): 16 cm;
- Slotline length (without taper sections): 8 cm.

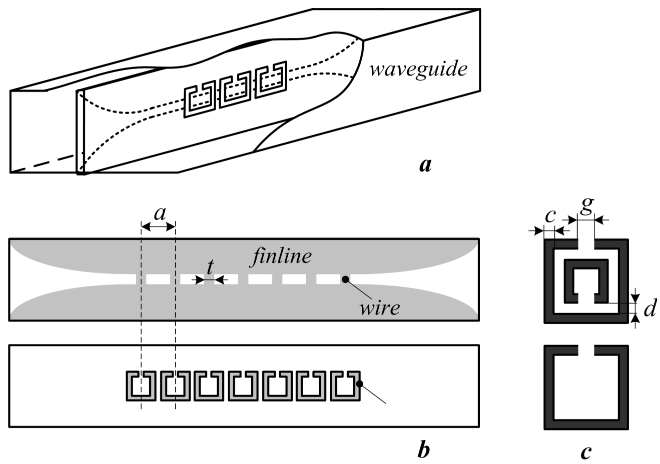


Fig. 1. Full waveguide structure (a), loaded finline – front- and backside view (b) and the two types of SRR used.

To ensure good electrical contact between the finline strips and the waveguide, springing elements (bendable wires) have been soldered in longitudal direction at the finline edges. The full waveguide structure is presented in Fig.1(a).

[1] Antoniya R. Georgieva is with the Faculty of Electronics, Radioengineering Department, Technical University of Varna, 1 Studentska str., 9000 Varna, Bulgaria, E-mail: antonia_ross@yahoo.com

The finline is loaded periodically by wires, as shown in Fig.1(b). The SRR are etched on the back side of the substrate (opposite to the strips) to ensure that they are excited by the transversal magnetic field component. The operation of the SRR is based on the Lenz law – the external field, coaxial with the split-ring resonator, excites currents along the SRR strips. The magnetic field of the excited current is opposite to the external field so the particle is diamagnetic in nature. At the resonance frequency of the particle the excited magnetic field is greatly enhanced and, as it opposes the incident field, the macroscopic effect is negative permeability.

Two types of structures have been constructed and studied. Type 1 consists of double SRR (Fig.1c) with dimensions l=7.5mm, d=0.2 mm, c = 0.9 mm, g = 0.9 mm, which are resonant around 3 GHz. The wires are 2 mm wide. The period is a=9mm and the finline is loaded by 9 SRR/wires. Type 2 employs single SRR (Fig.1c) with dimensions l = 7mm, c = 0.7 mm, g=1 mm and resonance frequency of 3.66 GHz. The period is a=8 mm and the number of SRR/wires is 5.

The finline characteristic impedance, effective dielectric permittivity and propagation constant as a function of frequency have been assessed using the formulas in [4]. The characteristic impedance varies from 140.8 to 143.2 Ω in the frequency range 2.7 ÷ 3.9 GHz. The calculated finline wavelength for the highest frequency of interest (3.9 GHz) is 5.9 cm, so the SRR and wires sizes and period are much smaller than the finline wavelength.

III. Measurement Setup

The purpose of the experiment was to measure the level of the signal transmitted through different structures – SRR-only loaded finline, wires-only loaded finline, and the combination of SRR-and-wire loaded finline. As each media separately exhibits negative effective permittivity/permeability in certain frequency range, which corresponds to non-transmission, a transmission band should occur, according to theory, in the range where $\varepsilon_{r,\,eff}$ and $\mu_{r,\,eff}$ are simultaneously negative.

The measurement setup consists of: a signal generator Rohde-Shwartz in the frequency range of 2.5 ÷ 4 GHz, which excites a TE_{10} mode in the waveguide; a spectral analyzer to measure the transmission level.

IV. Results

Fig. 2 and 3 present the measured level of the transmitted signal in function of frequency. Fig. 2 corresponds to type 1 structure with double SRR and Fig.3 – to single SRR structure type 2. In each figure four graphs are plotted – transmission through the finline (for reference), SRR-loaded finline, wires-loaded finline and SRR-and-wire-loaded finline. The SRR resonance causes transmission level degradation between 2.81-3.12 GHz for double SRRs (Fig. 2) and between 3.55-3.78 GHz for the single SRR structure (Fig. 3). Wires-loaded finline in both cases exhibits poor transmission (minimum 25dB below the finline transmission level). When combining SRRs and wires, an enhanced transmission band (designated with arrows) occurs near the band of non-transmission of the SRR-loaded finline.

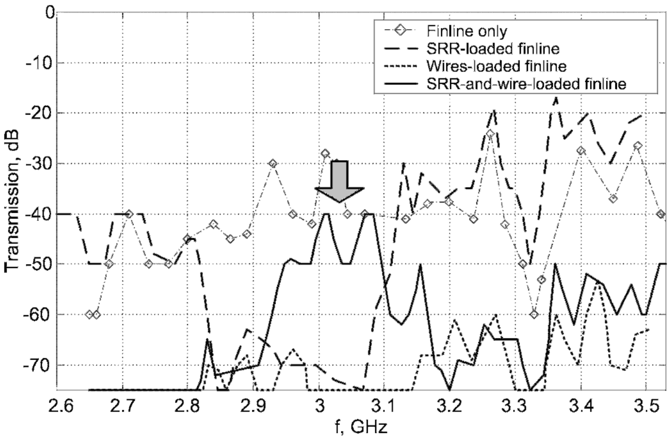


Fig. 2. Measured transmission through structure type 1

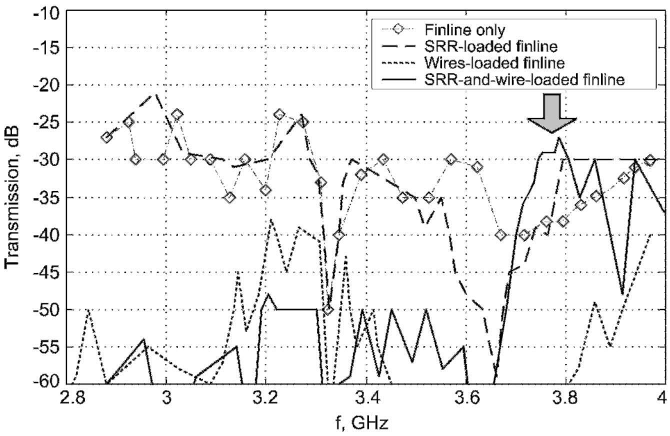


Fig. 3. Measured transmission through structure type 2

V. Conclusion

The paper presented experimental results of transmission measurement of metamaterial-loaded finline. An enhanced transmission band is observed near the frequency of SRR resonance. This behaviour is in agreement with theory [1] and other experimental studies [3]. The conducted experiments, however, do not explicitly prove the existence of negative effective parameters. More detailed analysis should include phase measurement of the transmitted signal through finlines of various length.

References

[1] V. Veselago, "The electrodynamics of substances with simultaneously negative values of ε and μ", Sov. Phys. Usp., vol. 10, No.4, 1968, pp. 509-514

[2] J. Pendry, A. Holden, D. Robbins, W. Stewart, "Magnetism from conductors and enhanced nonlinear phenomena", IEEE Trans. MTT, vol. 47, Nov. 1999, pp. 2075-2084

[3] R. Shelby, D. Smith, S. Nemat-Nasser, S. Schultz, "Microwave transmission through a two-dimensional, isotropic, left-handed metamaterial", Appl. Phys. Lett., Vol. 78, No. 4, Jan. 2001, pp. 489-491

[4] K. Gupta, *Microstrip Lines and Slotlines*, Artech House, 1996

Diversity System with *L* Brunches for the Demodulation of n-FSK Signals

Dragana Krstić[1], Mihajlo Stefanović[1]
and Petar Spalević[2]

***Abstract* – The diversity system for the demodulation of n-ary digital frequently modulated signal (n-FSK) with *L* brunches is considered in this paper. Such receiver consists of *n* brunches. The signals from the corresponding brunches of all receivers are added. The combiner determines the maximal signal from EGC output. The joint probability density of the signal and the signal derivative at the combiner output is determined in this paper. The average number of the signal crossing rate at the system output can be calculating by this joint probability density. In the paper the probability density of the signal at the combiner output is also calculated.**

***Key words* – diversity system, n-FSK signals, probability density, EGC combiner**

I. Introduction

In this paper the statistical characteristics of the signal at the output of the diversity system with *L* diversity brunches will be calculated. This system uses for the n-FSK signal demodulation. The signals from the same brunches of each receiver are added. The combiner determines the maximal signal at the output. These diversity systems apply in wireless telephony. The Gaussian noise appears at the receiver input. This noise becomes the narrowband Gaussian noise at the narrowband filter output. The post detection filtering is used. Some models of the diversity systems which are used for the binary FSK signal demodulation are considered in paper [1]. One of models considered is: the combiner determines maximal value of the signal from all brunch outputs of the all receiver. The combiner we consider in this paper is more efficient because of the addition of signals from corresponding brunches of each receiver. In this paper the n-FSK signal will be analyzed whereas in paper [1] the BFSK signal is analyzed. The demodulation in the receivers is coherent. Each receiver consists of narrowband filter and of the correlator. The correlator consists of the multiplier and the low pass filter. We consider the extraction of the referent carrier as ideal. In this paper the error probability will be calculated, the probability density of the signal at the system output and the joint probability density of the signal and its derivative at the system output will be determined also.

We can calculate the number of the signal crossing rate at the system output by the joint probability density of the signal and its derivative.

[1] Dragana Krstić and Mihajlo Stefanović are with Faculty of Electronic Engenering, Niš, Aleksandra Medvedeva 14, SCG, e-mail: dragana@elfak.ni.ac.yu
[2] Petar Spalević is with Faculty of Tehnical Science, Kosovska Mitrovica, e-mail: petarspalevic@yahoo.com

II. The Error Probability Calculation

The model of the diversity system which is used for the n-FSK signal demodulation is given at Fig. 1. The diversity system has *L* diversity brunches. Such receiver has *n* brunches. The Gaussian noise appears at all receiver inputs. The narrow band Gaussian noise is obtained at the narrow band filter outputs. The signals at the first receiver brunches outputs are $z_{111}, z_{121}, ..., z_{1n1}$ for the hypothesis H_l. The signals at the second receiver brunches outputs are $z_{112}, z_{122}, ..., z_{1n2}$ etc. The signals derivatives are $\dot{z}_{111}, \dot{z}_{121}, ..., \dot{z}_{1n1}$, $\dot{z}_{112}, \dot{z}_{122}, ..., \dot{z}_{1n2}$, etc. The signals at the inputs of the receiver for the hypothesis H_0 are:

$$\begin{aligned} r_1(t) &= A\cos\omega_1 t + n_1(t) \\ r_2(t) &= A\cos\omega_2 t + n_2(t) \\ &\cdots \\ r_n(t) &= A\cos\omega_n t + n_n(t) \end{aligned} \qquad (1)$$

A is the signal amplitude, ω_l is the signal frequency for the hypothesis H_l. The Gaussian noises appearing at the diversity brunches inputs are $n_1(t), n_2(t), ..., n_n(t)$. These noises have the mean values zero and variances $\sigma_1^2, \sigma_2^2, ..., \sigma_n^2$. The narrow band Gaussian noise at the filter output in l-th diversity brunch and k-th receiver brunch is

$$n_{lk} = x_{lk}\cos\omega_k t + y_{lk}\sin\omega_k t,$$

$$l = 1,2,..., L; \quad k = 1,2,..., n \qquad (2)$$

where x_{lk} and y_{lk} are the Gaussian components in quadrature, independent at the same time instant.

The signals at the corelator outputs are:

$$\begin{array}{llll} z_{111} = A + x_{11} & z_{112} = A + x_{21} & \cdots & z_{11L} = A + x_{L1} \\ z_{121} = x_{12} & z_{122} = x_{22} & \cdots & z_{12L} = x_{L2} \\ \cdots & & & \\ z_{1n1} = x_{1n} & z_{1n2} = x_{2n} \cdots & & z_{1nL} = x_{Ln} \end{array} \qquad (3)$$

The signals at the EGC combiner outputs are:

$$\begin{aligned} z_{11} &= z_{111} + z_{112} + \cdots + z_{11L} = L \cdot A + x_{11} + x_{21} + \cdots + x_{L1} \\ z_{12} &= x_{12} + x_{22} + \cdots + x_{L2} \\ &\cdots \\ z_{1L} &= x_{1n} + x_{2n} + \cdots + x_{Ln} \end{aligned} \qquad (4)$$

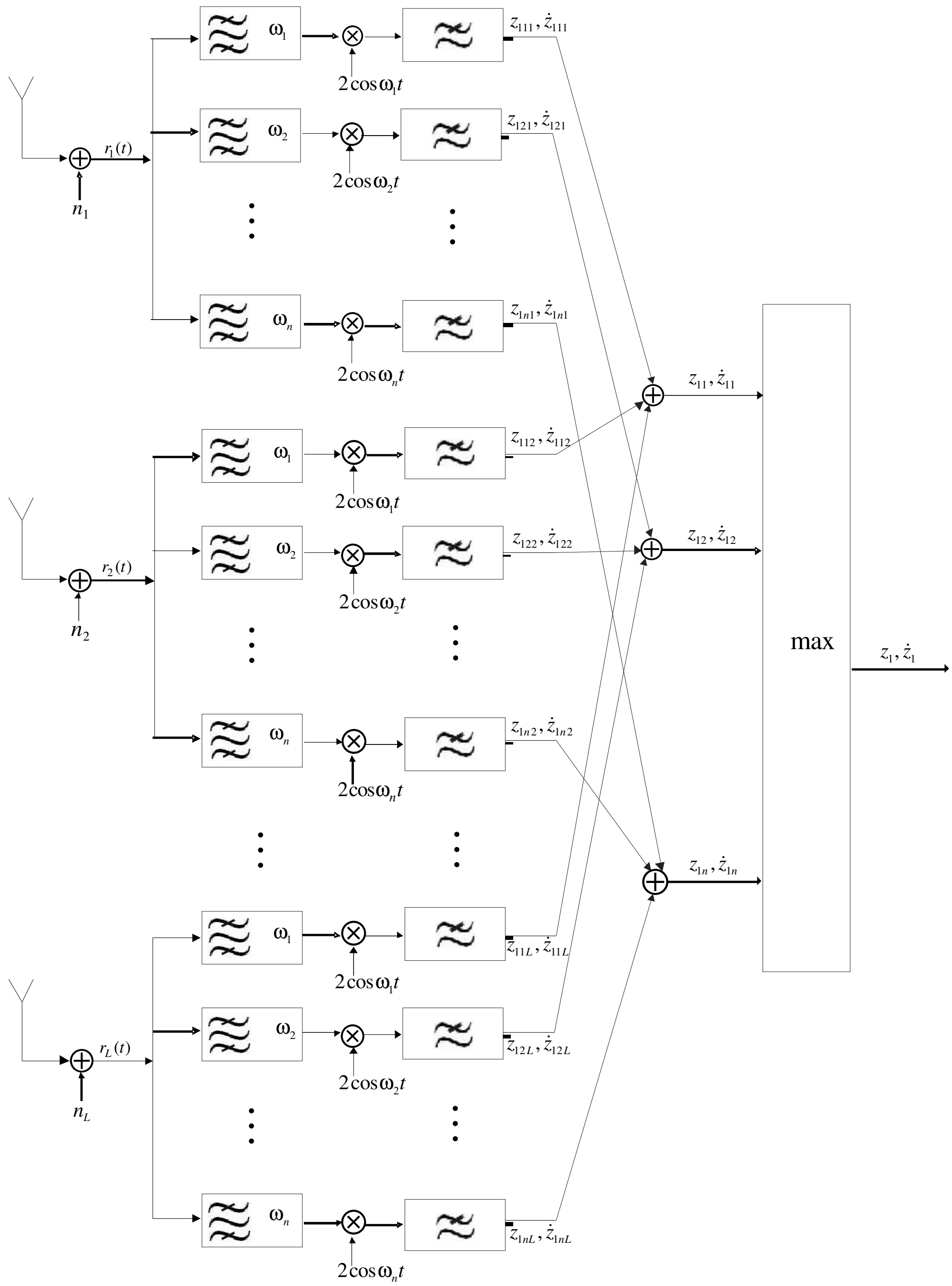


Fig. 1. The model of the diversity system

The signals z_{11l}, $l = 1,2,...,L$ have Gaussian probability density function with mean values A and variances $\sigma_1^2, \sigma_2^2, ..., \sigma_L^2$.

$$p_{z_{11l}}(z_{11l}) = \frac{1}{\sqrt{2\pi}\sigma_l} e^{-\frac{(z_{11l}-A)^2}{2\sigma_l^2}} \tag{5}$$

The signals z_{1kl}, $k = 1,2,3,...,n;\ l = 1,2,...,L$ have mean values zero and variances $\sigma_1^2, \sigma_2^2, ..., \sigma_L^2$.

$$p_{z_{1kl}}(z_{1kl}) = \frac{1}{\sqrt{2\pi}\sigma_l} e^{-\frac{(z_{1kl})^2}{2\sigma_l^2}} \tag{6}$$

The random variable z_{11} has mean value $\overline{z_{11}} = L \cdot A$ and variance

$$\overline{\left(z_{11} - \overline{z_{11}}\right)^2} = \sigma^2 = \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_L^2 \tag{7}$$

The random variables $z_{12}, z_{13}, ..., z_{1L}$ have mean values zero and variances:

$$\overline{\left(z_{1l} - \overline{z_{1l}}\right)^2} = \sigma^2 = \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_L^2, \qquad l = 2,3,..., L \tag{8}$$

Because of that we have:

$$p_{z_{11}}(z_{11}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z_{11}-LA)^2}{2\sigma^2}} \quad (9)$$

$$p_{z_{1l}}(z_{1l}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{z_{1l}^2}{2\sigma^2}} \quad , \quad l = 2,3,...,L \quad (10)$$

The probability for correct decision for hypothesis H_1 is

$$P_C = P\{z_{11} > z_{12}, z_{11} > z_{13}, ..., z_{11} > z_{1L}\} \quad (11)$$

This probability is:

$$P_C = \int_{-\infty}^{\infty} p_{z_{11}}(z_{11}) \cdot dz_{11} \int_{-\infty}^{z_{11}} p_{z_{12}}(z_{12}) \cdot dz_{12} \cdot$$

$$\cdot \int_{-\infty}^{z_{11}} p_{z_{13}}(z_{13}) \cdot dz_{13} \cdots \int_{-\infty}^{z_{11}} p_{z_{1L}}(z_{1L}) \cdot dz_{1L} =$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z_{11}-LA)^2}{2\sigma^2}} \cdot \left(\int_{-\infty}^{z_{11}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{z_{12}^2}{2\sigma^2}} \cdot dz_{12} \right)^{L-1} \cdot dz_{11} \quad (12)$$

The conditional error probability for hypothesis H_1 is:

$$P_e = 1 - P_C = 1 - \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z_{01}-LA)^2}{2\sigma^2}} \cdot$$

$$\cdot \left(\int_{-\infty}^{z_{01}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{z_{02}^2}{2\sigma^2}} \cdot dz_{02} \right)^{L-1} \cdot dz_{01} \quad (13)$$

We can calculate now in the similar way the other conditional error probabilities for hypotheses H_l and then the error probability.

III. STATISTICAL CARACTERISTICS OF THE SIGNAL AT THE DIVERSITY SYSTEM OUTPUT

For the hypothesis H_1, the signal at the system output is

$$z_1 = \max\{z_{11}, z_{12}, ..., z_{1n}\} \quad (14)$$

The probability density of the signal z_1 is

$$p_{z_1}(z_1) = \sum_{i=1}^{n} p_{z_{1i}}(z_1) \cdot \prod_{\substack{j=1 \\ j \neq i}}^{n} F_{z_{1j}}(z_1) \quad (15)$$

where $F_{z_{1j}}(z_1)$ is cumulative probability of the random variable z_{1j}. The cumulative probability of the random variable z_{11} is

$$F_{z_{11}}(z_{11}) = \int_{-\infty}^{z_{11}} p_{z_{1i}}(x) \cdot dx = \int_{-\infty}^{z_{11}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-LA)^2}{2\sigma^2}} dx \quad (16)$$

If: $$t = \frac{x - LA}{\sqrt{2}\sigma}, \qquad dt = \frac{1}{\sqrt{2}\sigma} dx \quad (17)$$

we obtain

$$F_{z_{11}}(z_{11}) = \int_{-\infty}^{\frac{z_{11}-LA}{\sqrt{2}\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = Q\left(\frac{z_{11} - LA}{\sqrt{2}\sigma} \right) \quad (18)$$

where

$$Q(x) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{x} e^{-t^2} dt \quad (19)$$

The cumulative probability of the random variables $z_{1k}, k = 2,3,...,n$ is

$$F_{z_{1k}}(z_{1k}) = Q\left(\frac{z_{1k}}{\sqrt{2}\sigma} \right) \quad (20)$$

The probability density of the random variable z_1 can be determined at the manner below. The joint cumulative probability of the random variables $z_{11}, z_{12}, ..., z_{1n}$ is

$$F_{z_{11}, z_{12}, ..., z_{1n}}(z_{11}, z_{12}, ..., z_{1n}) = \prod_{k=1}^{n} F_{z_{1k}}(z_{1k}) \quad (21)$$

The cumulative probability of the random variable z_1 is

$$F_{z_1}(z_1) = F_{z_{11}, z_{12}, ..., z_{1n}}(z_{11}, z_{12}, ..., z_{1n}) = \prod_{k=1}^{n} F_{z_{1k}}(z_{1k}) \quad (22)$$

The probability density of the random variable z_1 is

$$p_{z_1}(z_1) = \frac{dF_{z_1}(z_1)}{dz_1} = \frac{d}{dz_1}\left(\prod_{k=1}^{n} F_{z_{1k}}(z_1) \right) \quad (23)$$

By using Leibniz formula we can obtain the same expression for the probability density of the signal at the combiner output as previously is derived. The joint probability density of the envelope and the signal derivative at the output is

$$p_{z_1, \dot{z}_1}(z_1, \dot{z}_1) = \sum_{i=1}^{n} p_{z_{1i}\dot{z}_{1i}}(z_1, \dot{z}_1) \cdot \prod_{\substack{j=1 \\ j \neq i}}^{n} F_{z_{1j}}(z_1) \quad (24)$$

where

$$p_{z_{11}, \dot{z}_{11}}(z_{11}, \dot{z}_{11}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z_{11}-A)^2}{2\sigma^2}} \frac{1}{\sqrt{2\pi}\beta} e^{-\frac{\dot{z}_{11}^2}{2\beta^2}}$$

$$p_{z_{12}, \dot{z}_{12}}(z_{12}, \dot{z}_{12}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{z_{12}^2}{2\sigma^2}} \frac{1}{\sqrt{2\pi}\beta} e^{-\frac{\dot{z}_{12}^2}{2\beta^2}}$$

$$\cdots$$

$$p_{z_{1n}, \dot{z}_{1n}}(z_{1n}, \dot{z}_{1n}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{z_{1n}^2}{2\sigma^2}} \frac{1}{\sqrt{2\pi}\beta} e^{-\frac{\dot{z}_{1n}^2}{2\beta^2}} \quad (25)$$

β^2 is the variance of the signal derivative. The signal and his derivative are independent and have Gaussian probability density.

The signals at the system output in two time instants t_1 and t_2 are z_1 and z_2. These signals are:

$$z_1 = \max\left(z_{111}, z_{121}, ..., z_{1n1}\right)$$

$$z_2 = \max\left(z_{112}, z_{122}, ..., z_{1n2}\right) \tag{26}$$

$z_{111}, z_{121}, ..., z_{1n1}$ are the signals at the combiner outputs at time instant t_1 and $z_{112}, z_{122}, ..., z_{1n2}$ at time instant t_2.

We will determine now the joint probability density of the signals z_1 and z_2 at the system output. This probability density is:

$$p_{z_1,z_2}\left(z_1, z_2\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} p_{z_{1i1},z_{1j2}}\left(z_1, z_2\right) \cdot \cdot \prod_{\substack{k=1\\k\neq i}}^{n}\prod_{\substack{s=1\\s\neq j}}^{n} F_{z_{1k1}}\left(z_1\right) \cdot F_{z_{1s2}}\left(z_2\right) \tag{27}$$

where

$$p_{z_{111},z_{112}}\left(z_{111}, z_{112}\right) = \frac{1}{\sqrt{2\pi}\sigma^2\sqrt{1-r^2}} \cdot \cdot e^{-\frac{(z_{111}-A)^2 - 2r(z_{111}-A)(z_{112}-A) + (z_{112}-A)^2}{2\sigma^2\left(1-r^2\right)}}$$

$$p_{z_{121},z_{122}}\left(z_{121}, z_{122}\right) = \frac{1}{\sqrt{2\pi}\sigma^2\sqrt{1-r^2}} e^{-\frac{z_{121}^2 - 2rz_{121}z_{122} + z_{122}^2}{2\sigma^2\left(1-r^2\right)}}$$

...

$$p_{z_{1n1},z_{1n2}}\left(z_{1n1}, z_{1n2}\right) = \frac{1}{\sqrt{2\pi}\sigma^2\sqrt{1-r^2}} e^{-\frac{z_{1n1}^2 - 2rz_{1n1}z_{1n2} + z_{1n2}^2}{2\sigma^2\left(1-r^2\right)}} \tag{28}$$

r is the coefficient of correlation. $\dot{z}_1$ and $\dot{z}_2$ are signals derivatives at the system output at two time instants. The joint probability density of the signals z_1 and z_2 and theirs derivatives is

$$p_{z_1z_2\dot{z}_1\dot{z}_2}\left(z_1, z_2, \dot{z}_1, \dot{z}_2\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} p_{z_{1i}z_{1j}\dot{z}_{1i}\dot{z}_{1j}}\left(z_1, z_2, \dot{z}_1, \dot{z}_2\right) \cdot \cdot \prod_{\substack{k=1\\i\neq k}}^{n}\prod_{\substack{s=1\\s\neq j}}^{n} F_{z_{1k1}}\left(z_1\right) F_{z_{1j2}}\left(z_2\right) \tag{29}$$

where

$$p_{z_{111}z_{112}\dot{z}_{111}\dot{z}_{112}}\left(z_{111}, z_{112}, \dot{z}_{111}, \dot{z}_{112}\right) = = \frac{1}{\sqrt{2\pi}\sigma^2\sqrt{1-r^2}} e^{-\frac{(z_{111}-A)^2 - 2r(z_{111}-A)(z_{112}-A) + (z_{112}-A)^2}{2\sigma^2\left(1-r^2\right)}} \cdot \cdot \frac{1}{2\pi\beta^2} e^{-\frac{\dot{z}_{111}^2 + \dot{z}_{112}^2}{2\beta^2}}$$

$$p_{z_{121}z_{122}\dot{z}_{121}\dot{z}_{122}}\left(z_{121}, z_{122}, \dot{z}_{121}, \dot{z}_{122}\right) = = \frac{1}{\sqrt{2\pi}\sigma^2\sqrt{1-r^2}} e^{-\frac{(z_{121})^2 - 2r(z_{111})(z_{122}-) + (z_{122})^2}{2\sigma^2\left(1-r^2\right)}} \cdot \cdot \frac{1}{2\pi\beta^2} e^{-\frac{\dot{z}_{121}^2 + \dot{z}_{122}^2}{2\beta^2}}$$

...

$$p_{z_{1n1}z_{1n2}\dot{z}_{1n1}\dot{z}_{1n2}}\left(z_{1n1}, z_{1n2}, \dot{z}_{1n1}, \dot{z}_{1n2}\right) = = \frac{1}{\sqrt{2\pi}\sigma^2\sqrt{1-r^2}} e^{-\frac{(z_{1n1})^2 - 2r(z_{1n1})(z_{1n2}-) + (z_{1n2})^2}{2\sigma^2\left(1-r^2\right)}} \cdot \cdot \frac{1}{2\pi\beta^2} e^{-\frac{\dot{z}_{1n1}^2 + \dot{z}_{1n2}^2}{2\beta^2}} \tag{30}$$

IV. CONCLUSION

The diversity system for coherent demodulation of n-ary digital frequently modulated signal in the presence of Gaussian noise at the system input is considered in this paper. The diversity system has L diversity brunches and we considered general case. The signals from the corresponding brunches of each receiver are added. The combiner determines (chooses) the brunch of the receiver for whom the signal from the output of the EGC combiner is the biggest. The decision is based on this signal. The results for the error probability are better than in some earlier suggested systems for the FSK signal demodulation [1] because of adding of signals from the same brunches of each receiver previously. The statistical characteristics of the signal and the error probability for the system are calculated in this paper. The probability density function of the signal at the system output, joint probability density function of the signal and signal derivatives at the system output, the joint probability density of the signal at the system output at two time instants and the joint probability density of the signal and signal derivatives at two time instants are derived.

REFERENCES

[1] Elisabeth A. Neasmith, Norman C. Beaulieu, “New Results on Selection Diversity”, *IEEE Trans. on Commun.*, vol. 46, No 5, May 1998.

[2] Y.Miyakgaki, N.Morinaga, and T. Namekawa, “Error probability characteristics for CPSK signal through m-distributed fading channel’, *IEEE Trans. Commun.*, vol. COM-26, pp. 88-89, Jan.1978.

[3] N. L. Johnson and S. Kotz, *Continuous Univariate Distributions,* vol. 2. New York: Wiley, 1976.

[4]W.C.Lee, *Mobile Communications: Design Fundamentals,* 2nd ed. New York: Wiley, 1993, pp. 202-211.

[5] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions.* New York: Dover, 1972.

MRC diversity systems in the presence of Log-Normal and Nakagami-m fading

Petar Nikolić [1] and Dragana Krstić [2]

Abstract – **A performance of dual combined diversity communication systems in the presence of Log-normal and Nakagami-m fading over two uncorrelated branches, is presented in this paper. Performance measures of fading communication systems such as Probability density function of SNR, Amount of fading and Outage probability are calculated and graphically represented for Maximal Ratio Combining.**

Keywords – **Log-Normal fading, Nakagami-m fading, Maximal ratio combining, Amount of fading, Outage probabilty.**

I. INTRODUCTION

In wireless communications, fading causes difficulties in signal recovery. When a received signal experiences fading during transmission, its envelope and phase both fluctuate over time. The overall fading process for land mobile satelite systems is a comlex combination of multipath fading and a log-nomal shadowing. Multipath fading caused by the constructive and destructive combination of randomly delayed, reflected, scattered and diffracted signal components. This type of fading is relatively fast and is responsible for the short-term signal variation. In terrestrial and satellite land-mobile systems, the link quality is also affected by slow variation of the mean signal level due to the shadowing from terrain, buildings and trees.

One of the methods used to mitigate this degradation is diversity. Diversity combining has been considered as an efficient way to combat multipath fading and improve the received signal-to-noise ratio (SNR) because the combined SNR compared with the SNR of each diversity branch, is being increased. In this combining, two or more copies of the same information-bearing signal are combining to increase the overall SNR

Maximal-Ratio Combining (MRC) is one of the most widely used diversity combining schemes whose SNR is the sum of the SNR's of each individual diversity branch. MRC is the optimal combining scheme, but its price is complexity, since MRC requires cognition of all fading parameters of channel.

This paper presents Maximal-Ratio Combining procedure for communication system where the diversity combining is applied over two uncorrelated $(\rho = 0)$ branches, which are given as channels with Log-Normal and Nakagami-m fading.

[1] Petar B. Nikolić is with the Tigar MH company, Nikole Pašica 99, 18300 Pirot, Serbia and Montenegro, E-mail: nikpetar@tigar.com
[2] Dragana S. Krstić is with the Faculty of Electronic Engineering, University of Nis, Aleksandra Medvedeva 14, 18000 Niš, Serbia and Montenegro, E-mail: dragana@elfak.ni.ac.yu

In this environment the receiver does not average the envelope fading due to multipath but rather, reacts to the instantaneous composite multipath/shadowed signal. This is often the scenario in congested downtown areas with slow-moving pedestrians and vehicles. This type of composite fading is also observed in land mobile satellite systems that are subjected to vegetative or urban shadowing.

II. SYSTEM AND CHANNEL MODELS

Here, a dual-branch diversity system over two uncorrelated channels in the presence of log-normal and Nakagami-m feding, is being considered. Under these conditions, instantaneous SNR $p(\gamma)$ is obtained by averaging the instantenous Nakagami-m feding average power over the conditional pdf (probability density function) of the log-normal shadowing, which results with combination of gamma distribution (for Nakagami-m feding) and log normal distribution (for log-normal shadowing) [1]

$$p(\gamma/\Omega) = \frac{m^m \gamma^{m-1}}{\Omega^m \Gamma(m)} e^{-\frac{m\gamma}{\Omega}}, \qquad \gamma \geq 0 \tag{1}$$

$$p_\Omega(\Omega) = \frac{\xi}{\sqrt{2\pi}\sigma_i \gamma_i} e^{-\frac{(10\log_{10}\gamma_i - \mu_i)^2}{2\sigma_i^2}} \tag{2}$$

$$p(\gamma) = \int_0^\infty p_\gamma(\gamma/\Omega) p_\Omega(\Omega) d\Omega \tag{3}$$

Substituting Eq. (1) and Eq. (2) in Eq. (3), $p(\gamma)$ can be written as

$$p(\gamma) = \int_0^\infty \frac{m^m \gamma^{m-1}}{\Omega^m \Gamma(m)} \exp\left(-\frac{m\gamma}{\Omega}\right) \times \left\{ \frac{\xi}{\sqrt{2\pi}\sigma\Omega} \exp\left[-\frac{(10\log_{10}\Omega - \mu)^2}{2\sigma^2}\right]\right\} d\Omega$$
$$, \gamma \geq 0 \tag{4}$$

where $\xi = 10/\ln 10 = 4.3429$, μ_i (db) is mean of $10\log_{10}\gamma$, σ_i (db) is standard deviation of $10\log_{10}\gamma$, and m is Nakagami-m factor.

In fig 1., 2. and 3. are shown pdf $p(\gamma)$ for various values of the factors μ, σ and m. For the case where $m = 1$, an overall fading consists of Rayleigh and log-normal feding.

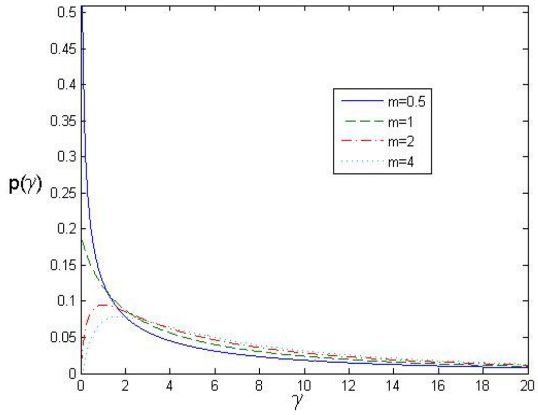


Fig. 1. $p(\gamma)$ for $\mu = 10db$, $\sigma = 5db$, $m = 0.5,1,2,4$

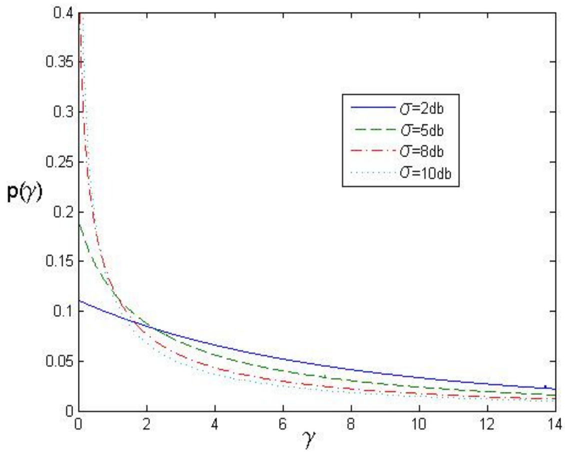


Fig. 2. $p(\gamma)$ for $\mu = 10db$, $\sigma = 2,5,8,10db$, $m = 1$

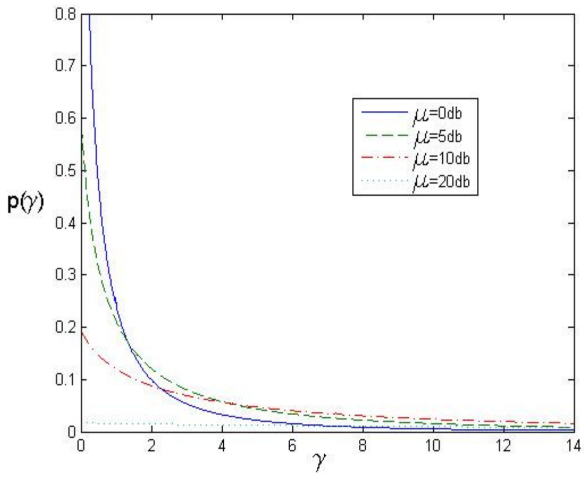


Fig. 3. $p(\gamma)$ for $\mu = 0,5,10,20db$, $\sigma = 5db$, $m = 1$

For further evaluations is important to find moments of combined SNR.

N-th moment of the Output SNR is given by [3]

$$E[\gamma^n] = \int_0^\infty \gamma^n p_\gamma(\gamma) d\gamma \tag{5}$$

N-th moment for distribution in Eq. (4) can be written as [1]

$$E[\gamma^k] = \frac{\Gamma(m+k)}{\Gamma(m) m^k} \exp\left[\frac{k}{\xi}\mu + \frac{1}{2}\left(\frac{k}{\xi}\right)^2 \sigma^2\right] \tag{6}$$

Amount of fading (AF) is a unified measure of the severity of fading for particular channel model and is typically independent of the average fading power, but is dependent of the instantaneous SNR.

Amount of fading is defined by

$$AF = \frac{\mathrm{var}(\alpha^2)}{(E[\alpha^2])^2} = \frac{E[(\alpha^2 - \Omega)^2]}{\Omega^2} \tag{7}$$

$$AF = \frac{E[\gamma^2] - (E[\gamma])^2}{(E[\gamma])^2} = \frac{E[\gamma^2]}{(E[\gamma])^2} - 1 \tag{8}$$

where α is fading amplitude, Ω is average fading power, $E[]$ denotes statistical average and var( ) denotes variance.

Join probability density function in the case of two uncorrelated fading channels is given by [2]

$$p_{\gamma 1,\gamma 2}(\gamma_1, \gamma_2) = p_{\gamma 1}(\gamma_1) p_{\gamma 2}(\gamma_2) \tag{9}$$

III. MAXIMAL RATIO COMBINING

The total SNR at the output of the MRC combiner is given by

$$\gamma_{MRC} = \sum_{l=1}^{L} \gamma_l \tag{10}$$

where L is number of branches.

The average combined SNR at the MRC (maximal ratio combining) output with two branceh is given by:

$$\gamma_{MRC} = \gamma_1 + \gamma_2 \tag{11}$$

Probability density function of the sum of the first and second branch can be written as

$$p_{\gamma_{MRC}}(\gamma_{MRC}) = \int_0^{\gamma_{MRC}} p_{\gamma 2}(\gamma_{MRC} - \gamma_1) p_{\gamma 1}(\gamma_1) d\gamma_1 \tag{12}$$

Substituting Eq. (4) in Eq. (11), $p_{\gamma_{MRC}}(\gamma_{MRC})$ can be obtained as

$$p_{\gamma_{MRC}}(\gamma_{MRC}) = \int_0^{\gamma_{MRC}} \left[\int_0^\infty \frac{m^m (\gamma_{MRC} - \gamma_1)^{m-1}}{\Omega^m \Gamma(m)} \exp\left(-\frac{m(\gamma_{MRC} - \gamma_1)}{\Omega}\right) \times \left\{ \frac{\xi}{\sqrt{2\pi}\sigma\Omega} \exp\left[\frac{10\log_{10}\Omega - \mu)^2}{2\sigma^2}\right] \right\} d\Omega \right] \cdot \left[\int_0^\infty \frac{m^m \gamma_1^{m-1}}{\Omega_1^m \Gamma(m)} \exp\left(-\frac{m\gamma_1}{\Omega_1}\right) \times \left\{ \frac{\xi}{\sqrt{2\pi}\sigma\Omega_1} \exp\left[\frac{10\log_{10}\Omega_1 - \mu)^2}{2\sigma^2}\right] \right\} d\Omega_1 \right] d\gamma_1 \tag{13}$$

Relatively simple closed form expressions to represent $p_{\gamma_{MRC}}(\gamma_{MRC})$ can not be derived, because Eq. (13) is too complex for tractable communication system analyses. This pdf can be evaluated numerically using some of software tools (Matlab, Mathematica).

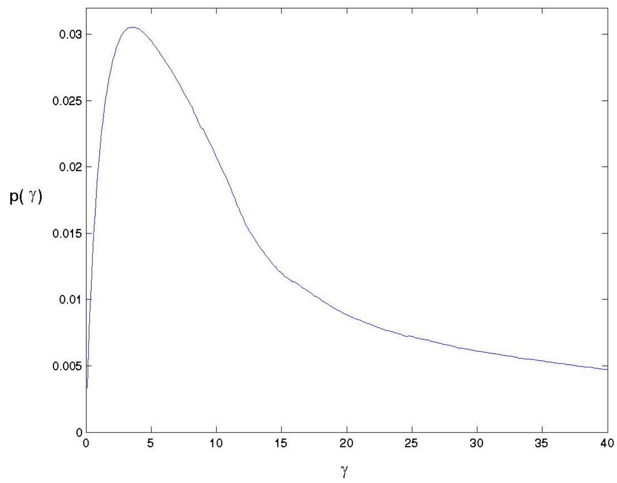


Fig. 4. $p_{\gamma_{MRC}}(\gamma_{MRC})$ for $\mu_i = 10db$, $\sigma_i = 5db$, $m = 1$

Very often it is assumed for performance analysis of communication systems, that channel coefficients are uncorrelated and identically distributed. Fig 4. depicts pdf of γ_{MRC} at the output of MRC combiner for two uncorrelated identically distributed channels with identical parameters $\mu_i = 10db$, $\sigma_i = 5db$, $m = 1$.

N-th moment of γ_{MRC} can be expressed as [1]

$$E\left[\gamma^n{}_{MRC}\right] = E\left[(\gamma_1 + \gamma_2)^n\right] \tag{14}$$

Using binomial expansion, (14) can be written as

$$E\left[\gamma^n{}_{MRC}\right] = E\left[\sum_{k=0}^{n}\binom{n}{k}\gamma_1{}^k\gamma_2{}^{n-k}\right] = \sum_{k=0}^{n}\binom{n}{k}E\left[\gamma_1{}^k\gamma_2{}^{n-k}\right] \tag{15}$$

The average combined SNR $\bar{\gamma}_{MRC}$ at the MRC output can be written as

$$\bar{\gamma}_{MRC} = E\left[\gamma^1{}_{MRC}\right] = 2\bar{\gamma} \tag{16}$$

The second moment of γ_{MRC} is given by

$$E\left[\gamma^2{}_{MRC}\right] = (\gamma_1{}^2 + 2\gamma_1\gamma_2 + \gamma_2{}^2) \tag{17}$$

For the parameters given in Fig 4. ($\mu_i = 10db$, $\sigma_i = 5db$, $m = 1$) it is obtained $\bar{\gamma}_{MRC} = 38.802$.

Amount of fading can be calculated from Eq. (16) and Eq. (17), and for these parametars is $AF_{MRC} = 7.528$.

Outage probabilty is standard performance criterion of diversity systems opereting over fading chanels and it is defined as the probability that the instantaneous error rate exceeds a specified value, or equivalently, that combined SNR of MRC falls below a predetermined threshold γ_{th}

$$P_{out}^{MRC} = P\left[\gamma_{MRC} = \gamma_1 + \gamma_2 \le \gamma_{th}\right] \tag{18}$$

P_{out}^{MRC} is defined in form of integral by

$$P_{out}^{MRC} = \int_0^{\gamma_{th}} p_{\gamma MRC}(\gamma_{MRC})d\gamma_{MRC} \tag{19}$$

Substituting Eq. (10) in Eq. (16) P_{out}^{MRC} can be written as

$$P_{out}^{MRC} = \int_0^{\gamma_{MRC}}\left[\int_0^{\infty}\frac{m^m(\gamma_{MRC}-\gamma_1)^{m-1}}{\Omega^m\Gamma(m)}\exp\left(-\frac{m(\gamma_{MRC}-\gamma_1)}{\Omega}\right)\times\left\{\frac{\xi}{\sqrt{2\pi}\sigma\Omega}\exp\left[\frac{10\log_{10}\Omega-\mu)^2}{2\sigma^2}\right]\right\}d\Omega\right]$$

$$\cdot\left[\int_0^{\infty}\frac{m^m\gamma_1{}^{m-1}}{\Omega_1{}^m\Gamma(m)}\exp\left(-\frac{m\gamma_1)}{\Omega_1}\right)\times\left\{\frac{\xi}{\sqrt{2\pi}\sigma\Omega_1}\exp\left[\frac{10\log_{10}\Omega_1-\mu)^2}{2\sigma^2}\right]\right\}d\Omega_1\right]d\gamma_1 d\gamma_{MRC} \tag{20}$$

Fig. 5 shows Outage probabilty versus instantaneous SNR for the same factors as in fig. 4.

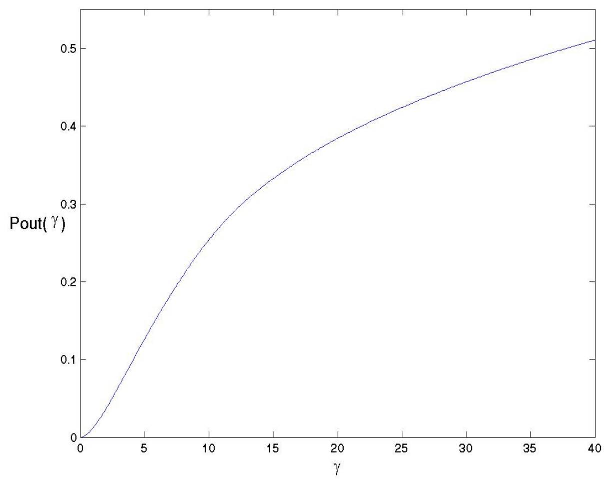


Fig. 5. P_{out}^{MRC} for $\mu_i = 10db$, $\sigma_i = 5db$, $m = 1$

IV. Conclusion

In this letter, a unified performance analysis for the dual diversity MRC over uncorrelated Nakagami-M fading and Log-Normal fading channels is presented. Probability density function of SNR, Amount of fading and Outage probability are derived in the form of multiple integral. It can not be obtained relatively simple closed-form expressions for evaluation of this parametars, because system structure is too complex and it was performed numerical calculation of them. As ilustration of this aproach, caracteristics of receiver are shown for MRC dual diversity case to point out the effect of the overall fading.

REFERENCES

[1] Marvin K.Simon, Mohamed-Slim Alouni, *Digital Communication over Fading Channels*, New York, 2000.

[2] Mihajlo Stefanović, *Detekcija signala u belom i obojenom Gausovom šumu*, Elektonski fakultet, Niš, 1999.

[3] Mohamed-Slim Alouni, Marvin K.Simon, ”Dual Diversity over correlated Log-normal Fading Channels”, *IEEE Trans. Commun*., vol. 50, pp.1946-1959, Dec 2002.

[4] Norihiko Morinaga, Ryuji Kohno, Seiichi Sampei, *Wireless communication technologies : new multimedia systems*, Kluwer Academic publisher, New York , 2002.

[5] Jianxia Luo, James R. Zeidle, ”Performance Analysis of Compact Antenna Arrays with MRC in Correlated Nakagami Fading Channels”, *IEEE Trans. Commun*., vol. 50, pp.267-277, Jan 2001.

[6] Bengt Holter, Geir E. Øien, *Performance analysis of a rate-adaptive dual-branch switched diversity system*, Norwegian University of Science and Technology, 2003.

[7] George K. Karagiannidis, Dimitris A. Zogas, Stavros A. Kotsopoulos, “ Performance Analysis of Triple Selection Diversity Over Exponentially Correlated Nakagami-m Fading Channels”, *IEEE Trans. Commun*., vol 51, no. 8, august 2003.

Microcomputer supported microwave transmitter and pseudo-monopulse receiver of GCS for UAV complex

Vladimir Smiljaković, Zoran Golubičić, Predrag Manojlović[1]

Abstract - **Paper presents microwave transmitter and pseudo-monopulse receiver supported by microprocessor developed and realized at Institute IMTEL as a part of telecommand-telemetry link for unmanned aerial vehicle (UAV) communication with ground control station (GCS). The role of described receiver is not only to receive signal from UAV's transmitter properly, but together with other parts of GCS and UAV telecommunication system to measure UAV-GCS distance and angular azimuthal error between boresight axe of GCS antenna system and actual position of UAV, thus forming autonomous measuring system of UAV position.**

Keywords - **microwave pseudo-monopulse receiver, microcomputer, unmanned aerial vehicle (UAV), autonomous position measuring system, intelligent interface**

I. Introduction

Global positioning systems that cover complete Earth exist nowadays and are commercially available to practically every potential user. However, in some cases still exists need for autonomous position measuring system for unmanned aerial vehicles (UAV) [1],[2]. This kind of position measuring system determines position in space of UAV only by use of parts of UAV system itself and not supported by anything out of it. Global systems are satellite based: GPS (USA), GLONASS (Russia), Galileo (European Union – under construction) or ground fixed reference stations based like: LORAN, Omega. Nowadays, a big number of UAV systems use GPS data for position determining and navigation because of low cost, small dimensions and weight and power supply of commercially available GPS receivers.

Still, some of UAV systems use autonomous position measuring system as an alternative way, or even as a main position determining system. Usually autonomous UAV position measuring system is referred to control station – whether it is at the ground, sea or in the air.

In the realized autonomous UAV position measuring system that is described in this paper polar coordinate system tied to ground control station (GCS) antenna system center as a reference point is used. It is field proven by test flights in the beginning with airborne communication part mounted on the top of automobile, then on the test helicopter and after proper functioning is approved airborne part is mounted in UAV and test flights are continued.

Autonomous UAV position determining system measures distance between GCS and UAV (so called slant distance-radius) applying the principle of secondary radar [3], while angular position in azimuth plane by use of monopulse receiver principle [4]. Radius is indirectly obtained by measurement of propagation time of emitted signals from GCS to UAV and vice versa. At the same time angular position of UAV is obtained by measuring azimuth error between GCS antenna system boresight axe and GCS-UAV direction of equal intensity signals. Angular position of UAV in elevation plane is obtained by finding ratio between UAV height above ground level and GCS-UAV radius length (sinus of elevation angle). UAV's height is obtained by barometric static air pressure measurement (absolute height above sea level). By these three parameters (radius, azimuth angle, elevation angle) UAV's position in space related to GCS as a center of spherical coordinate system is completely defined, and by referring to the geographical position of GCS consequently UAV's geographical position is also defined.

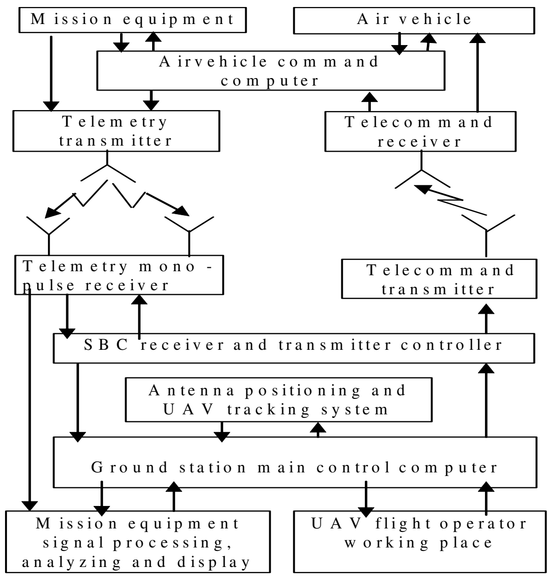


Fig. 1. Block diagram of UAV complex – airborne and ground part

The main reason for use of autonomous position (especially angular) measuring system in UAV systems is because of directional antenna system use for telemetry-telecommand links in GCS part of system. Use of directional antennas makes opportunity to obtain longer communication ranges keeping transmitted power the same. Due to unknown relative position (orientation) of UAV during flight omnidirectional antennas at UAV are preferred, though there are cases when

[1] Authors are with Institute of microwave technique and electronics IMTEL, Bulevar Mihajla Pupina 165b, 11 070 Novi Beograd, Serbia and Montenegro, email: [smiljac, golubicic, pedja]@insimtel.com

both at the UAV and GCS directional antennas are used. It is common practice when very long transmitting ranges are requested because energy supply at the UAV is extremely restricted, having direct influence to ability of obtaining as long mission duration as possible. This implies basic conclusion that fundamental fulfilled condition for acceptable quality of UAV-GCS communication is that UAV during the mission have to be in the main lobe of GCS directional antenna system. Hence the importance of monopulse receiver role for the whole UAV's mission success is clear. In cases when monopulse receiver principle is not applied in two orthogonal planes as usual, but in one plane only (in this case azimuth), some authors use term pseudo-monopulse receiver.

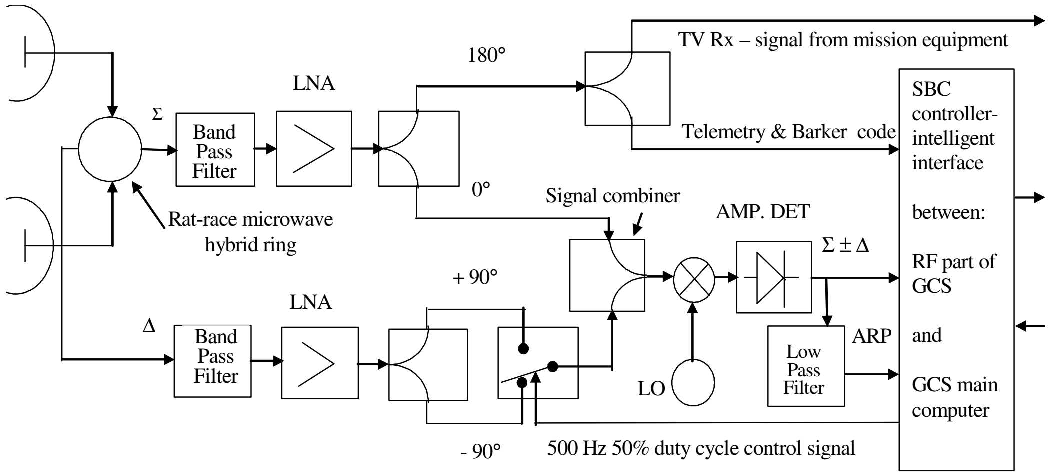


Fig. 2. Realized monopulse receiver without reference chanel

II. REALIZED UAV COMPLEX

Block diagram of realized UAV complex consists of one GCS and one UAV (figure 1). Air vehicle is airplane of classical configuration, wingspan of 4 meters, tricycle lending gear, requesting smooth terrain for take-off and landing. Cruising speed of UAV is about 100 km/h, maximal speed about 125 km/h and endurance is more than 1,5 hour. Propulsion is by two-stroke one cylinder petrol engine.

As electronic equipment at the UAV there are: telecommand receiver and actuators, air computer, mission equipment – day light TV camera and low light level camera, telemetry transmitter, antennas and set of sensors for flight and air frame status data. Signals from Pitot tube for air-speed measurement, barometer for height measurement, termopairs for motor cylinder head and exaust gases temperature measurement, air temperature, temperature in electronic compartment, fuel quantity, revolutions per minute of propeller and accumulator voltage are collected and inserted in telemetry message that UAV transmits towards GCS receiver together with mission equipment signal.

GCS is modular, easy to transport and use, consisting of modified radio-control for RC planes, telecommand transmitter, telemetry and mission equipment signal receiver, GCS main computer for virtual cockpit and digital terrain map display and monitor for mission equipment signal display (TV picture). Also received signal archiving device and antenna systems (transmitting and receiving) positioning, as well as power supply equipment are parts of GCS.

III. GCS-UAV COMMUNICATION SYSTEM

Block diagram of GCS-UAV communication system is depicted at figure 2. It is important to stress mutual coupling of functions between various systems in GCS-UAV complex to proper functioning of the complex as a whole.

Communication part at GCS consists of transmitter, transmitting antenna system, receiver antenna system and receiver. Transmitting and receiving antenna systems are mounted on common antenna positioning system, having possibility of azimuth and elevation axial rotation. Transmitter, receiver and antenna position controller are interfaced to main GCS computer, that also supports man-machine communication with UAV's operator in GCS. The main part of this interface is single board computer system that will be presented in this paper, and it is part of communication system of GCS.

Communication system part at the UAV consists of telecommand receiving antenna, telecommand receiver, telemetry and mission signal equipment transmitter and transmitting antenna.

IV. SINGLE BOARD COMPUTER FUNCTIONS

The role of this SBC is to properly interface digital command data obtained from main GCS computer to GCS telecommand link transmitter concerning data level, timing etc, to properly interface telemetry data obtained from GCS receiver to GCS main computer and to fulfill all measurement procedures necessary for obtaining data about UAV's position: distance between GCS and UAV [5], azimuth error between GCS receiving antenna boresight axe and direction of arrival of receiving telemetry signal from UAV [6] and other receiving signal parameters if possible (signal intensity).

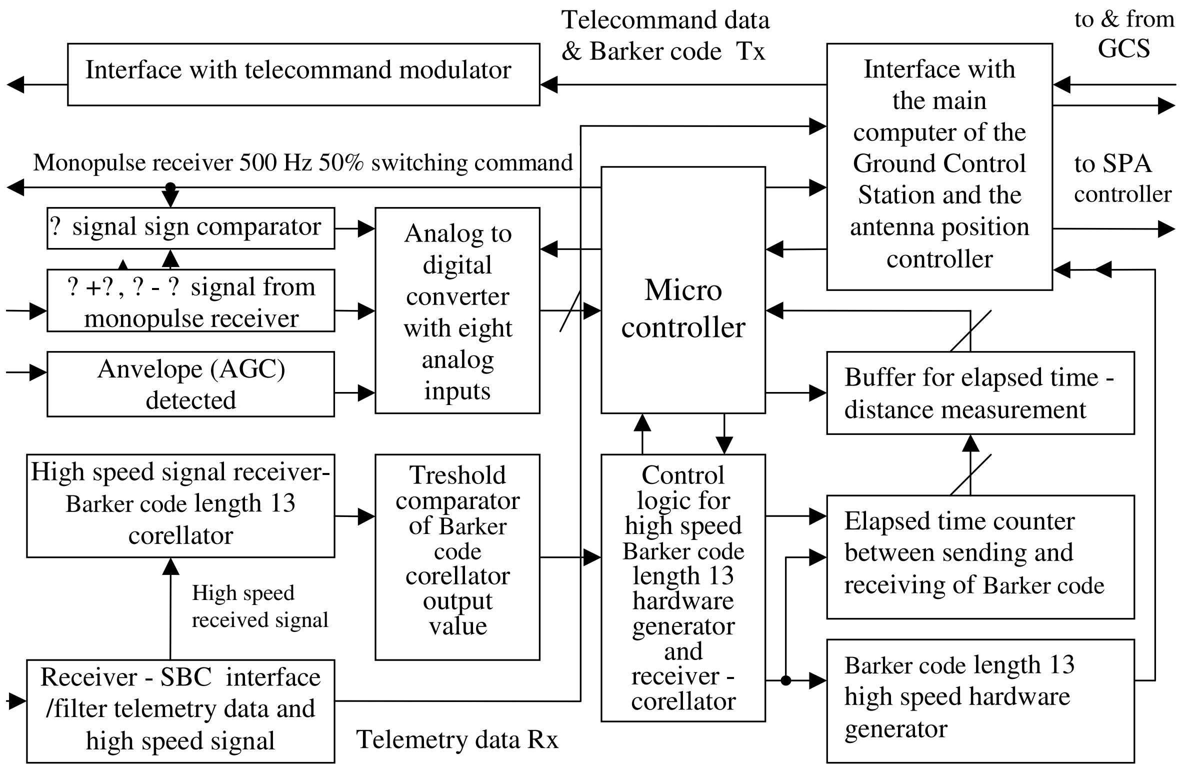


Fig. 3. Block diagram of microcontroller based single board computer for GCS communication system

SBC's activity is synchronized with the main computer activity, i.e. periodical transmitting of command data message – string of bytes with predefined format. When command string starts microcontroller directs it towards modulator. At the same moment microcontroller preloads – initiates hardware realized Barker code generator which is specialized periphery of microcontroller and then waits to detect end of command string. When end of command message is detected microcontroller issues command to open receiver and waits till end of receiving message is detected, at the same time starting time out counter. Time out counter threshold is so adjusted that is bigger than the longest ever needed time for telecommand – telemetry signal propagation along both directions together. If time-out happens and telemetry signal is not detected, unsuccessful communication is flagged-marked and another attempt is tried to accomplish communication. There are two kinds of flags-no signal at all and poor quality of received signal.

If the number of successive attempts without established communication is achieved above threshold, main GCS computer is alarmed and it commands position system controller to start acquisition of signal trying to establish communication. At the beginning search for telemetry signal starts at the angular position at which communication disappeared. Initial sweeping angle calculation is realized by applying worst case design approach: UAV trajectory is perpendicular to UAV-GCS direction (tangential moving), UAV is moving with maximal speed perpendicularly to last detected distance of UAV, and it lasts during the time equal to elapsed time from break of communication. Sweeping angle by azimuth is increased each time when attempt to establish communication fails.

Communication is established when monopulse receiver SBC detects good telemetry message. At that moment microprocessor starts hardware generator transmitting of Barker code train of pulses, starting at the same moment hardware quartz stabilized counter-timer that measures the time interval with resolution of 20ns. Barker code sequence has length of 13 bits and is well known for its property of clearly defined correlation peak value during one bit above all other values of correlation of this sequence (figure 3). Hardware generator and time counter are used because wanted time resolution is not possible to obtain using microprocessor only.

Transmitted Barker code is received by UAV's receiver of telecommands, regenerated and retranslated through telemetry link towards GCS receiving system. Counter-timer counts until receiving corellator gives value of correlation of 13 bits long Barker code and received signal above threshold. In that moment timer counting stops and interrupt to microprocessor is generated. Microprocessor loads value of counter in input buffer, and restarts the whole procedure, repeating it predefined number of times.

At the same time microprocessor generates symmetric train of command pulses to monopulse receiver. The role of this pulse train is to generate alternate commands to microwave switch in monopulse receiver [6] that alternatively enables synchronous detection of signals $\sum+\Delta$ and $\sum-\Delta$ in baseband.

After analog to digital conversion of these detected signals, appropriate numerical values of their values are fed to microcontroler. Finding sum and difference of these two signals by arithmetic operations in microcontroller numerical equivalents of original microwave intensities of signals are obtained. Symbols $\sum$ and Δ designate amplitude of receiving microwave telemetry signal sum and difference respectively at the appropriate outputs of rat-race modul of monopulse receiver (figures 2, 3). Signal designated by Δ under fulfilled circumstances corresponds to angular error in azimuth plane, as it is known from theory of monopulse receiver. Knowing data of angular position of antenna positioner and angular error, absolute position in azimuth plane of UAV is obtained. Comparing phase of command signal to monopulse receiver and measured Δ signal, antenna position error orientation is obtained, i.e. whether UAV is left or right (positive or negative error) from azimuth boresight axe of antenna system.

During the whole time interval between successive telecommand – telemetry message rounds is used for repetitive measurements of distance and angular position of UAV. The resultant measurement of one period (round) is obtained by averaging of successive measurement values in one round. Resultant measurements are fed to main computer of GCS and to controller of GCS antenna positioner.

Antenna position controller task is to track UAV's actual trajectory minimizing azimuth angular error [7]. Keeping angular error close enough to zero enables UAV to be in the main lobe of the GCS antenna system and thus to have the best possible quality of telemetry – telecommand signal, avoiding break of communication. Due to applied antenna construction (planar array of radiating elements) antenna diagram in elevation plane is like fan-stacked beams about 2 degrees wide in azimuth and 30 degrees wide in elevation plane. Because of that it is important to realize good tracking in azimuth plane only.

UAV by itself is subjected to its local disturbances influencing position (orientation) change in pitch, rotation and roll (all three axes of UAV's space orientation) due to unpredictable wind gusts, changing angles of pitch, rotation and roll (all three axes of UAV's space orientation) in spite of continual efforts to stabilize flight. Because of impossibility to obtain smooth flight trajectory, omnidirectional antenna at UAV is used as a simple solution. The cost of this choice is necessity of bigger transmitting power of telemetry signal, lowering of GCS telemetry receiver's signal threshold or increasing GCS receiving antenna gain to compensate lack of UAV antenna gain. This implies significant increasing of necessary energy of electrical supply of UAV vehicle, independently whether its electrical equipment is powered by accumulator (battery) or by electrical generator driven by internal combustion motor for airplane propulsion.

Increased weight has as a consequence shortened autonomy (and consequently shorter duration of useful part of mission) compared to directional antenna use at UAV and also growing cost of airvehicle as a flying platform itself. At the other side, use of directional antenna is not only much more complicated and expensive to realize, but also unreliable in functioning and maintenance due to rotary joints for microwave frequencies. So engineering trade-off in decision making is what we had to do also in this case.

V. Conclusion

Microcomputer with specialized peripheries realized on single board as a integral unit of ground control station communication system part that supports functioning of microwave telecommand transmitter and telemetry and signal mission signal receiver is developed and realized in Institute IMTEL. Single (azimuth) plane monopulse receiver operation is applied (called by some authors pseudo – monopulse receiver because of lack of other spatial plane – elevation channel receiver) for angular position in azimuth plane measurement and also secondary radar principle for slant distance measurement.

It is successfully field tested and proven in UAV complex as a part of the ground control station communication system. Proper functioning of this SBC enables autonomous determining of UAV's position in space and tracking of its movement along trajectory during mission, thus enabling communication between UAV and GCS. Further improvements of the system are under way.

Acknowledgement:

The work on this paper is partially funded by Ministry of Science and Environmental Protection of Republic of Serbia under the project No 6107.

References

[1] Slobodan Tirnanić *Bespilotne letilice* (in Serbian), Vojnoizdavački zavod, Beograd, Yugoslavia, 2001

[2] www.uavforum.com

[3] Michael Stevens *Secondary Surveillance Radar*, Artech House, Inc, Boston, Ma, USA, 1988.

[4] Donald Rhodes *Introduction to monopulse*, Artech House, Inc, Dedham, Ma USA, 1959.

[5] Vladimir Smiljaković, Zoran Golubičić, Predrag Manojlović, Zoran Živanović "Autonomous Distance Finding Microwave System for Middle Range Remotely Piloted Vehicle", *Proceedings of ICEST 2002 Conference*, pp 229-232, Niš, Yugoslavia, 2.-4. October 2002,

[6] Vladimir Smiljaković, Zoran Golubičić, Predrag Manojlović, Zoran Živanović "Microwave Autonomous Angular Position Finding System for Middle Range Unmanned Aerial Vehicle, *Proceedings of ICEST 2003 Conference*, pp 121-124, Sofia, Bulgaria, 16.-18. October 2003,

[7] George Biernson *Optimal Radar Tracking Systems*, New York, USA, John Willey and Sons, 1990

Vector's model of spatial-temporal signal, interference and noise for simulating radar's optimal processor

Le Quoc Vuong[1]

Abstract – **By construction mathematical model of spatial – temporal signal as a vector, with same way the article will construct the model of spatial – temporal interference, noise and combinatorial interference-noise. In the result, one can define the spatial-temporal interference-noise covariance matrix. This is the key step for simulation optimal interference suppressive processor using in radar.**

Keywords – **Vector model of spatial-temporal signal; Interference-noise covariance matrix; Optimal interference suppressive processor.**

I. INTRODUCE

For increasing probability of target detection (such as predetective process), in modern radar systems one often apply the spatial – temporal optimal processor (STOP). The basic task of STOP is to combine the signals in such a way that the interference is reduced to the level of the thermal noise while the desired signal is preserved. Stated another way, the goal of STOP would like to maximize the output *Signal to Interference plus Noise Ratio* (SINR). The general schema of STOP is shown in Figure 1.1.

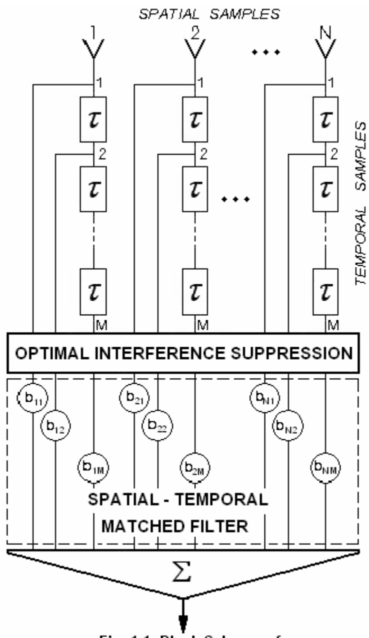


Fig. 1.1. Block Schema of Spatial-Temporal Optimal Processor (STOP)

From this schema, we can see that the STOP consists of two basic steps: Optimal interference suppression and spatial - temporal matched filter. Wherein, the optimal interference suppression filter plays most importance role and is an essential step to increased SINR. By nature, it is a digital filter, knows as *optimal interference suppressive filter* and its frequency response has form of inverse interference spectrum. To satisfying the special requirement of this frequency response, the coefficients (or weights) of the digital filter are defined by some algorithm (often it is adaptive algorithm), that base on estimation of interference spectrum. By close mathematical provableness, one has gained result: The coefficients of the digital filter accurately equal components of inverse interference-noise covariance matrix $\mathbf{Q}^{-1}$. Therefore in simulative models of the STOP, one often use the block $\mathbf{Q}^{-1}$ instead of the optimal interference suppressive filter. For example, the simulative model of Auxiliary Channel Processor (ACP) shown on Figure 1.2 is a characteristic Spatial - Temporal Near-Optimal Processor. In this model, the optimal interference suppressive filter has been simplifying such as the first row of inverse of interference-noise covariance matrix.

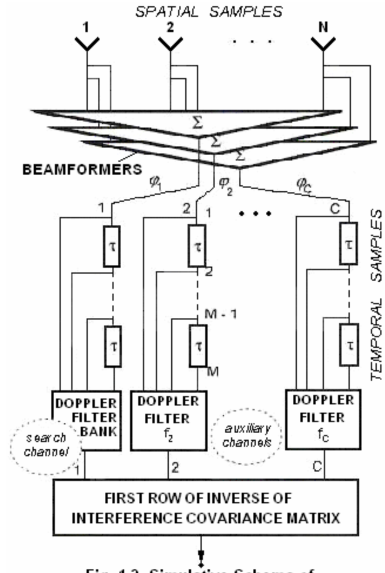


Fig. 1.2. Simulative Schema of Auxiliary Channel Processor (ACP)

For these reasons, the most important problem in simulative process of Spatial – Temporal Optimal Processor is definition of interference-noise covariance matrix $\mathbf{Q}$.

II. VECTOR'S MODELS OF SPATIAL, TEMPORAL AND SPATIAL-TEMPORAL SIGNALS

The construction for *mathematical models of signals* is to organize and arrange data in some fixed order. The vector's model of signal consist of 2 parts: the scalar part is represent

[1] Ph.D. Le Quoc Vuong is with Faculty of Ship's Electric - Electronic, Vietnam Maritime University, 484 Lachtray str., Haiphong city, Vietnam, E-mail: lqvuong05@yahoo.com.vn.

for the length of this vector, which is frequently called module or amplitude and the directional part of vector is an unit vector (having unit norm), that is frequently called *steered vector* of signal.

In the spatial domain, by [1, 2] we have a *vector of spatial signal*, that is the set of received signals from series outputs of N sensors, which is on a line with equal spacing between them (or uniform linear). It can be expressed as:

$$\mathbf{s}^{(K)}(n) = \sqrt{N}.\mathbf{v}^{(K)}\left(f^{(Ks)}\right).s(n) \tag{2.1}$$

where: $s(n)$ is amplitude (scalar part) of the signal vector, received from any output of a arbitrary sensor at a time instant n;

$f^{(Ks)}$ is normalized spatial frequency of signal, which is defined by:

$$f^{(Ks)} = \frac{d.\sin\varphi_s}{\lambda_0} \tag{2.2}$$

d – uniform space between 2 successive sensors;

φ_s – azimuth angle of arrival signal;

λ_0 – wavelength of carried signal;

The $\mathbf{v}^{(K)}\left(f^{(Ks)}\right)$ is *steering vector of spatial signal*, that is a Vandermonde vector. In this case we can write:

$$\mathbf{v}^{(K)}\left(f^{(Ks)}\right) = \frac{1}{\sqrt{N}}\left[1 \quad e^{-j2\pi f^{(Ks)}} \quad \dots \quad e^{-j2\pi(M-1)f^{(Ks)}}\right]^{/} \tag{2.3}$$

(Note: From (2.3) it can be deduce $\mathbf{v}^{(K)*}\mathbf{v}^{(K)} = 1$, or the norm of steering vector $\mathbf{v}^{(K)}$ equals unity);

Notation *K* - refers to space;

Notation *s* - refers to signal.

Likewise, in temporal domain, we can introduce the concepts about *vector of temporal signal* and *steering vector of temporal signal*. In accordance to [2], the reflective signal received from a moving target and defined at a time instant *n* on output of any arbitrary sensor, supposed being on the first sensor, has the form:

$$s(n) = A.e^{j2\pi nT_P f_{Ds}} \tag{2.4}$$

where: T_P is a pulse's period;

f_{Ds} is Doppler frequency of target's signal, which defined as:

$$f_D = \frac{2v_p}{\lambda_0}\cos\varphi \tag{2.5}$$

If set:

$$f^{(Ts)} = T_P f_{Ds} = \frac{f_{Ds}}{f_P} \tag{2.6}$$

Then $f^{(Ts)}$ is called the normalized temporal frequency of signal and now (2.4) can rewrite as:

$$s(n) = A.e^{j2\pi nf^{(Ts)}} \tag{2.7}$$

When this signal pass through a delay line consisted of $M-1$ tabs, what the lag of each tab equals T_p, it is correspond to phase rotation at an angle $-j2\pi T_p f_D$ (radian).

Thus, the input signal of delay line is: $s_1^{(T)}(n) = s(n) = A.e^{j2\pi nf^{(Ts)}}$. The signal after first lag is: $s_2^{(T)}(n) = A.e^{j2\pi(n-1)f^{(Ts)}}$. The signal after second lag is: $s_3^{(T)}(n) = A.e^{j2\pi(n-2)f^{(Ts)}}$. The signal after $M-1$ -st lag is: $s_M^{(T)}(n) = A.e^{j2\pi[n-(M-1)]f^{(Ts)}}$.

Denoted $\mathbf{s}^{(T)}(n)$ to be the temporal signal vector, which comprised of output signals of all successive delay taps, and then we have:

$$\begin{aligned}\mathbf{s}^{(T)}(n) &= \left[s_1^{(T)}(n) \quad s_2^{(T)}(n) \quad \dots \quad s_M^{(T)}(n)\right]^{/} = \\ &= \left[A.e^{j2\pi nf^{(Ts)}} \quad A.e^{j2\pi(n-1)f^{(Ts)}} \quad \dots \quad A.e^{j2\pi[n-(M-1)]f^{(Ts)}}\right]^{/} = \\ &= A.e^{j2\pi nf^{(Ts)}}\left[1 \quad e^{-j2\pi f^{(Ts)}} \quad \dots \quad e^{-j2\pi(M-1)f^{(Ts)}}\right]^{/} = \\ &= \sqrt{M}\,\mathbf{v}^{(T)}\left(f^{(Ts)}\right).s(n)\end{aligned} \tag{2.8}$$

where: $\mathbf{v}^{(T)}\left(f^{(Ts)}\right)$ is *temporal signal steering vector,* in which:

$$\mathbf{v}^{(T)}\left(f^{(Ts)}\right) = \frac{1}{\sqrt{M}}\left[1 \quad e^{-j2\pi f^{(Ts)}} \quad \dots \quad e^{-j2\pi(M-1)f^{(Ts)}}\right]^{/} \tag{2.9}$$

(Note: From (2.9) it can be deduce $\mathbf{v}^{(T)*}\mathbf{v}^{(T)} = 1$, or the norm of $\mathbf{v}^{(T)}$ equals unity).

The term "steering" expresses a common meaning, which in detail as "rotating to a direction..." in the spatial domain or "reaching to frequency..." in the time domain.

By use of one dimension signal vectors above, we can construct the vector's model of spatial-temporal signal. From the steering vector of spatial signal (1.3) and steering vector of temporal signal (1.8) we have the steering vector of spatial-temporal signal corresponding to the normalized temporal frequency $f^{(Ks)}$ and Doppler normalized frequency $f^{(Ts)}$:

$$\mathbf{v}^{(KT)}\left(f^{(Ks)}, f^{(Ts)}\right) = \mathbf{v}^{(K)}\left(f^{(Ks)}\right) \otimes \mathbf{v}^{(T)}\left(f^{(Ts)}\right) \tag{2.10}$$

where $\otimes$ is Kronecker product of 2 vectors; notation *(KT)* expresses the quantities in two-directional spatial-temporal domain. This vector is similar two correspond one-directional steering vector which has unit norm, that is $\mathbf{v}^{(KT)*}\mathbf{v}^{(KT)} = 1$. From (2.10) shown that spatial-temporal steering vector has dimension $NM \times 1$. With spatial-temporal steering vector $\mathbf{v}^{(KT)}$ we can construct the *vector model of spatial-temporal signal*, which has normalized temporal frequency $f^{(Ks)}$ and normalized Doppler frequency $f^{(Ts)}$:

$$\mathbf{s}^{(KT)}(n) = \sqrt{NM}.\mathbf{v}^{(KT)}\left(f^{(Ks)}, f^{(Ts)}\right).s(n) \tag{2.11}$$

Note that, the time variance *n* in (2.11) for quantities in two-dimension spatial-temporal domain "spreads" on an interval, which is defined:

– In space, it is the time for the signal passing over *N* sensors: T_N

– In time, it is the time for the signal passing over (*M-1*) delay taps: $(M-1)\tau$.

Normally, $(M-1)\tau \ll T_N$, therefore the actual spreading time is $(M-1)\tau$. Due to interested quantities are examined at the same starting time, therefore the variance n is not necessary. Thus, (2.11) can be rewritten:

$$\mathbf{s}^{(KT)} = \sqrt{NM}.\mathbf{v}^{(KT)}\left(f^{(Ks)}, f^{(Ts)}\right).s$$

(2.12)

The dimension of vector $\mathbf{s}^{(KT)}$ is $(NM \times 1)$.

The simulation algorithm of spatial-temporal signal vector followed (2.12) need input data, which comprise amount of sensors N, delay taps M, signal amplitude s, normalized spatial frequency $f^{(Ks)}$ and normalized Doppler frequency $f^{(Ts)}$. Where, the steering vector is defined follow (2.10).

III. VECTOR'S MODEL OF SPATIAL-TEMPORAL INTERFERENCE AND NOISE. CONSTRUCTION COVARIANCE MATRIX OF INTERFERENCE - NOISE

Suppose that there are C interference sources, divided into:

– C_K spatial interference sources, corresponding to the spatial frequencies $f_p^{(Ki)}$ with ($p = 1 \div C_K$);

– C_T temporal interference sources, corresponding to the temporal frequencies $f_q^{(Ti)}$ with ($q = 1 \div C_T$);

where, notation i expresses the quantities related to interference and $C = C_K + C_T$.

With interference, the spatial and temporal frequencies have relation following (2.6) (φ corresponding to the spatial frequency defined in (2.2), f_D corresponding to temporal frequency in (2.6)). Therefore, each spatial frequency $f_p^{(Ki)}$ corresponds to the temporal frequency $f_p^{(Ti)}$ and each temporal frequency $f_q^{(Ti)}$ corresponds to spatial frequency $f_q^{(Ki)}$. In the other words, C interference sources correspond to C couple of spatial frequency $f_l^{(Ki)}$ and temporal frequency $f_l^{(Ti)}$, with $l = 1 \div C$.

It can be deduce that, similarly to the vector of spatial-temporal signal (2.12), we can construct the *vector of spatial - temporal interference* for an interference source l, amplitude σ_{il}. It is written as:

$$\mathbf{i}_l^{(KT)} = \sqrt{NM}.\sigma_{il}.\mathbf{v}^{(KT)}\left(f_l^{(Ki)}, f_l^{(Ti)}\right) \quad (3.1)$$

The dimension of vector $\mathbf{i}_l^{(KT)}$ is $(NM \times 1)$.

The *general spatial - temporal interfe-rence vector* of all interference sources is:

$$\mathbf{i}^{(KT)} = \sum_{l=1}^{C} \mathbf{i}_l^{(KT)} \quad (3.2)$$

Assume that, the interference intensity on each of all directions (corresponding to all Doppler frequencies) is equal, that is $\sigma_{il} = \sigma_i$ with all l. Then (3.2) is now expanded:

$$\mathbf{i}^{(KT)} = \sqrt{NM}.\sigma_i \sum_{l=1}^{C} \mathbf{v}^{(KT)}\left(f_l^{(Ki)}, f_l^{(Ti)}\right) = \sqrt{NM}.\sigma_i.v_C^{(KT)} \quad (3.3)$$

where $\mathbf{v}_C^{(KT)} = \sum_{l=1}^{C} \mathbf{v}^{(KT)}\left(f_l^{(Ki)}, f_l^{(Ti)}\right)$ is *general spatial – temporal interference steering vector* with dimension $(NM \times 1)$. One element at k^{th} row ($k = p,q$) of this vector is detail expressed in form:

$$v_C^{(KT)}(p,q) = \sum_{l=1}^{C} e^{-j2\pi\left(pf_l^{(Ki)} + qf_l^{(Ti)}\right)} \quad (3.4)$$

where: $p = 1 \div N$ and $q = 1 \div M$.

The *vector of spatial – temporal noise* is the set of energy existing over all spatial – temporal channels, is also constructed in way based on (2.12). Assumed that the white noise having amplitude σ_n contributes equally on all spatial – temporal channels, and then the vector of spatial – temporal noise has the form:

$$\mathbf{n}^{(KT)} = \sqrt{NM}.\sigma_n.\mathbf{v}^{(KT)}\left(f^{(Kn)}, f^{(Tn)}\right) \quad (3.5)$$

with dimension $(NM \times 1)$.

The simulation algorithm of the general spatial – temporal interference vector carrying out by (3.3) requires input data, which composed of amount sensors N, delay taps M, isotropic interference amplitude σ_i and C normal spatial frequency of the interference $f_l^{(Ki)}$. The normalized Doppler frequencies of the interference $f_l^{(Ti)}$ are defined through normalized spatial frequency of the interference $f_l^{(Ki)}$ by (2.2), (2.5) and (2.6). In which, steering vector is still defined by (2.10)

The simulation algorithm of the vector of spatial - temporal noise carrying out by (3.5) requires input data, which composed of amount sensors N, delay taps M, isotropic noise amplitude σ_n, normalized spatial frequency $f^{(Kn)}$ and normalized Doppler frequency $f^{(Tn)}$. In which, steering vector is still defined by (2.10).

The combinative spatial - temporal signal vector also commonly referred to as *spatial - temporal data vector* $\mathbf{x}^{(KT)}$ has the form:

$$\mathbf{x}^{(KT)} = \mathbf{s}^{(KT)} + \mathbf{i}^{(KT)} + \mathbf{n}^{(KT)} \quad (3.6)$$

where: $\mathbf{s}^{(KT)}$ – the spatial - temporal signal vector;

$\mathbf{i}^{(KT)}$ – the general spatial– temporal interference vector

$\mathbf{n}^{(KT)}$ – the spatial – temporal noise vector.

Note that, all the spatial – temporal vectors shown in (3.6) have the dimension corresponding to the signal vector that is equal to $NM \times 1$.

The summation of the interference and noise vector, denoted as $\mathbf{q}^{(KT)}$, is called the *spatial – temporal interference – noise vector*:

$$\mathbf{q}^{(KT)} = \mathbf{i}^{(KT)} + \mathbf{n}^{(KT)} \tag{3.7}$$

The spatial – temporal interference-noise covariance mat–rix - $\mathbf{Q}$, is defined as:

$$\mathbf{Q} \triangleq E\left\{\mathbf{q}^{(KT)}.\left(\mathbf{q}^{(KT)}\right)^*\right\} \tag{3.8}$$

Thus $\mathbf{Q}$ has the dimension $NM \times NM$ and it is determined as followings.

Due to the linearity of the mathematical expectation operator and note that interference and noise are not correlative, therefore with respect to (3.8), we can expand as follow:

$$\mathbf{Q} = E\left\{\left(\mathbf{i}^{(KT)} + \mathbf{n}^{(KT)}\right)\left(\mathbf{i}^{(KT)} + \mathbf{n}^{(KT)}\right)^*\right\} =$$

$$= E\left\{\mathbf{i}^{(KT)}.\mathbf{i}^{(KT)*}\right\} + E\left\{\mathbf{n}^{(KT)}.\mathbf{n}^{(KT)*}\right\} = \mathbf{Q}_i + \mathbf{Q}_n \tag{3.9}$$

where: $\mathbf{Q}_i$ - is the spatial - temporal interference covariance matrix;

$\mathbf{Q}_n$ - is the spatial - temporal noise covariance matrix.

*Determining the noise covariance matrix $\mathbf{Q}_n = E\left\{\mathbf{n}^{(KT)}.\mathbf{n}^{(KT)*}\right\}$:

Replacing the spatial-temporal vector value in (3.5), we get:

$$\mathbf{Q}_n = \sigma_n^2\mathbf{I} = P_n\mathbf{I} \tag{3.10}$$

where: $\mathbf{I}$ - is unit diagonal matrix;

P_n - is the noise mean power.

Physically, we may clearly understand the meaning of (3.10), because of un-correlation of the noises in both spatial and temporal, therefore the only elements lying on diagonal are equal to 1 (due to the auto-correlation), the others are equal to 0.

* Determining the interference covariance matrix $\mathbf{Q}_i = E\left\{\mathbf{i}^{(KT)}.\mathbf{i}^{(KT)*}\right\}$:

Replacing the spatial – temporal interference vector value in (3.3), we get:

$$\mathbf{Q}_i = \sigma_i^2.E\left\{\mathbf{v}_C^{(KT)}.\mathbf{v}_C^{(KT)*}\right\} = P_i.E\left\{\mathbf{v}_C^{(KT)}.\mathbf{v}_C^{(KT)*}\right\} \tag{3.11}$$

where: P_i - is the interference mean power.

The solution to determine any one of elements of the interference covariance matrix is presented as following:

– As we knew, an element lying on the row $p\&q$ of the general steering vector has the form shown in (3.4).

– By similarly way, we may determine any one of elements lying on the column (g,h) of the vector $\mathbf{v}_C^{(KT)*}$:

$$v_C^{(KT)}\left(g,h\right) = \sum_{l=1}^{C} e^{j2\pi\left(gf_l^{(Ki)} + hf_l^{(Ti)}\right)} \tag{3.12}$$

with $g = 1 \div N$ and $h = 1 \div M$.

– From the (3.4) and (3.12), we can find out an element lying on row (p,q) and at the column (g,h) of the interference covariance matrix Q_i determined as form:

$$Q_i\left(p,q;g,h\right) = P_i.E\left\{v_C^{(KT)}\left(p,q\right).v_C^{(KT)*}\left(g,h\right)\right\} =$$

$$= P_i.\sum_{l=1}^{C} e^{j2\pi\left((g-p)f_l^{(Ki)} + (h-q)f_l^{(Ti)}\right)} \tag{3.13}$$

IV. CONCLUSION

In fact, mathematical models of spatial-temporal signals may be shown by alternative 3 methods:

The classically expression method that common used is *analytic* form. By this method, the mathematical model of signal is shown in form as a function. In the case of spatial-temporal signal, the analytic model is a multi-variable function (or multidimensional function). This method is very general and often uses to analyze any signal or system. But it has the *basic weakness*, that is: It can't be use in simulation, because can't construct the spatial-temporal interference-noise covariance matrix from it.

For multidimensional signals such as the spatial-temporal signals, we may apply matrix solution to mathematically simulate them. However, the interference-noise covariance matrix expressing the correlation among elements of the two matrices will have spatial dimension very large and become extremely complex. Therefore, this method is unprofitable to simulate spatial-temporal optimal processor.

Thus the expression of vector model provided in this article is rather simple to simulate process of interference suppression, using in spatial-temporal signal vector model. This affirms the special advantages of the spatial-temporal signal vector model.

REFERENCE

[1] V. Oppenheim, R. W. Schafer, J. R. Buck (1990), *Discrete-Time Signal Processing*, Second editon, Prentice Hall, New Jersey.

[2] R. Klemm (2002), *Principles of Space-Time Adaptive Processing*, The Institution of Electrical Engineers, London, UK.

[3] Д. Е. Даджион, Р.М. Мерсеро (1988), *Цифровая обработка многомерных сигналов*, Изд. "Мир" Москва, USSR.

[4] Lê Quốc Vượng (2006), *Ph. D. Thesis "Spatial Signal Processing - Xử lý tín hiệu không gian"*, Institute of Military Techniques, Vietnam.

Studies of Linear and Quadratic Approximations of GPS Satellite Radius Alteration in Time

Dimitar V. Dimitrov[1] and Marin S. Marinov[2]

Abstract –**This paper presents some studies of linear and quadratic approximation of radius alteration of satellite (SV) in GPS using Taylor's series. Studies are made, using real data from satellites ephemeris message. The results for the best and the worst case are shown and the errors of approximations are obtained.**

Keywords – **GPS, approximation.**

I. Introduction

Presently, GPS is only fully operational satellite positioning system. The satellites broadcast ranging codes and navigation data. The navigation data provides the means for a receiver to determine the location of satellites, whereas the ranging codes enable the user's receiver to determine the propagation times of signals and thereby determine the satellite-to-user ranges.

The ephemeris message is a part of navigation data and contains the following orbital parameters [1, 2]:

- t_{0e} – Reference time of ephemeris;
- $\sqrt{a}$ – Square root of semi major axis;
- e – Eccentricity;
- i_0 – Inclination angle (at time t_{0e});
- Ω_0 – Longitude of the ascending node (at weekly epoch);
- ω – Argument of perigee (at time t_{0e});
- M_0 – Mean anomaly (at time t_{0e});
- di/dt – Rate of change of inclination angle;
- $\dot{\Omega}$ – Rate of change of longitude of the ascending node;
- Δn – Mean motion correction;
- C_{uc} – Amplitude of cosine correction to argument of latitude;
- C_{us} – Amplitude of sine correction to argument of latitude;
- C_{rc} – Amplitude of cosine correction to orbital radius;
- C_{rs} – Amplitude of sine correction to orbital radius;
- C_{ic} – Amplitude of cosine correction to inclination angle;
- C_{is} - Amplitude of sine correction to inclination angle;

These parameters are used to calculate the rest of satellite orbital parameters and coordinates: n-corrected mean motion, E-eccentric anomaly, ϕ-argument of latitude, r-corrected radius, (x, y, z)-satellite coordinates in Earth-Centered Earth-Fixed (ECEF) coordinate system, utilize a standard algorithm [1, 2, 3].

There is an indirect relationship between radius of satellite and time. Usually it is more convenient when this relation is direct [6]. An approximation is needed to convert indirect to direct dependence.

In this paper, studies of linear and quadratic approximation of radius for each satellite, are presented. A U-blox GPS receiver ANTARIS AEK-4P is used to collect the necessary ephemeris data.

II. Approximation of Radius

In the standard algorithm the satellite radius r_{sk} is calculated by the following equation:

$$r_{sk} = a(1 - e\cos E_k) + C_{rs}\sin(2\phi_k) + C_{rc}\cos(2\phi_k)\,, \quad (1)$$

where index "k" denote the k-th satellite; E_k and ϕ_k are functions of time.

One of the most popular approximation techniques is representation of a function in a Taylor series [4, 5]. By expanding (1) in Taylor series about time instant $t=t_n$, for radius is obtained:

$$r_{sk}(t) \approx r_{sk}(t_n) + \frac{dr_{sk}}{dt}\bigg|_{t=t_n}(t-t_n) + \frac{1}{2}\frac{d^2 r_{sk}}{dt^2}\bigg|_{t=t_n}(t-t_n)^2 + \frac{1}{6}\frac{d^3 r_{sk}}{dt^3}\bigg|_{t=t_n}(t-t_n)^3\,, \quad (2)$$

The close form for derivatives in (2) is found and they result in the following equation for satellite radius:

$$r_{ak}(t) \approx r_{sk}(t_n) + A(t-t_n) + B(t-t_n)^2 + C(t-t_n)^3\,, \quad (3)$$

where: A, B and C are approximation coefficients, calculated at the instant $t=t_n$ by the following equations:

$$A = \frac{ane\sin E_k}{1-e\cos E_k} + \frac{2n\sqrt{1-e^2}}{(1-e\cos E_k)^2}\times \times(C_{rs}\cos 2\phi_k - C_{rc}\sin 2\phi_k)\,, \quad (4)$$

$$B = \frac{an^2 e}{2}\frac{\cos E_k}{(1-e\cos E_k)^3} - \frac{2n^2\sqrt{1-e^2}}{(1-e\cos E_k)^4}\times \times\left[\sqrt{1-e^2}(C_{rs}\sin 2\phi_k + C_{rc}\cos 2\phi_k) + e\sin E_k(C_{rs}\cos 2\phi_k - C_{rc}\sin 2\phi_k)\right], \quad (5)$$

[1] Dimitar V. Dimitrov, is with the Aviation Faculty, National Military University, 5856 Dolna Mitropolia, Bulgaria, e-mail: Lz5wx@abv.bg.

[2] Marin S. Marinov, PhD is with the Aviation Faculty, National Military University, 5856 Dolna Mitropolia, Bulgaria, e-mail: mmarinov2000@yahoo.com.

$$C = \frac{an^3 e}{6}\left[\frac{\sin E_k}{(1-\cos E_k)^4} + \frac{3e\sin E_k(\cos E_k - e)}{(1-\cos E_k)^5}\right] + \\ + \frac{2n^3\sqrt{1-e^2}}{3(1-\cos E_k)^6}[6e\sqrt{1-e^2}\sin E_k(C_{rs}\sin 2\phi_k + \\ + C_{rc}\cos 2\phi_k) + (C_{rs}\cos 2\phi_k + C_{rc}\sin 2\phi_k)\times \\ \times(4e^2\sin^2 E_k - 2\sqrt{1-e^2} - e\cos E_k + e^2\cos^2 E_k)] \quad (6)$$

In case of linear approximation the third and fourth terms in equation (3) are rejected. It is known that the error of approximation, when Taylor's series is used, does not exceed value of first rejected term [7]. Consequently the error of linear radius approximation is given by:

$$\Delta_{apr1} \le B(t-t_n)^2 . \quad (7)$$

In case of quadratic approximation the fourth term in equation (3) is rejected. The error of quadratic radius approximation is given by:

$$\Delta_{apr2} \le C(t-t_n)^3 . \quad (8)$$

Limited user range accuracy (URA) is provided by data in ephemeris message. Nominal value of URA is in range between 2m and 4000m [2]. In normal operational conditions this value is between 2m and 8m. Information for current URA value is also transmitted in navigation data.

If the errors of approximation (7) and (8) are of much more less value than URA, then they could be neglected. The conditions meeting these requirements are defined by:

$$\Delta_{apr1} \le 0.1\text{URA}; \qquad \Delta_{apr2} \le 0.1\text{URA} \quad (9)$$

Consequently the errors of approximation should be bellow 0.2 m.

III. Researches

Researches about radiuses of all GPS satellite and corresponding approximation coefficients are made. The navigation data from satellites is collected for a 42 hours period using GPS receiver ANTARIS AEK-4P. Some of the results are shown below.

Studies show that the radius alteration and corresponding approximation coefficients are unique for each satellite. The alteration for satellite SV27 is the greatest, while for satellite SV17 it is the smallest. The similar results are obtained for approximation coefficients – their values for SV27 are the biggest, while for SV17 they are the smallest.

Fig. 1 illustrates radius of satellite SV 27 as function of time.

It is seen that the change of radius r_{sk} is periodic and the difference between minimum and maximum values is over 1000 km.

In Fig.2 approximation coefficients B and C for SV27 are shown. As it was mentioned the approximation accuracy depends on them.

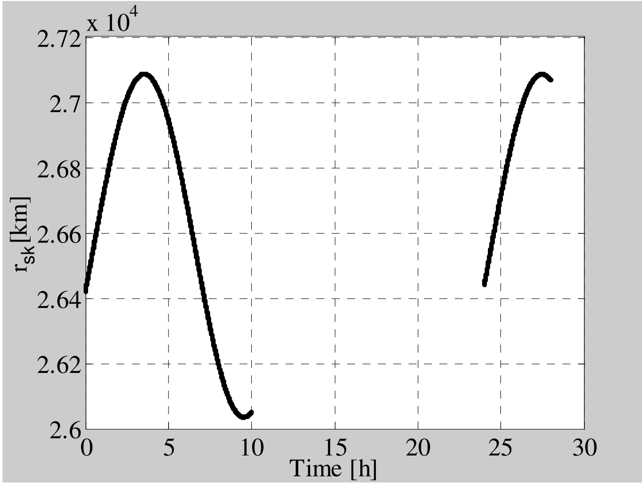


Fig. 1. Radius r_{sk} of SV27 as function of time.

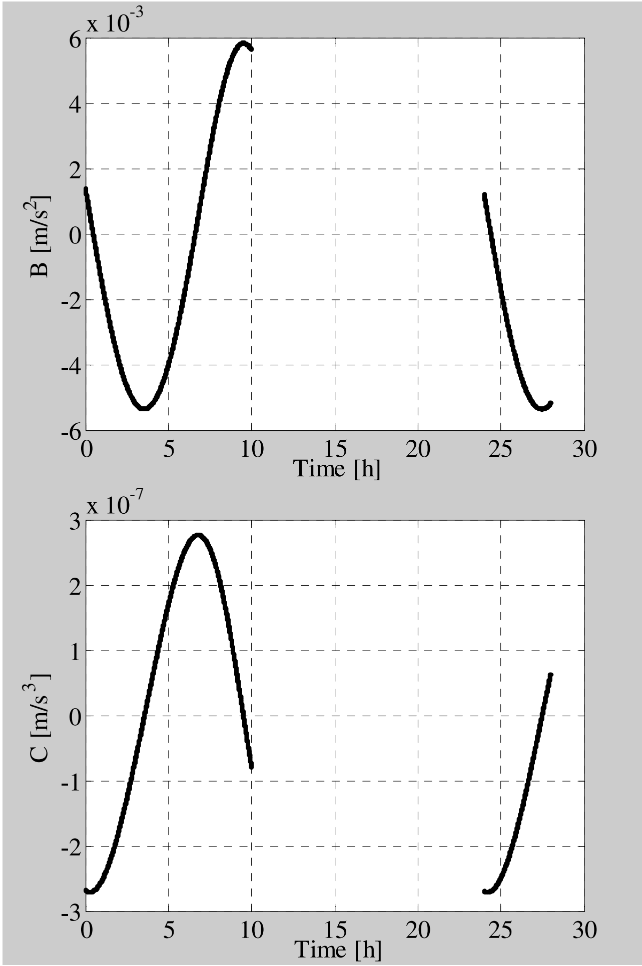


Fig. 2. Approximation Coefficients *B* and *C* for SV27.

As one can see there are periodical changes of coefficients values. The magnitude of B is up to 5.83×10^{-3} m/s^2 and the magnitude of C does not exceed 2.7643×10^{-7} m/s^3.

The radius of satellite SV17 is represented in Fig. 3. It is also periodic and the change of its value is less than 100 km.

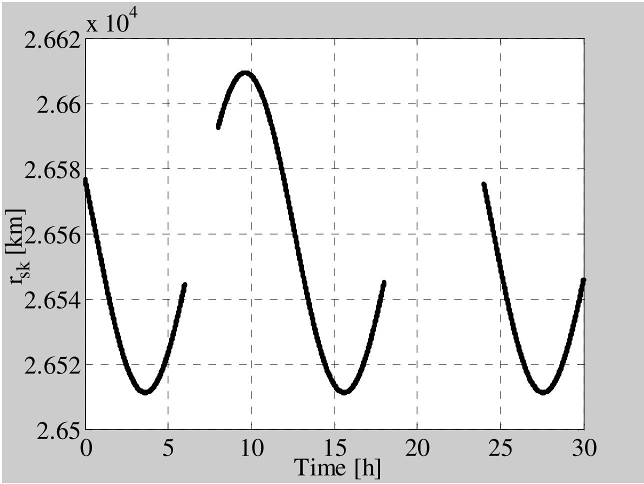


Fig. 3. Radius r_{sk} of SV17 as function of time.

Corresponding coefficients B and C for this satellite are shown in Fig. 4.

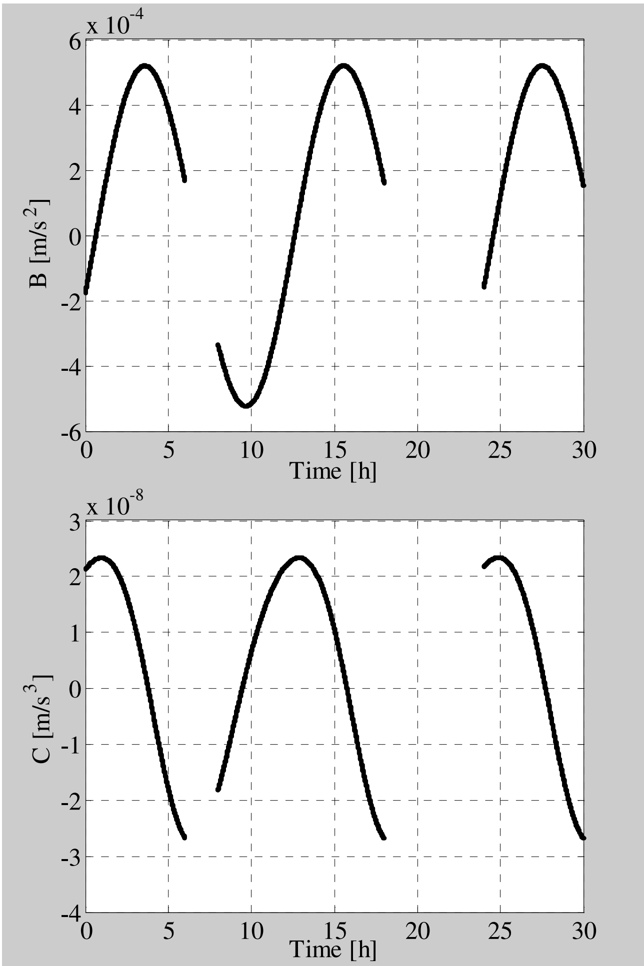


Fig. 4. Approximation Coefficients *B* and *C* for SV17.

Results indicate that the magnitude of B does not exceed 5.24×10^{-4} m/s^2, and the magnitude of C is up to 2.6867×10^{-8} m/s^3.

Analysis of results shows that there is a huge difference of radius alteration for these two satellites. This is because of the different orbital planes of two satellites. There is a difference in radius changes in all other satellites in constellation.

As it is seen in figures above there is absence of data for radius and respectively in calculated coefficients. This is because the satellite has not been visible from point of the GPS receiver reception at this time. Periodical behavior of the radius shows that the changes of its magnitude could not be extremely different.

The magnitudes of approximation coefficients B and C for SV27 are the worst case, because none coefficient for all other satellites exceeds their values. That's why these values are put in (7) and (8) to calculate the maximum of approximation errors Δ_{apr1} and Δ_{apr2}.

Next researches are focused to obtain the approximation errors as function of time interval of approximation (t-t_n).

The relation between Δ_{apr1} and the time interval, using linear approximation of the radius to both SV17 and SV27, is shown in Fig. 5.

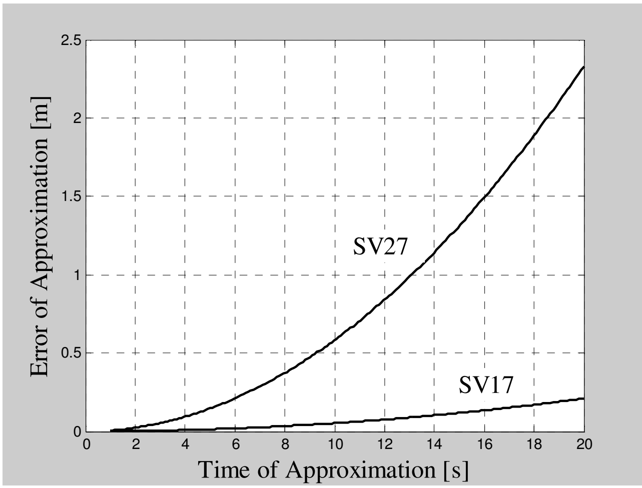


Fig. 5. Approximation error Δ_{apr1} for SV27 and SV17.

It is seen that Δ_{apr1} for SV27 is bigger than Δ_{apr1} for SV17 using the same time of approximation. Exact values of time interval, in case of maximum allowed error Δ_{apr1} = 0.2 m, are 5.8 seconds for SV27 and 19.5 seconds for SV 17. From these results it is obvious that the linear approximation is useable only if the rate of measurements is high enough. For instance, if the rate is 1 measurement per second, there are only five measured values in approximation time interval for satellite SV27.

Fig. 6 compares Δ_{apr2} for the same satellites, in case of quadratic radius approximation. As one can see the approximation time interval when Δ_{apr2} approach the value of 0.2 m, is larger than the time interval, when linear approximation is used. Exact values of time interval, in case of maximum allowed error, are 89 seconds for SV27 and 195 seconds (not shown in the Figure) for SV17.

Studies shows that the possible approximation intervals are large enough to use proposed approximation even when the rate of measurement is low. The GPS receiver used for researches outputs the user position coordinates four times per second. In this case 356 measurements are available for one approximation time interval in the worst case.

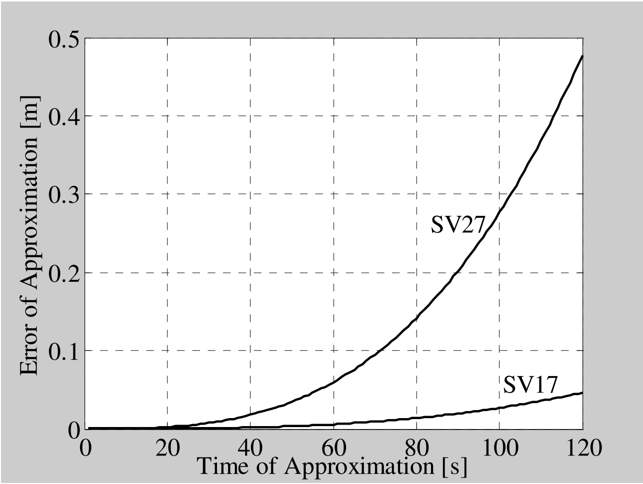


Fig. 6. Approximation error Δ_{apr2} for SV27 and SV17.

IV. Conclusion

The results of researches show that the radius alteration for every satellite in constellation could be approximated with quadratic or even linear function of time. This approximation can be used for future models.

The key features to approximate the radius alteration are the approximation errors and the approximation time interval. In case of using short time interval of approximation it is recommended to approximate radius with linear function of time.

When quadratic approximation is utilized longer time interval of approximation could be used , but the number of calculation operations will be increased because of additional coefficient B needed to compute, even though in this case, the number numerical calculation of Kepler's equation in standard algorithm [1, 2, 3] will be more seldom.

References

[1] E. D. Kaplan, *Understanding GPS Principles and Applications*, Artech House, London, 1996.

[2] *Interface Specification IS-GPS-200 Revision D*, ARINC Engineering Services, El Segundo, 2004.

[3] D. J. Dailey, B. M. Bell, *A Method for GPS Positioning*, IEEE T-AES, Vol. 32, No. 3, 1996.

[4] M. Moore, J. Wang, *Adaptive Dynamic Modelling for Kinematic Positioning*, The University of New South Wales, Sydney, Australia, 2001.

[5] M. Moore, J. Wang, *An Extended Dynamic Model for Kinematic Positioning*, The University of New South Wales, Sydney, Australia, 2002.

[6] M. Marinov, G. Stanchev, *Use of Space Correlation of Satellite Move in GPS*, Proceeding of XXXVII International Scientific Conference ICEST, 16-18 October 2003, Sofia, Bulgaria, 2003.

[7] G. A. Korn, T. M. Korn, *Mathematical Handbook for Scientists and Engineers,* McGraw-Hill, 1968.

Studies of an Approximation of User-to-Satellite Range in GPS

Dimitar V. Dimitrov[1] and Marin S. Marinov[2]

Abstract – **This paper presents some studies of an approximation of user-to-satellite range in GPS. Studies are made, using real data from satellites navigation message. The results for the best and the worst case are shown and the errors of approximations are obtained.**

Keywords – **GPS, approximation, user-to-satellite range.**

I. Introduction

Different methods, as Kalman filtering, the least square algorithms and etc., are used by GPS receivers to calculate user position. In these algorithms, models of various parameters are usually used. In the most of the used algorithms an approximation of user-to satellite range is utilized [4, 6, 7].

A user-to-satellite range approximation, using approximation of satellite radius proposed in [3] is presented in this paper. The studies are made to prove that the errors of proposed approximation are much smaller than errors in pseudorange measuring.

A U-blox GPS receiver ANTARIS AEK-4P is used to collect the necessary navigation data from satellites.

II. Approximation of User-to-Satellite Range

Relation between radius of user R_u and distance user-to-k^{th} satellite D_k is [8]

$$D_k = \sqrt{r_{sk}^2 + R_u^2 - 2r_{sk} R_u \cos\alpha_k} \quad (1)$$

where α_k is the angle between radius of user R_u and radius of k^{th} satellite r_{sk}.

As was obtained in [3] the radius approximation for k^{th} satellite is

$$r_{ak}(t) \approx r_{sk}(t_n) + A(t-t_n) + B(t-t_n)^2 + C(t-t_n)^3 \quad (2)$$

where A,B and C are approximation coefficients derived in [3]. If quadratic approximation is used only the first three terms are needed.

Substituting Eq. (2) into Eq. (1) yields:

$$\begin{aligned} D_{ak} = \{ & D_{k0}^2 + B^2(t-t_n)^4 + 2AB(t-t_n)^3 + \\ & + (A^2 + 2r_{sk0}B)(t-t_n)^2 + 2r_{sk0}A(t-t_n) - \\ & - 2R_u \cos\alpha_k [B(t-t_n)^2 + A(t-t_n)]\}^{\frac{1}{2}} \end{aligned} \quad (3)$$

where
$$D_{k0}^2 = r_{sk0}^2 + R_u^2 - 2r_{sk0}R_u \cos\alpha_k ; \quad (4)$$

$$r_{sk0} = a(1 - e\cos E_k) + C_{rs}\sin(2\phi_k) + C_{rc}\cos(2\phi_k) \quad (5)$$

is the satellite radius, computed by standard algorithm from received ephemeris data [1, 2] at instant $t=t_n$; a is the semi major axis of satellite orbit, e is the eccentricity, E_k is the eccentric anomaly, C_{rs} and C_{rc} are sine and cosine corrections to orbital radius, ϕ_k is the argument of latitude.

The Eq. (3) could be rewritten in the following form:

$$D_{ak} = D_{k0}\sqrt{1+X} , \quad (6)$$

where

$$\begin{aligned} X = & \frac{B^2(t-t_n)^4 + 2AB(t-t_n)^3}{D_{k0}^2} + \\ & + \frac{(A^2 + 2r_{sk0}B)(t-t_n)^2 + 2r_{sk0}A(t-t_n)}{D_{k0}^2} - \\ & - \frac{2R_u \cos\alpha_k [B(t-t_n)^2 + A(t-t_n)]}{D_{k0}^2} \end{aligned} \quad (7)$$

It is useful to expand the root square in Eq. (6) in series to simplify expression. The condition that must be meet is X<<1 [8]. The range of distance D_{k0} is the widest in comparison with all other variables in Eq. (7). On the other hand distance D_{k0} is raised to the second power. That's why the influence of D_{k0} on X is the greatest. The worst case for approximation is when the value of D_{k0} is minimum. This takes place when the user is situated on the geodetic normal from the satellite position. The geometric relations for above situation are shown in Fig. 1.

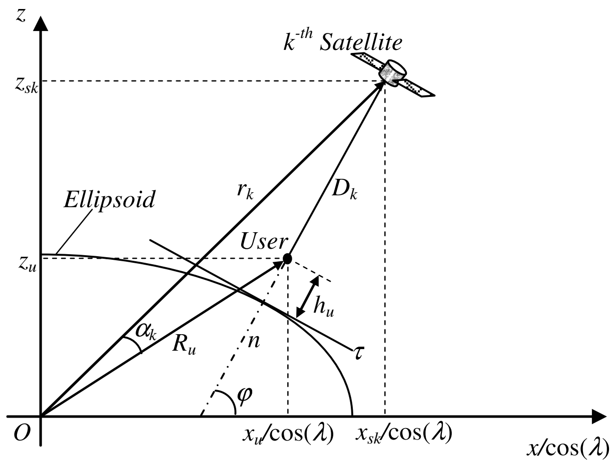


Fig. 1. Geometric relation

[1] Dimitar V. Dimitrov, is with the Aviation Faculty, National Military University, 5856 Dolna Mitropolia, Bulgaria, e-mail: Lz5wx@abv.bg.

[2] Marin S. Marinov, PhD is with the Aviation Faculty, National Military University, 5856 Dolna Mitropolia, Bulgaria, e-mail: mmarinov2000@yahoo.com.

The earth is represented as a reference ellipsoid according to WGS-84 [2]. In Fig. 1 O is the origin of both Earth-Centered Earth-Fixed (ECEF) and geodetic coordinate systems; φ and λ are corresponding geodetic latitude and longitude of satellite position; R_u is radius of user situated on the geodetic normal n; τ is tangent to the ellipsoid; h_u is user height above the ellipsoid.

Transformation from ECEF coordinates of k^{-th} satellite to corresponding geodetic coordinates is done to compute the minimum user-to-satellite range D_{k0}. From Fig.1 it is obvious that the geodetic height of satellite position is the distance from ellipsoid surface to the satellite. Maximum value of user height h_u must be subtracted from satellite height to obtain the minimum value of D_k. Usually the maximum value of user height h_u is 20 000 meters. Transformation from ECEF to Geodetic coordinate system (latitude φ , longitude λ and height h) are given by the following equations[5]:

$$\varphi = \tan^{-1}\left[\frac{z+\left[\left(\frac{a^2-b^2}{b^2}\right)b\left[\sin\left[\tan^{-1}\left[\frac{za}{b\sqrt{x^2+y^2}}\right]\right]\right]^3\right]}{\sqrt{x^2+y^2}-\left(\frac{a^2-b^2}{a^2}\right)a\left[\cos\left[\tan^{-1}\left[\frac{za}{b\sqrt{x^2+y^2}}\right]\right]\right]^3}\right], \tag{8}$$

$$\lambda = \begin{cases} \tan^{-1}\left(\frac{x}{y}\right), & x \ge 0 \\ 180^0 + \tan^{-1}\left(\frac{x}{y}\right), & x<0, y\ge 0 \\ -180+\tan^{-1}\left(\frac{x}{y}\right), & x<0, y<0 \end{cases}, \tag{9}$$

$$h = \frac{\sqrt{x^2+y^2}}{\cos(\varphi)} - \frac{a}{\sqrt{1-\left(\frac{a^2-b^2}{a^2}\right)\sin^2(\varphi)}}, \tag{10}$$

where a and b are the semi major and semi minor axis of the earth ellipsoid; x, y and z are ECEF coordinates.

Minimum values of D_{k0} are calculated using a real ephemeris data for each satellite, collected for 42 hours time of observation with U-blox GPS receiver ANTARIS AEK-4P. Using these data and Eqs. (8), (9), (10) the geodetic coordinates of satellites are computed. The ECEF coordinates of satellites x_{sk}, y_{sk}, z_{sk} are also calculated by ephemeris data, using standard algorithm [1, 2].

The ECEF coordinates of user could be calculated, using transformation equations from geodetic to ECEF coordinates given by [1, 5]:

$$x = \left[\frac{a}{\sqrt{1-\left(\frac{a^2-b^2}{a^2}\right)\sin^2(\varphi)}}+h\right]\cos(\varphi)\cos(\lambda) \tag{11}$$

$$y = \left[\frac{a}{\sqrt{1-\left(\frac{a^2-b^2}{a^2}\right)\sin^2(\varphi)}}+h\right]\cos(\varphi)\sin(\lambda) \tag{12}$$

$$z = \left[\frac{a}{\sqrt{1-\left(\frac{a^2-b^2}{a^2}\right)\sin^2(\varphi)}}\left[1-\left(\frac{a^2-b^2}{a^2}\right)\right]+h\right]\sin(\varphi) \tag{13}$$

User radius R_u is given by:

$$R_u = \sqrt{x_u^2+y_u^2+z_u^2}, \tag{14}$$

where x_u, y_u, z_u are ECEF coordinates of user position.

When the user is located on the geodetic normal of satellite the user longitude and latitude are coincided with the satellite longitude λ and latitude φ, and the user height h_u is considered to be 20 000 meters (the worst case).

The values of $\cos\alpha_k$ could be computed using the following expression [8]:

$$\cos\alpha_k = \frac{\vec{r}_{sk}\vec{R}_u}{r_{sk}R_u}, \tag{15}$$

where r_{sk} and R_u are the magnitudes of $\vec{r}_{sk}$ and $\vec{R}_u$.

From the above equation the product $R_u\cos\alpha_k$ could be derived as follows [8]:

$$R_u\cos\alpha_k = \frac{x_{sk}x_u + y_{sk}y_u + z_{sk}z_u}{r_{sk}}, \tag{16}$$

Eqs. (4), (5), (14) and (16) allow the values of X to be calculated, in the worst case for each satellite.

III. RESEARCHES

Various time intervals of approximation (t-t_n) are used to compute the values of X. Studies show that the values of X are unique for each satellite. They are the greatest for satellite SV27, while for satellite SV17 they are the smallest. So the worst case of approximation takes place for satellite SV27. In [3] is shown that the maximum possible time interval of satellite radius approximation, using Eq. (2) for SV27, is 89 seconds. This determines the maximum allowed time interval of user-to-satellite range approximation. The studies are made for time intervals from 10 to 90 seconds with a step of 10 seconds.

In Figs. 2 and 3 the values of X for satellite SV27 are illustrated. In Fig. 2 the results for 20 second time interval are shown. In Fig. 3 the results for 90 second time interval are presented.

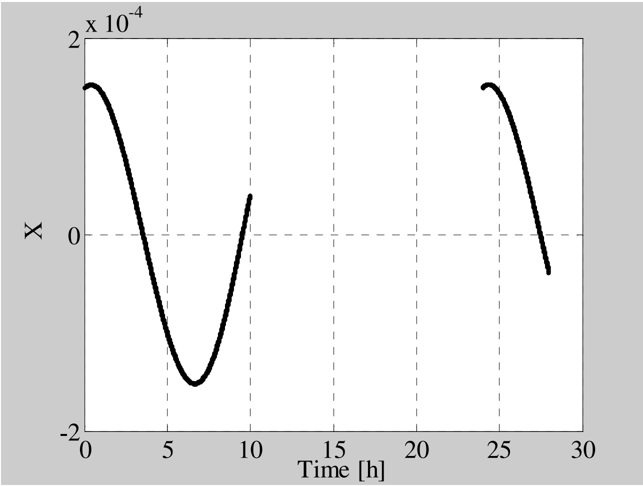


Fig. 2. Values of X in case of time interval of approximation 20 seconds

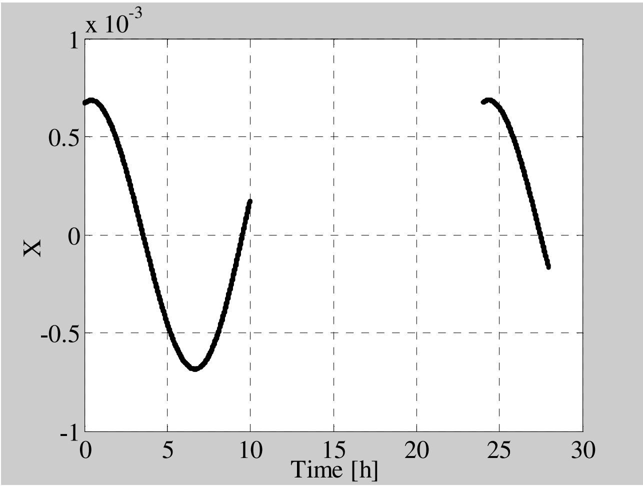


Fig. 3. Values of X in case of time interval of approximation 90 seconds

As it is seen in above figures there is absence of data for X. This is because the satellite has not been visible from the point of the GPS receiver at this time.

It is obvious that the values of X do not exceed 1×10^{-3} which is small enough to expand user-to-satellite range in series [8]:

$$D_{ak1} \approx D_{k0}\left(1+\frac{X}{2}-\frac{X^2}{6}+\frac{X^3}{48}-...\right) \qquad (17)$$

In order to simplify (17) the non-linear terms are rejected. In this case the final result of user-to-satellite range approximation is obtained as:

$$\begin{aligned} D_{ak1} \approx D_{k0}\left(1+\frac{X}{2}\right) = D_{k0} + \\ +\frac{B^2(t-t_n)^4+2AB(t-t_n)^3}{2D_{k0}}+ \\ +\frac{(A^2+2r_{sk0}B)(t-t_n)^2+2r_{sk0}A(t-t_n)}{2D_{k0}}- \\ -\frac{2R_u\cos\alpha_k[B(t-t_n)^2+A(t-t_n)]}{2D_{k0}} \end{aligned} \quad . \qquad (18)$$

It is known that the user–equivalent range error UERE budget depends on the space segment, control segment, and user segment sources [1]. UERE is 6.6 meters (1 sigma) for Precise Point Service (PPS), and 8 meters (1 sigma) for Standard Positioning Service (SPS) [1]. The errors due to ionospheric and tropospheric delays can be reduced at least with 50%, using the corresponding model [2], which parameters are transmitted in navigation data by every satellite. Thus the UERE could be decreased to 3 meters (1 sigma) [1, 2].

If the error of approximation

$$\Delta D_{apr} = D_{ak} - D_{ak1} \qquad (19)$$

is of much more less value than UERE, then it could be neglected. The conditions meeting these requirements are defined by:

$$\Delta D_{apr} \le 0.1 UERE \qquad (20)$$

Consequently the error of approximation should be bellow 0.3 meters.

Researches about the values of error approximation ΔD_{apr} for satellite SV27 (the worst case) are made for time approximation intervals from 10 to 90 seconds with a step of 10 seconds. Some of the results are shown in Fig. 4 and Fig. 5.

In Fig. 4 errors of approximation, when time interval of approximation is 60 seconds, as a function of observation time, are presented.

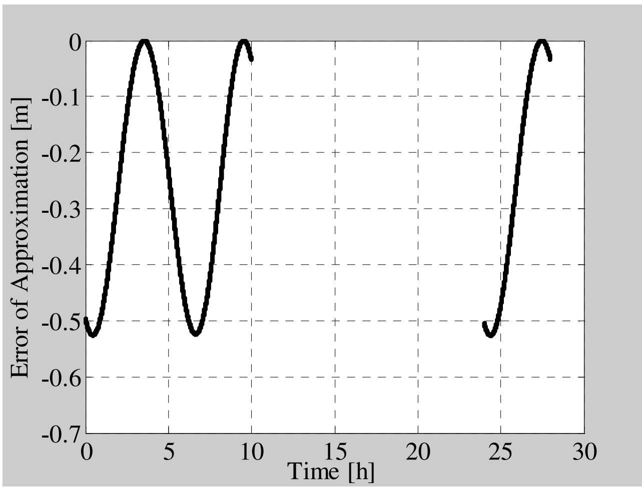


Fig. 4. Error of approximation in case of time interval of approximation 60 seconds

As it can be seen from Fig. 4 the maximum value of error is over 0.5 meters. This does not agree with Eq. (20) and consequently the time interval of approximation must be shorter.

In Fig. 5 errors of approximation, when time interval of approximation is 40 seconds, as a function of observation time, are presented.

Fig. 5 illustrates that the maximum value of error does not exceed 0.25 meters. This is in agreement with Eq. (20) and consequently the time interval of approximation 40 seconds could be used and even longer one could be used.

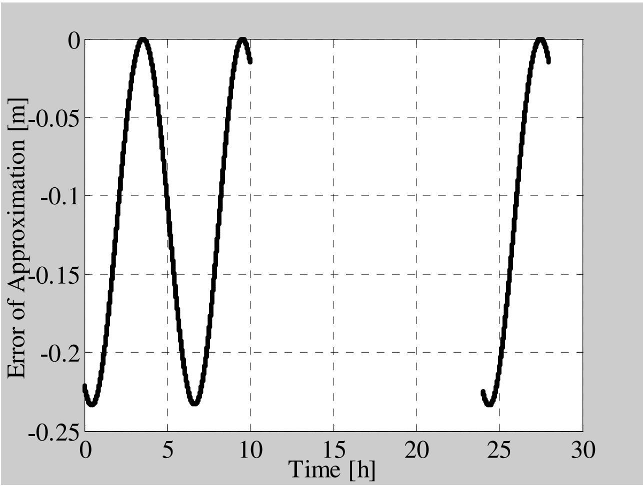


Fig. 5. Error of approximation in case of time interval of approximation 40 seconds

The results prove that the requirements to time interval of user-to-satellite range approximation are much stronger than the time interval of satellite radius approximation obtained in [3]. Hence the time interval of approximation must be chosen according to Eq. (20).

In Fig. 6 the results from studies of the error of approximation as a function of time interval of approximation are shown. The errors of approximation for the best (satellite SV17) and the worst (satellite SV27) case are presented.

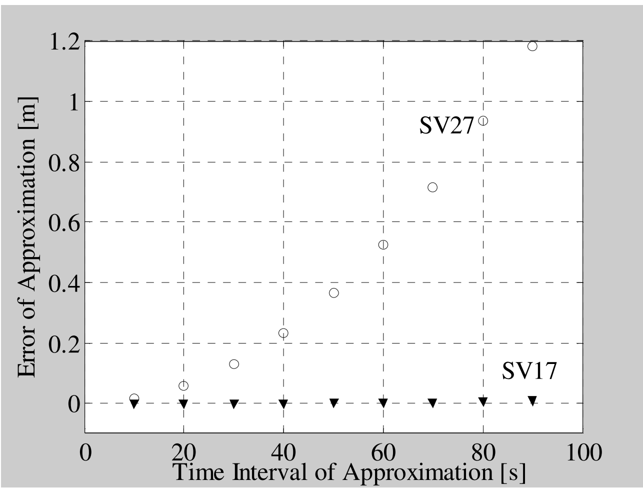


Fig. 6. Error of approximation for satellites SV27 and SV17.

As one can see in case of approximation for satellite SV17 the longer time interval of approximation could be used, because even for 90 seconds time interval the approximation error ΔD_{apr} is 0.0107 meters.

In the same time interval of approximation the approximation error ΔD_{apr} for satellite SV27 is 1.1839 meters, which goes far beyond the error bounder.

IV. Conclusion

Proposed approximation represents user-to-satellite range as direct function of time. Eqs. (16) and (18) show that this approximation can be used for modeling the measured pseudoranges in the way that allows using kalman filtering or least square algorithm. The obtained results indicate that the requirements to user-to-satellite range approximation are much stronger than to satellite radius approximation. Despite of this for one satellite radius approximation interval several time intervals of user-to-satellite range approximation could be used.

References

[1] E. D. Kaplan, *Understanding GPS Principles and Applications*, Artech House, London, 1996.

[2] *Interface Specification IS-GPS-200 Revision D*, ARINC Engineering Services, El Segundo, 2004.

[3] D.V.Dimitrov, M.S.Marinov, *Studies of Linear and Quadratic Approximations of GPS Satellite Radius Alteration in Time*, XLI International Scientific Conference ICEST, 29th June to 1st July 2006, Sofia, Bulgaria, 2006.

[4] J.Juang, *On GPS Positioning and Integrity Monitoring*, IEEE T-AES, Vol. 36, No. 1, 2000.

[5] J.Zhu, *Conversion of Earth-Centered Earth-Fixed Coordinates to Geodetic Coordinates,* IEEE T-AES, Vol. 30, No. 3, 1994

[6] K. Borre, K. de Jong, C. Pichot, *Subsidence Monitoring System Using Real-Time GPS Sensors*, ION GPS 2001, 11-14 September 2001, Salt Lake City, UT

[7] M. Marinov, G. Stanchev, *Use of Space Correlation of Satellite Move in GPS*, Proceeding of XXXVII International Scientific Conference ICEST, 16-18 October 2003, Sofia, Bulgaria, 2003.

[8] G. A. Korn, T. M. Korn, *Mathematical Handbook for Scientists and Engineers,* McGraw-Hill, 1968.

Second and Third Order Dispersion Influence on Pulse Propagation in the Presence of the Interference

Mihajlo C. Stefanovic[1], Dragan Lj. Draca[2], Aleksandra S. Panajotovic[3], Daniela M. Milovic[4]

Abstract – **Chromatic dispersion is one of the most important limiting factors in linear optical transmission systems. Pulse propagation under the common influence of second and third order dispersion is considered in this paper. Analytic solution of pulse shape at a receiver in the presence of interference, second and third order dispersion is determined in this paper, too. Interference is time and phase shifted in regard to useful signal. We consider the worst case, i.e. case when phase shift is π.**

Keywords – **Second and third order chromatic dispersion, Impulsive response, Coherent interference.**

I. Introduction

Chromatic dispersion limits transmission length in optical telecommunication systems that use single-mode fiber. Because of that, it plays a critical role in a propagation of optical pulse. In order to significantly increase transmission speed of optical system the influence of chromatic dispersion on signal propagation, especially its second and third order, must be considered [1].

Another limiting factor in optical transmission systems are incoherent and coherent interferences. Coherent interference is more problematic than incoherent interference because it cannot be controlled by optical filtering in the receiver [2, 3]. It is a reason for research such interference influence on pulse deformation when propagating along optical fiber under second and third order dispersion.

Until now, we studied an influence of interference on signal propagation along linear optical fiber solving nonlinear Schrödinger equation by "split-step" Fourier method [4, 5, 6]. In all of our previous papers, only second order dispersion was included in an investigation. In this paper we obtain general expression of Gaussian pulse at the receiver in the presence of interference, second and third order dispersion. Using this result we consider interference influence on pulse propagation along linear single mode fiber under second and third order dispersion.

[1]Mihajlo C. Stefanovic is with Faculty of Electronic Engineering, University of Nis, Aleksandra Medvedeva 14, 18000 Nis, Serbia and Montenegro, E-mail: misa@elfak.ni.ac.yu.

[2]Dragan LJ. Draca is with Faculty of Electronic Engineering, University of Nis, Aleksandra Medvedeva 14, 18000 Nis, Serbia and Montenegro, E-mail: draca@elfak.ni.ac.yu.

[3]Aleksandra S. Panajotovic is with Faculty of Electronic Engineering, University of Nis, Aleksandra Medvedeva 14, 18000 Nis, Serbia and Montenegro, E-mail: alexa@elfak.ni.ac.yu.

[4]Daniela M. Milovic is with Faculty of Electronic Engineering, University of Nis, Aleksandra Medvedeva 14, 18000 Nis, Serbia and Montenegro, E-mail: dacha@elfak.ni.ac.yu.

II. Analytic Solution of Resulting Pulse at the Receiver

Pulses from many lasers can be approximated by a Gaussian shape. Because of that signal with Gaussian envelope is very often found as a useful signal in optical systems [1, 7, 8]. Gaussian optical pulse is given by:

$$s(t)=\sqrt{P_0}\,e^{-\left(\frac{t}{T_0}\right)^2+j\omega_o t} \tag{1}$$

where ω_0 is the angular frequency of useful signal, P_0 is optical pulse peak power, $T_0^{'}=T_0/\sqrt{2}$ is pulse half width (at 1/e intensity point).

Coherent interference is of the same frequency as a useful signal and it can be time and phase shifted in regard to useful signal. It can be written as:

$$s_i(t)=\sqrt{P_i}\,e^{-\left(\frac{t}{T_0}-b^{'}\right)^2+j(\omega_o t+\varphi)} \tag{2}$$

where P_i is interference peak power, b ($b=b^{'}T_0$) and φ are time and phase shift, respectively. The envelope and phase of resulting signal $s_r(t)$ are [4]:

$$|s_r(t)|=$$

$$\sqrt{\left(\sqrt{P_0}e^{-\left(\frac{t}{T_0}\right)^2}\right)^2+2\sqrt{P_0}e^{-\left(\frac{t}{T_0}\right)^2}\sqrt{P_i}e^{-\left(\frac{t}{T_0}-b^{'}\right)^2}\cos\varphi+\left(\sqrt{P_i}e^{-\left(\frac{t}{T_0}-b^{'}\right)^2}\right)^2} \tag{3}$$

$$\psi(t)=arctg\frac{\sqrt{P_i}e^{-\left(\frac{t}{T_0}-b^{'}\right)^2}\sin\varphi}{\sqrt{P_0}e^{-\left(\frac{t}{T_0}\right)^2}+\sqrt{P_i}e^{-\left(\frac{t}{T_0}-b^{'}\right)^2}\cos\varphi} \tag{4}$$

A general expression of impulse response r (t, L) for arbitrary input pulse is [7, 9]:

$$r(t,L)=\frac{1}{2\pi}e^{-\alpha L}e^{j(\omega_0 t-\beta_0 L)}\cdot$$

$$\cdot\int_{-\infty}^{\infty}F(\omega)e^{j\left(\omega t-\frac{1}{2!}\beta_2 L\omega^2-\frac{1}{3!}\beta_3 L\omega^3\right)}d\omega \tag{5}$$

when optical fiber is under second and third order dispersion. Second order dispersion is very often dominated in linear optical fiber. Then influence of third order dispersion can be neglected. The situation changes for propagation of ultrashort pulses (with widths in the femtosecond range) in the vicinity of λ_D (λ_D – zero-dispersion wavelength) [1]. That case we consider in this paper. $F(\omega)$ is Fourier transform of the input pulse. If we assume that coherent interference appears at the beginning of the optical fiber and that phase shift of interference is π (the worst case), then Fourier transform of the resulting pulse is:

$$F(\omega)=\sqrt{P_0}T_0\sqrt{\pi}e^{-\frac{\omega^2T_0^2}{4}}\left(1-\sqrt{\frac{P_i}{P_0}}e^{-j\omega b'T_0}\right) \tag{6}$$

If we substitute (6) in (5) and use some mathematical operations, the pulse response for our case becomes:

$$r(t,L)=\frac{\sqrt{P_0}T_0}{2\sqrt{\pi}}e^{-\alpha L}e^{j(\omega_t t-\beta_0 L+\theta(t))}\sqrt{I_1^2(t)+I_2^2(t)} \tag{7}$$

$$\theta(t)=arctg\frac{I_2(t)}{I_1(t)} \tag{8}$$

with

$$I_1(t)=\int_{-\infty}^{\infty}e^{-\frac{\omega^2T_n^2}{4}}\left[\left(1-\sqrt{\frac{P_i}{P_0}}\cos\left(b'T_0\omega\right)\right)\cos\left(\omega t-b_2\omega^2-b_3\omega^3\right)-\right.$$

$$\left.-\sqrt{\frac{P_i}{P_0}}\sin\left(b'T_0\omega\right)\sin\left(\omega t-b_2\omega^2-b_3\omega^3\right)\right]d\omega \tag{9}$$

$$I_2(t)=\int_{-\infty}^{\infty}e^{-\frac{\omega^2T_n^2}{4}}\left[\left(1-\sqrt{\frac{P_i}{P_0}}\cos\left(b'T_0\omega\right)\right)\sin\left(\omega t-b_2\omega^2-b_3\omega^3\right)+\right.$$

$$\left.+\sqrt{\frac{P_i}{P_0}}\sin\left(b'T_0\omega\right)\cos\left(\omega t-b_2\omega^2-b_3\omega^3\right)\right]d\omega \tag{10}$$

where

$$b_2=\left(\frac{\beta_2L}{2!}\right)$$

$$b_3=\left(\frac{\beta_3L}{3!}\right) \tag{11}$$

Equations (7)-(10) represent a general expression of Gaussian pulse at the end of optical fiber under second and third order dispersion in the presence of coherent interference.

III. NUMERICAL RESULTS

Results of our research are shown on following figures. Fiber length is expresed via dispersion length which is for n-th order of dispersion given by:

$$L_D=\frac{\left(T_0'\right)^n}{|\beta_n|} \tag{12}$$

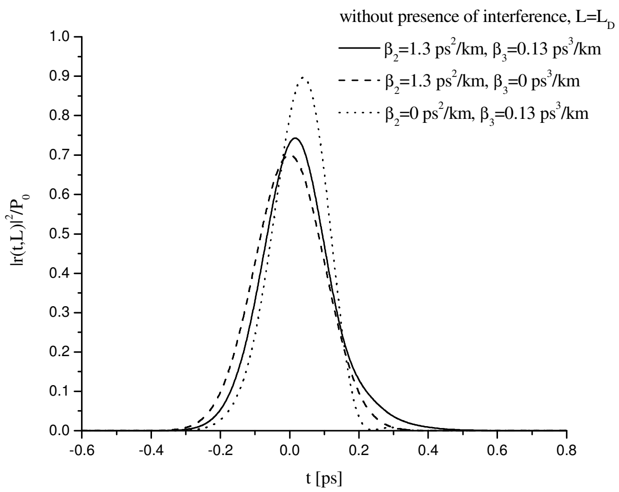
In optical fiber under common influence of second and third order dispersion L_D for $n = 2$ and $n = 3$ is equal. The following values of parameters are used in all cases: $T_0' = 0.1$ ps, $\alpha = 0$ dB/km (optical losses are neglected in the fiber), $\beta_2 = 1.3$ ps^2/km (dispersion coefficient of second order), $\beta_3 = 0.13$ (dispersion coefficient of third order) ps^3/km and $\varphi = \pi$.

Figure 1 shows propagation Gaussian pulse along the optical fiber. Well known facts can be noticed from Fig. 1:

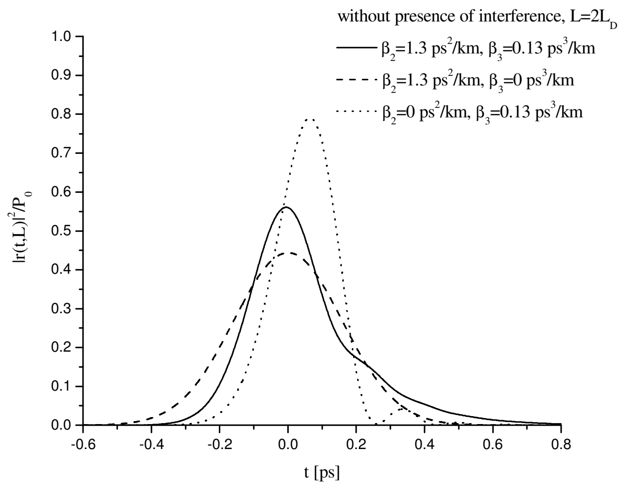
- second order dispersion influence on pulse broadening.
- third order dispersion influence on pulse deformation, i.e. induces strong pulse oscillation in positive region of t.

This figure shows that third ordere disperion decreases pulse broadening induced by second order dispersion. But, it provokes moving of peak power to left.

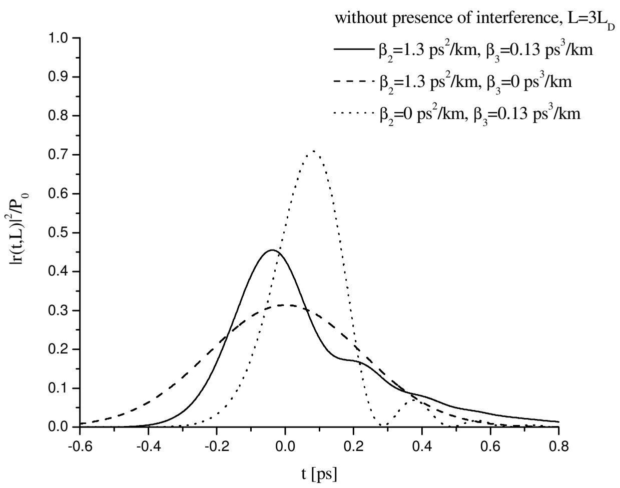
Figures 2 (SIR = 5 dB) and 3 (SIR = 10 dB) show resulting pulse shape at the receiver for different values of the interference time shift. Interference is the most ruinous when it is not time shifted in regard to useful signal. Negative time shift induced bigger deformation of resulting pulse than positive time shift. Following figures show that in this case we can hardly concluse that Gaussian pulse propagated along the fiber. Presence of interference time shift induces stronger intersymbol interference (ISI), too. Comparasion of Figs. 2 and 3 show that pulse deformation is insofar larger as SIR (Signal – to – Interference Ratio) is less.



a)

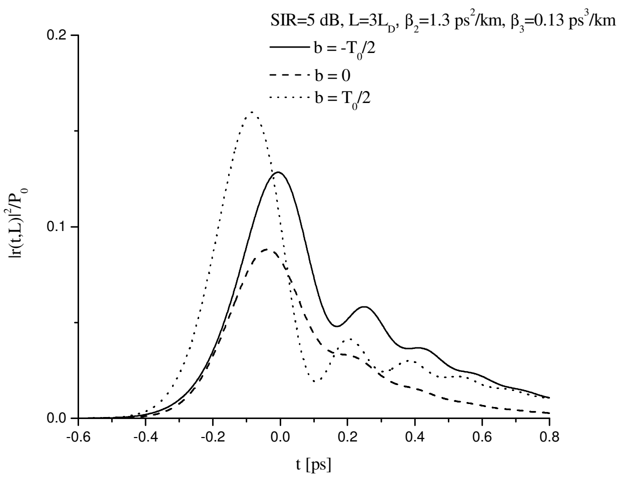


b)

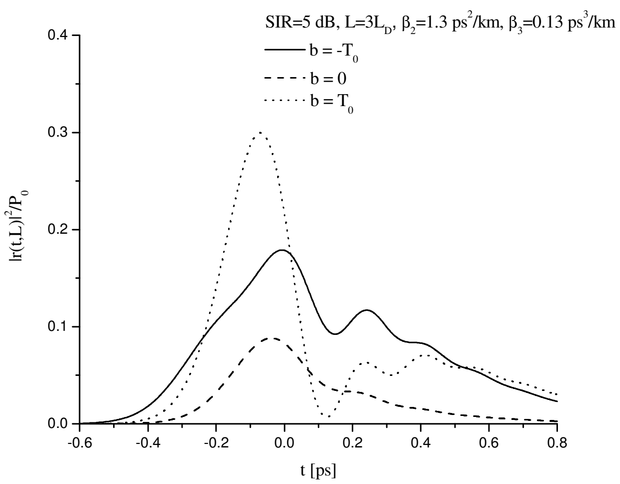


c)

Fig. 1. Propagation Gausian pulse along the linear optical fiber.

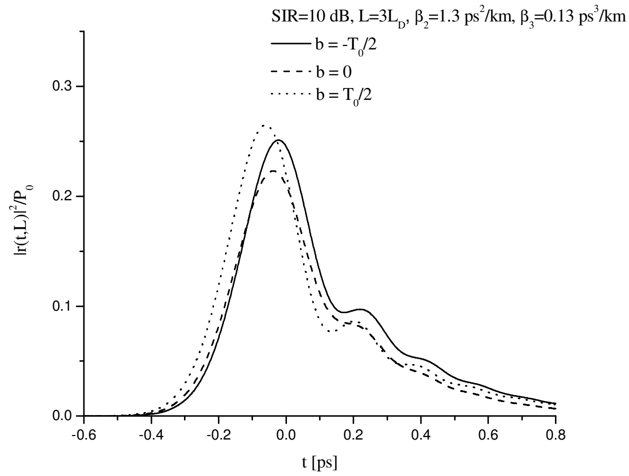


a)

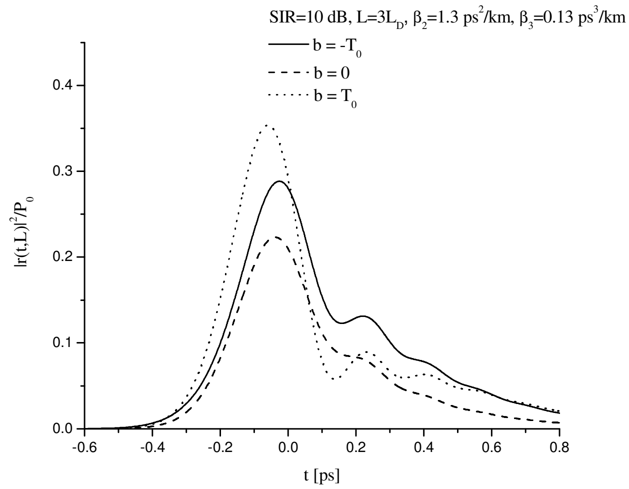


b)

Fig. 2. Resulting pulse shape at the receiver for different values of time shift (SIR = 5 dB).



a)



b)

Fig. 3. Resulting pulse shape at the receiver for different values of time shift (SIR = 10 dB).

IV. Conclusion

Linear optical fiber is very often used as medium for pulse propagation through telecommunication system. Chromatic dispersion is one of the main factors limiting transmission length in these systems. It is reason for studing common influence of second and third order dispersion on pulse shape along the fiber. Results of this paper show that third order dispersion decreases pulse broadening but moves peak power to left. If there is second order dispersion in the optical fiber, presence of third order dispersion will be advisable because error in detection process is smaler in that case. Coherent interference may be detrimental in optical telecommunication systems because it additionally distorts optical pulse and can cause significant rise to intersymbol interference. We analyzed the worst case of coherent interference, i.e. $\varphi = \pi$. The time shift of interference play significant role. It makes the pulse response take the shape of long trailing skirts which influence on distant pulses. This leads to stronger intersymbol interference. Until now, these problems we considered using "split-step" Fourier method for solving Schrödinger equation. This method is very sensitive on the number of fiber segments and their width. Obtained results are very important because we used analytical method that has great accuracy. Performances of optical system determined on this way are very objectiv.

REFERENCES

[1] P. Agrawal, *Nonlinear fiber optics*, Academic Press INC., Boston-San Diego-New York-London-Tokyo-Toronto, 1995.

[2] Y. Shen, K. Lu and W. Gu, "Coherent and incoherent crosstalk in WDM optical networks"; *J.Light. Techn.*, Vol.17, No. 5, pp. 759-764, 1999.

[3] P. J. Legg, M. Tur and I. Andonovic, "Solution paths to limit interferometric noise induced performance degradation in ASK/direct detection lightwave networks", *J. Light. Techn.*, Vol. 14, No. 9, pp. 1943-1954, 1996.

[4] M. Stefanovic, D. Draca, P. Spalevic and A. Panajotovic, "The influence of crosstalk signal interference to signal propagation along the nonlinear and dispersive fiber"; *J. Opt. Commun.*, Vol. 26, No. 1, pp. 9-12, 2005.

[5] M. Stefanovic, D. Draca, P. Spalevic and A. Panajotovic, "Performance of IM-DD optical system in the presence of interference at the input of the fiber"; *Nonlin. Phenom. Complex. Syst.*, Vol. 6, No. 4, pp. 870-877, 2003.

[6] M. Stefanovic, D. Draca, and A. Panajotovic, "The common influence of time shift and appearing place of interference on signal propagation along optical fiber"; *Electron. and Electric. Engin.*, Vol. 57, No. 1, pp. 14-19, 2005.

[7] M. Amemiya, "Pulse broadening due to higer order dispersion and its transmission limit", *J. Light. Techn.*, vol. 20, no. 4, pp. 591-597, 2002.

[8] P. Johannisson, D. Anderson, A. Berntson and J. Märtensson, "Generation and dynamics of ghost pulses in strongly dispersion-managed fiber optic communication systems", *Optical Letters*, vol. 26, pp. 1227, 2001.

[9] A. Panajotovic, D. Milovic, and A. Mitic, "Boundary Case of Pulse Propagation Analytic Solution in the Presence of Interference and Higher Order Dispersion"; *TELSIKS 2006, Conference Proceeding.*, pp. 547-550, Nis, Serbia and Montenegro,2005.

Impulse Response of SI Polymer Optical Fibres

Zwetomir Zwetkov[1]

Abstract – **Using the power flow equation to describe the mode coupling in the fibre and consider the power loss dependent on mode group we have built a mathematical model of the SI polymer fibre, which solution yields all quantities of interest of the fibre: power distribution, attenuation and impulse response as functions of length by UMD launching condition.**

Keywords – **POF, time-dependent power flow equation, impulse Response, Fourier Approach.**

I. Introduction

Presently, polymer optical fibres are of interest in communication systems in close-up range, automotive, home equipment etc. because of less stringent requirements ofoptical carriers and splices, cheaper connectors and low costs by the production. However, the mode dispersion caused by the relative thick core and high value of power loss reduce the transmission distance and throughput. Moreover the mode coupling caused by the random variations of the core index or the random irregularities of the fibre wall complicates the characteristics of the pulse propagation. It is purpose of this paper to present a satisfactory approach to obtain the impulse response of polymer optical fibres.

II. Power Loss

Step index polymer optical fibres are thick core light guides that lead as many as million of modes. The core is larger than the wavelength of the light. The fibre can be analysed with a geometric ray-tracing model. Moreover, an important consequence of this huge number of modes is that light polarization is not preserved along the fibre. There is no correlation between input and output polarization states, except for very short POF (< 0.5m). Further, we base our considerations on this ray model considering lengths in the range of many hundred meters.
One of the most important processes encountered by light as it passes through a fibre is attenuation. When passing through an optical fibre of the length *L*, the power of the light decreases exponentially according to

$$P(L) = P(0) \cdot exp(-\alpha \cdot L) \tag{1}$$

where $P(0)$ is the input power of the light and α is the attenuation coefficient in 1/km which depends on mode number. The attenuation coefficient can be split into three terms that describe three different phenomena of attenuation, i.e. length difference for each mode, the difference in the number of reflections for each mode and *Goos – Haenchen* effect.

At first, we consider the length difference for each mode. The accumulated path length can be written as

$$L_\theta = \frac{L}{cos(\theta)} \tag{2}$$

where θ is the angle of incidence. With the core attenuation coefficient α_{core} the path-dependent loss can be written as

$$\alpha_l = \alpha_{core} \cdot L_\theta \tag{3}$$

The second phenomenon of power loss is the attenuation caused by ray reflection on the clad-core interface. Rays with a greater propagation angle reflect more often than rays with a smaller angle of incidence θ. The length-dependent attenuation is given by

$$\alpha_r = \frac{L \cdot tan(\theta)}{2\rho} \cdot ln(R) \tag{4}$$

with the reflection coefficient *R* and the core radius ρ.

Now, the third cause for attenuation, the *Goos – Haenchen* effect, is taken into account. In [4] the extra attenuation coefficient caused by the *Goos - Haenchen* effect α_{GH} is described by

$$\alpha_{GH} = \frac{d(\theta)}{\rho} \cdot \frac{\alpha_{clad}}{cos(\theta)} \tag{5}$$

where α_{clad} is the attenuation coefficient of the cladding and the shift - depth of the reflection plane into the optically thinner cladding is expressed as $d(\theta)$ and the corresponding formula is given in [5] as

$$d(\theta) = \frac{\lambda}{2\pi\sqrt{n_{core}^2 cos^2(\theta) - n_{clad}^2}} \tag{6}$$

Now the attenuation coefficient α can be given as

$$\alpha = \alpha_{core}\frac{1}{\cos\theta} + \frac{\tan(\theta)}{2\rho} \cdot \ln(R) + \alpha_{GH} \tag{7}$$

for the power distribution we yield

[1] Zwetomir Zwetkov, Technical University of Munich,
e-mail: cvetomir.cvetkov@mytum.de

$$P(\theta,z) = P(\theta,0)\cdot exp\Big(-\alpha_{core}\frac{z}{cos(\theta)} - \frac{z\cdot tan(\theta)}{2\rho}\cdot ln(R) - \alpha_{GH}z\Big) \quad (8)$$

III. Diffusion Equation

By approximation of the modes by a continuum Gloge [1] obtained the fundamental coupled-power equation

$$\frac{\partial P(\theta,z)}{\partial z} = -\alpha(\theta)P(\theta,z) + \frac{D}{\theta}\frac{\partial}{\partial\theta}\Big(\theta\frac{\partial P(\theta,z)}{\partial\theta}\Big) \quad (9)$$

which describes the power flow of the coupled modes of stepp – index fibres, with

$P(\theta, z)$	angular power distribution
θ	propagation angle with respect to the core axis
D	coupling coefficient
α	modal attenuation

However, this equation does not consider the time differences of the different modes by the process of power diffusion and blur the time axis so that the delay of the different power parts cannot be determined. On the contrary, the determination of the impulse response requires the consideration of the time differences of the different modes. Gloge develops his idea further building in additionally the time into the equation. With the total variation in P

$$dP = \frac{\partial P}{\partial z}dz + \frac{\partial P}{\partial t}dt \quad (10)$$

he write in [2] the time-dependent power flow equation

$$\frac{\partial P}{\partial z} + \frac{dt}{dz}\frac{\partial P}{\partial t} = -\alpha(\theta)P(\theta,z) + \frac{D}{\theta}\frac{\partial}{\partial\theta}\Big(\theta\frac{\partial P(\theta,z)}{\partial\theta}\Big) \quad (11)$$

The derivate dt/dz is the delay of the power part $P(\theta)$ for a unit length and can be written for the ray model of the SI POF as

$$\frac{dt}{dz} = \frac{d}{dz}\frac{z\cdot n_{core}}{c_0\cdot cos(\theta)} = \frac{n}{c_0\cdot cos(\theta)} = \frac{n}{c_0} + \frac{n}{c_0}\Big(\frac{1}{cos(\theta)} - 1\Big) = \tau_0 + \tau_{rel} \quad (12)$$

The common delay to all modes n/c is ignored further.

IV. Fourier Approach

With the Fourier transformation

$$p(\theta,z,\omega) = \mathcal{F}\{P(\theta,z,t)\} \quad (13)$$

we can write Eq. (11) in the form

$$\frac{\partial p}{\partial z} = -(\alpha(\theta) + j\omega\tau_{rel})p + \frac{D}{\theta}\frac{\partial}{\partial\theta}\Big(\theta\frac{\partial p}{\partial\theta}\Big) \quad (14)$$

The first term in Eq. (14) describes attenuation and time shift of the different modes and the second one the diffusion respectively. There are two different physical phenomena, that act simultaneously. We can write formally Eq. (14) in the form

$$\frac{\partial p}{\partial z} = (\hat{L} + \hat{D})p \quad (15)$$

where $\hat{L}$ is a differential operator that accounts for attenuation and time difference of the different modes and $\hat{D}$ is the operator that governs the effect of mode coupling. We can give an approximate solution from (15) by assuming that over a small distance Δz the attenuation and coupling effects act independently. Mathematically,

$$p(\theta, z+\triangle z,\omega) \approx exp(\triangle z\hat{L})exp(\triangle z\hat{D})p(\theta,z,\omega) \quad (16)$$

The split approach ignores the no commutating nature of the operators $\hat{L}$ and $\hat{D}$. To improve the accuracy, the split method can be modified by adopting a different procedure to propagate the optical pulse over one segment Δz. In this procedure Eq. (16) is replaced by

$$p(\theta, z+\triangle z,\omega) \approx exp(\frac{\triangle z}{2}\hat{L})exp(\triangle z\hat{D})exp(\frac{\triangle z}{2}\hat{L})p(\theta,z,\omega) \quad (17)$$

For equation (14) the solution for the differential operator $\hat{L}$ can be given as

$$p(\theta, z+\frac{\triangle z}{2},\omega) = p(\theta,z,\omega)\cdot exp(-\alpha(\theta)\frac{\triangle z}{2})\cdot exp(-j\omega\tau_{rel}\frac{\triangle z}{2}) \quad (18)$$

An analytical solution for the differential operator $\hat{D}$ can be given with the Bessel functions. On the contrary, we use the finite-difference method to solve it numerically.

$$p(\theta, z+\triangle z,\omega) = \int_z^{z+\triangle z}\frac{D}{\theta}\frac{\partial}{\partial\theta}\Big(\theta\frac{\partial p}{\partial\theta}\Big)dz \quad (19)$$

Applying Eq. (18) und Eq. (19) successively with the procedure in Eq. (17) we obtain the solution at the end of the fibre

$$p(\theta, L, \omega) \quad (20)$$

The total output over all angles θ is obtained from the integration

$$H(L,\omega) = \int_0^{\theta_{max}} p(\theta,L,\omega)d\theta \quad (21)$$

With the reverse Fourier transformation we obtain the impulse response of the fibre

$$h(L,t) = \mathcal{F}^{-1}\{H(L,\omega)\} \qquad (22)$$

To find the quantity power distribution we do reverse Fourier transformation and integrate over all times

$$P(\theta,L) = \int_0^{t_{max}} \mathcal{F}^{-1}\Big\{p(\theta,L,\omega)\Big\}dt \qquad (23)$$

A last integral evaluation

$$P_{out}(L) = \int_0^{\theta_{max}} P(\theta,L)d\theta \qquad (24)$$

gives the total output power, which can be used to calculate the loss.

V. Simulation Results

Table I presents all important parameters, that we have adjusted to obtain the power distribution and impulse response along the fibre. Impulse responses by different fibre lengths are shown in Fig. 1. The process of diffusion does not influence heavy the impulse response after the first twenty meters, whose curve falls rapidly for longer transit times, that indicates the mode-dependent attenuation. For long fibre the diffusion effects a transition of the impulse form from an exponential shape to the Gaussian shape.

TABLE I
PARAMETER

Parameter	Value
Velocity of light c_0	$3 \cdot 10^8$ m/s
Radius of fibre ρ	$0.5 \cdot 10^{-3}$ m
Core refractive index n_{core}	1.492
Wavelength λ	650 nm
Numerical Aperture	0.5
Core attenuation coefficient α_{core}	120 dB/km
Cladding attenuation coefficient α_{clad}	5000 dB/km
Reflection factor R	$1 - 10^{-4}$
Diffusion coefficient D	$1.5 \cdot 10^{-4}$ rad^2/m

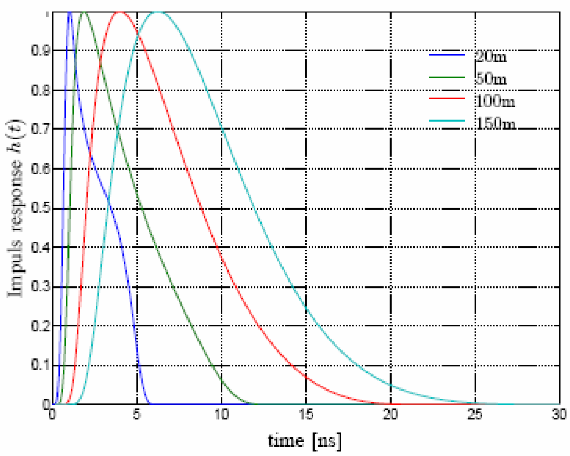


Fig. 1. Normalized impulse response according to Eq. (22) plotted for different fibre length

The power distribution along the fibre is shown in Fig. 2. We feed the fibre from a light emitting diode, that excites equally all modes. The blue curve shows the power distribution on the front of fibre. Modes with a greater number carry more power than lower modes, that propagate with a smaller angle of incidence because of the power contained in their differential solid angle. The diffusion concentrates slowly the power in the middle of the fibre and affects at the same time an averaging of the transit times of the different modes. The coupling length is achieved after 150m. The brown-ish curve results from the overlap of lilac and dark yellow. That indicates that the impulse shape does not change anymore after the coupling length.

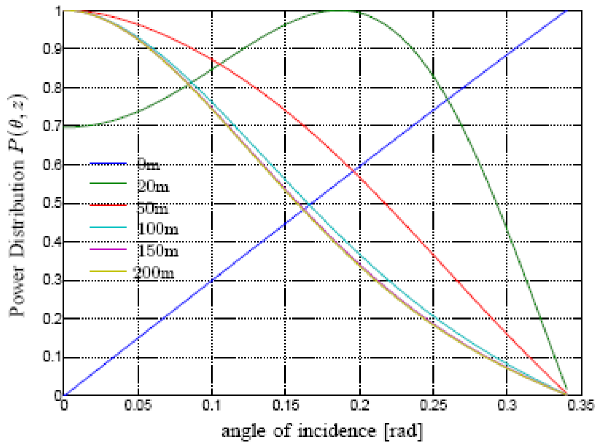


Fig. 2. Normalized power distribution at different fibre lengths calculated with (23) by UMD launching condition

VI. Conclusion

We have introduced shortly the attenuation of a SI polymer fibre based on optical rays model. Under the assumptions of mode coupling occurring between nearest neighbours using the time-dependent power diffusion equation we have built the base description of a step-index polymer optical fibre. We have developed a method to solve this equation transforming it into the frequency domain and giving an analytical and numerical solution for attenuation and diffusion respectively so that power distribution, attenuation and impulse response can be obtained.

Two further extensions of this work seem desirable: the building-in of the effect of mode conversion and to take the stochastical nature of diffusion into account.

As an example of applications, the impulse response can be useful to develop channel model of SI POF's, which can be used by several electrical methods of equalization to improve their transmission characteristics.

References

[1] D. Gloge, Optical Power Flow in Multimode Fibers, *Bell Syst. Tech. J.*, vol. 51, pp. 1767-1783, 1972.
[2] D. Gloge, Impulse response of Clad Optical Fibers, *Bell Syst. Tech. J.*, vol. 52, pp. 801-816, 1973.
[3] Oppenheim, R. Schafer, *Digital Signal Processing*, Prentice Hall, 1975.
[4] C.A. Bunge, Polymerfaser Dämpfungs- und Ausbreitungsmodell, *M.S.Thesis*, Technology Center of Deutsche Telekom AG, 1999.
[5] S. Geckeler, Lichtwellenleiter für die optische Nachrichten übertragung, Springer Verlag, Berlin 2001.

Analysis of the reasons for limiting the dynamic range of the signals in CATV systems

Lidia T. Jordanova[1], Valentin I. Topchiev[2]

Abstract: **In this article are considered the principals of building up a CATV System and is given a general characteristic of the elements that build them up. It is defined the dynamic range of the signals in CATV systems and also are analyzed the reasons for their limitation. It is discussed the different sources of noise and nonlinear distortion in the optical and coaxial part of the system.**

Keywords: **CATV, dynamic range, noise, nonlinear distortion, laser, photodiode receiver, optical fibre, optical and RF amplifier.**

I. INSTRUCTION

Community antenna television (CATV) systems are utilized for providing general-access and extra services of enormous number of customers. General-accessible services include fiery and satellite radio and TV programs. In this case the signals are transmitted from one provider to many subscribers. The extra services include the high-speed internet, the IP telephony, the video-on-demand, the home automation and control etc. The extra services are individual and the signals are transmitted from one provider to one subscriber, which requires also the building up of reverse channel [1].

In order to project a CATV system and optimize the dynamic range of the RF signals in critical points, it is necessary to analyze the reasons for the noise and nonlinear distortion appearance. Moreover, a very good knowledge of the architecture of the CATV system and the parameters of the components that build it up (lasers, photodiode receivers, optical fibres, optical and RF amplifiers) are required [2].

II. HIERARCHICAL PRINCIPLE OF BUILDING UP A CATV SYSTEM

The contemporary CATV system is a type of hybrid fibre coaxial (HFC) system [3]. As a carrier area in the optical part are used single-mode optical fibres for two ranges of the wavelength – 1310 nm and 1550 nm [4]. In the coaxial part is used standard coaxial cable with wave resistance 75 Ω.

The architecture of the HFC system is built in a hierarchal principle and is shown on figure 1. In the general case the system includes four hierarchical levels [1].

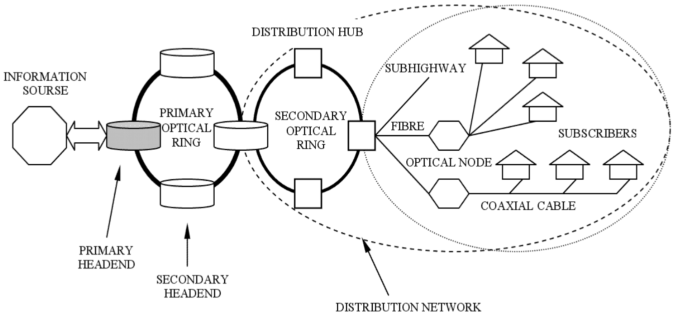


Fig. 1. Architecture of CATV system

The first level in the hierarchy of the CATV system is a superhighway system, which includes primary optical ring, primary and secondary headends. The primary headend combines few small autonomous systems and provides service to more than 200 000 subscribers, while the seconddary ones are local and serve usually to between 50 000 and 200 000 subscribers [5].

The second hierarchical level is a highway system (secondary optical ring), providing connection between the distribution hubs and the optical rings [6]. One distribution hub serve up to 16 distribution points or 40 000 subscribers.

The third level in the hierarchy of the system is sub-highway network. It connects the distribution hubs with the fibre nodes, where the conversion of the light stream into electrical current occurs and vice versa. The fibre node has two or four inputs/outputs and is able to serve between 250 and 1 000 subscribers [7].

The fourth hierarchical level is the subscriber system, which is a coaxial cable distribution network (CCDN). It connects the optical nodes with the subscriber's home.

The tendency in building the CATV system is the optical fibre to reach the customer's home. Fibre to the home (FTTH) is the ideal optical communication architecture. FTTH meets the requirements for high-speed data transmission, high-fidelity sound, as well as high-fidelity video signal [8, 9].

III. GENERAL CHARACTERISTIC OF THE CATV SYSTEM ELEMENTS

The main components which build up the optical channel of one CATV system when using direct (a) and external (b) modulation of the laser are shown on figure 2. In the first case, the modulating signal is added to the laser pre-voltage, whilst in the second – goes to the input of the external modulator [10].

In CATV System, as a light-source are preferred laser diodes of the distributed feedback (DFB), distributed Bragg

[1]Lidia T. Jordanova is with the Faculty of Communications and Communications Technologies, Technical University, Kliment Ohridski 8, 1000 Sofia, Bulgaria, e-mail: jordanova@tu-sofia.bg
[2] Valentin I. Topchiev is with the Faculty of Communications and Communications Technologies, Technical University, Kliment Ohridski 8, 1000 Sofia, Bulgaria, e-mail: vtsc@mail.bg

reflector (DBR), vertically cavity surface emitting laser (VCSEL). They emit in a very narrow frequency band and have low level of the insertion noise and nonlinear distortions [11, 12]. The most frequently used optical modulators in these systems are: Mach-Zehnder modulator, electro-absorption modulator and phase modulator [13, 14].

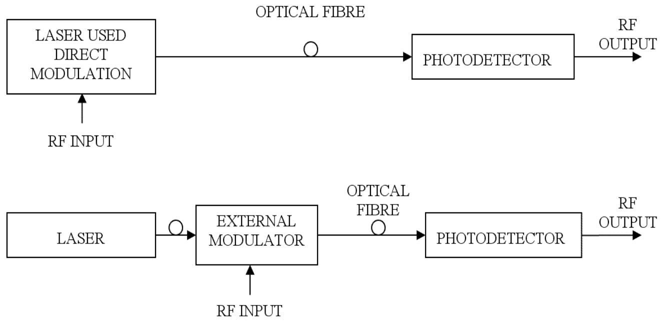


Fig. 2. Optical channel with direct a) and external b) modulation

The optical fibers used in the range of 1310 nm have insertion loss about 0.35 dB/km, while in the range of 1550 nm – about 0.2 dB/km. It is typical for the optical fibres is the phenomenon dispersion (chromatic and polarization), which leads to transmission speed and optical line length limitation. If $\lambda = 1310$ nm then the chromatic dispersion in single-mode fibre decreases to zero, while if $\lambda = 1550$ nm it has comparatively high values, which requires the use of fibres with displace dispersion. To compensate the loss in the optical fibre and the passive components are mainly used erbium-doped fiber amplifiers (EDFA) and Raman amplifier [15, 16].

The main element in the structure of the optical receiver is the photodiode (avalanche and PIN), whose way of working is based on the photoelectric effect. Moreover, the optical receiver includes circuit for impedance coherences, RF amplifier and control and alarm system [16].

The CCDN of the CATV System ensures the connection of its optical part with the subscribers. To compensate the insertion loss in the coaxial cable and the passive elements are used RF amplifiers [17].

IV. Determination of the Dynamic Range of the Sygnal in Catv Systems

The dynamic range is defined as the difference between the largest and the smallest signal power, which could be feed to the input of the communication channel. The maximal value of the input RF power is limited by the largest acceptable level of the nonlinear distortions products in the channel output. The minimum level of the input signals is limited by the minimal carrier to noise ratio (CNR) [18].

The bigger area range of the CATV System leads to increased number of sequentially plugged into the network sources of noises and nonlinear distortions which reduces the dynamic range of the input RF signals. This is one of the reasons for the limitation of the communication channel length.

The dynamic range of the signals in the optical channel is defined by the minimum and maximal level of RF signals, entering into the input of the laser transmitter. There are three types dynamic range – compression dynamic range (CDR), spurious-free dynamic range (SFDR) and working dynamic range (WDR), which are defined on figure 3 [1, 2, 19].

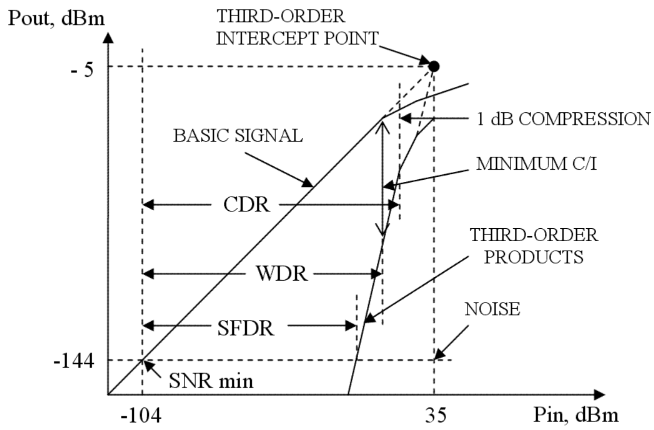


Fig. 3. Dynamic range of the signals in the optical channel

The dynamic range of the signals of CCDN depends mainly on the parameters of the RF amplifier, the number of the cascade connected amplifiers and the number of transmitted radio- and TV channels.

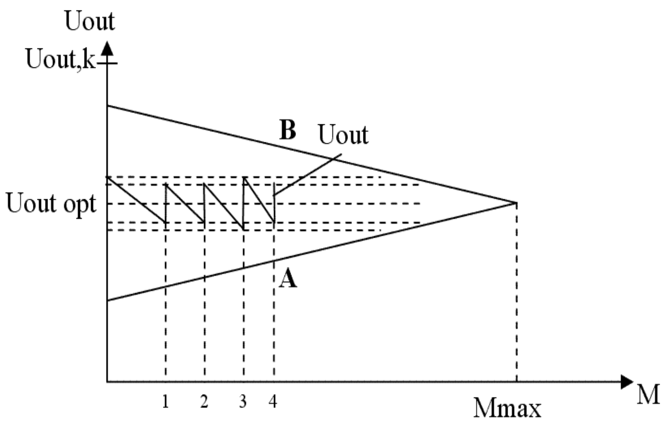


Fig. 4. Dynamic range of the signals in CCDN

The acceptable limits of variation of the RF signal levels in each RF amplifier outputs are shown on figure 4. The lowest level ("A") is defined by the minimum CNR in the subscriber's contact. The limitation of the RF signal maximal level ("B") is related to the acceptable nonlinear distortions in the subscriber's contact [20].

V. Analisis of the Reasons for Occurrence of the Noises in Catv System

When transmitting signals through CATV System, noises from different sources are added. It is very important when defining the total noise to know the kind, the typical properties and the area of generating the noise. The most significant influence when defining the total system noise has the thermal noise, the shot noise and the relative intensity noise (RIN) of the laser. In the CATV System are also generated many other noises, whose influence is not so essential.

The thermal noise is generated while electrical current passes through each electronic element and depends on the working temperature, the frequency band and the value of the equivalent active resistance. The main sources of thermal noise are circuits for impedance coherence in the optical transmitter and receiver and also the optical and RF amplifier. CCDN which has a resistance 75 Ω is also considered as a source of thermal noise, whose level is about 2 dBμV (T = 17 °C, B = 5 MHz) [21].

When the light sources convert the electrical signal into light stream different noise components are added to the useful signal. The most significant is the influence of the laser RIN, its phase noise and the refractive noise.

RIN has a quantum nature and is caused by the imperfection of the conversion process (electron – photon) in the laser material. In the active laser layer occur random emissions of a wide number of photons, causing random modulation og the generated light. RIN depends on: working temperature, laser exciting current and optical feedback, created by the reflection of the optical fibre [22, 23].

The phase noise has random nature and is caused by the spontaneous photon emission in the emitting layer of the laser. It leads to phase fluctuations of the output signal and extending of the spectral line, which for FP lasers ranges between 1 – 10 nm, while for DFB is less than 100 MHz [24].

In the optical modulator except the considered thermal noise is generated signal-spontaneous noise and spontaneous-spontaneous noise. The former depends on the total light power and the latter, which is dominating – on the modulator gain.

Besides the noises, generated in the exciting laser, the optical amplifier creates also noises by signal-spontaneous beating, spontaneous-spontaneous beating, multipath interference, double-Rayleigh backscatter and amplified spontaneous emission (ASE) [25]. ASE is due to the natural redistribution in the different power levels and generated shot noise in the receiver. The level of the caused by ASE noise is much lower than that of the generated by the mean optical power of the signal noise and for this reason could be neglected [26].

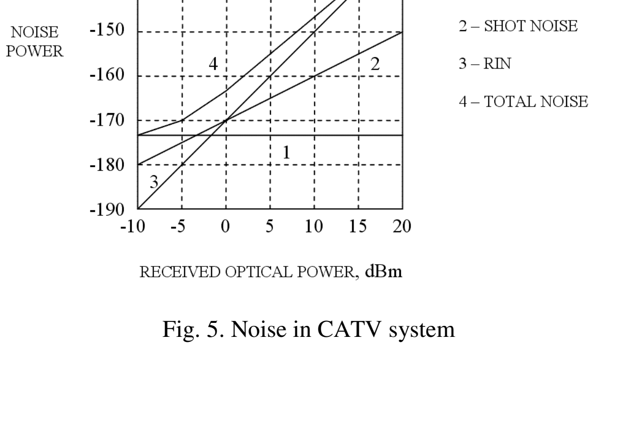


Fig. 5. Noise in CATV system

Defining for the optical receiver are the shot and thermal noise. The shot noise has quantum nature and is generated in the receiver photodiode. The reason for its occurrence is the imperfect conversion “photon-electron”.

The polarized noise has a random nature and is related to the gain dependence on the polarization and the polarized mode dispersion of the fibre. It is characterized with unequal attenuation of each mode in the output signal and leads to serious losses as a result of the polarization and therefore to reducing the CNR at the input of the optical receiver.

The noises created in the optical part of the system are given on figure 5. When the level of the received optical power is low, then the thermal noise of the photodiode receiver dominates, while in case of high level – dominates the laser RIN [1].

VI. Nonlinear Distortions in CATV System

The nonlinear distortions are reason for creating new frequency products in the spectrum of the useful signal, as a result of which the reliability of the received information deteriorates. The main sources of the nonlinear distortions are the optical transmitter, the optical fibre and the RF amplifiers. In the firm documentation of the optical transmitters are given the optimal value of the parameters which guarantee the defined level of the nonlinear distortions. The nonlinear distortions in the optical fibre are due to few events which happen when the level of the transmitted signals are very inadmissible high. They could be divided into two groups – nonlinearity, related to the scattering (stimulated Brillouin scattering and stimulated Raman scattering) and nonlinearity related to the Kerr effect (four fave mixing, cross phase modulation, self-phase modulation) [27].

Stimulated Brillouin scattering (SBS) is a nonlinear event which happens when optical power, larger than a defined threshold value, is feed to the fibre. SBS threshold depends on the width of the laser spectrum line and its output power. SBS leads to scattering of large part of the optical power transmitted through the optical fibre and this is the reason for reducing SNR at the input of the optical transmitter. In general when using the sources with narrow spectral line and external modulation, the SBS threshold varies from 5 to 10 mW. If using lasers with direct modulation then this power varies between 20 – 30 mW [28, 29].

The event stimulated Raman scattering (SRS) is similar to SBS. SRS threshold power is much bigger than SBS one and reaches values about 1 W. The final effect of SRS is transferring signal power from the channels with low wavelength to channels with large wavelength. That power transfer in some cases has a positive effect, for example Raman amplifier [28, 29].

Self-phase modulation (SPM) is an event that happens when digital signals are transmitted through the optical fibre. It is a result of changing the reflect coefficient of the optical fibre depending on the power feed to it. SPM leads to parasite phase modulation of pulse fronts (chirp). The pulse widen, interference between the symbols increases and limits the transmission speed. SPM is one of the reasons for decreasing the step between each channel [29, 30].

Cross-phase modulation (XPM) mainly happens in systems with WDM and DWDM. It is similar to the SPM and dues to the signal interaction from two neighbour optical channels. When the signals transmit through one optical fibre each of them changes its reflect coefficient keeping the law of changing of the optical power. Investigations show that the effective fibre surface is necessary to be increased in order to decrease the XPM [29, 30].

In WDM and DWDM systems is also typical the event four wave mixing (FWM). It is the reason for creating products with new frequencies, some of which may fall into the used channels. There are two main factors which affect to the level of FWM products, therefore the mixing efficiency – channel spacing and the dispersion of the chosen optical fibre [1, 29].

As mentioned above, the main source of nonlinear distortions in the coaxial part of the CATV System are the RF amplifiers. To escape these distortions it is necessary the maximal output level of the RF signal not to exceed the defined by half-line “B” level on the diagram on figure 4 [20].

VII. Conclusion

In this article it is performed an architecture of a CATV System, built on a hierarchal principle. There is given criteria for determining the dynamic range of the signals in the optical as well as the coaxial part of the system. The reasons for creating noise and nonlinear distortions in the main elements of the system – laser, photodiode receiver, optical fibre, optical and RF amplifiers are also analyzed.

References

[1] Д. Добрев. *Синтез, анализ и проектиране на широколентови кабелни радиокомуникационни системи – Хабилитационен труд*, ТУ-София, 2005.

[2] R.Watson. *Receiver Dynamic Range*, WJ Communications, Inc., 2001.

[3] J. Learned. *Hybrid Fiber/Coax options explode for last mile,* Communication System Design, August 2002, www.CommsDesign.com

[4] R. Grébol. *Analysis of the optical spectrum in the communications by means of optical fibre*, Telecommunications Engineer, Promax Electronica, S.A. Telecom Lab.

[5] http://www.cabledatacomnews.com/cmic/headend.html

[6] http://www.cabledatacomnews.com/cmic/hub.html

[7] *BXX Optical node*, Teleste access systems, January, 2005.

[8] P. Jr. Green. *Fiber to the home – The new empowerment,* John Wiley & Sons, Inc., 2006.

[9] http://iec.org, *Fiber to the home*, The International Engineering Consortium.

[10] J. Powers. *An introduction to fiber optic communication system*, The McGraw-Hill, 1986.

[11] K. Yelen. *Design method for ultimate efficiency in linear-cavity continuous-wave laser using distributed-feedback*, PhD Thesis, University of Southampton, 2004.

[12] J. Rajeev, Ch. Cox III, E. Goobar, M. Abraham, H. Lee, S. Patterson, F. Rana. *Semiconductor Lasers: Device Physics and Applications*, Part I, RLE Progress Report Number 141.

[13] G. Agrawal. *Fiber-optic communications systems*, John Wiley & Sons, Inc., 2002.

[14] E. Sackinger. *Broadband circuits for optical fiber communication*, John Wiley & Sons, Inc., 2005.

[15] Ph. Gallion,V. Dallot, E. Bridoux. *Study of noise properties in optical distributed Raman amplifiers using a semiclassical model*, Ecole Nationale Superieure des Telecommunications.

[16] M. Bass, E. Stryland. *Fiber optic handbook – fiber, devices and system for optical communication*, The McGraw-Hill, 2002.

[17] P. Hudec. *Noise in CATV Networks*, Applied Microwave & Wireless, September/October, 1997.

[18] http://www.deetc.isel.ipl.pt/sistemastele/ST1/arquivo/Amplif_nLin.pdf

[19] Д. Добрев, Л. Йорданова. *Радиокомуникации – 2 части*, С., Сиела, 2000.

[20] Д. Добрев, Л. Йорданова. *Приемане на радио- и телевизионни програми чрез спътници и по кабел*, С., Електрон-инвест, 1996.

[21] J. Randa. *Thermal noise measurements,* NIST/ARFTG Short Course, Boulder, December 2003.

[22] F. Jeremie, Jean-Luc Vey, Ph. Gallion. *Optical corpuscular theory of semiconductor laser intensity noise and intensity squeezed-light generation*, J. Opt. Soc. Am. B Vol. 14, No. 2, 1997.

[23] http://www.physics.montana.edu/students/meng/physics/noise/RIN.htm

[24] http://www.ece.drexel.edu/courses/ECE-T603/laser.pdf

[25] D. Baney, Ph. Gallion, R. Tucker. *Theory and Measurement Techniques for the Noise Figure of Optical Amplifiers*, Optical Fiber Technology 6, pp. 122-154, 2000.

[26] Ph. Gallion. *Classical Phase-Amplitude Description of Optical Amplifier Noise*, Ecole Nationale Supérieure des Télécommunications, Télécom Paris.

[27] Р. Фриман. *Волоконно – оптические системы связи,* М., Телеком транспорт, 2003.

[28] Z. Yusoff. *Applications of highly nonlinear holey fibres in optical communications*, PhD Thesis, University of Southampton, 2004.

[29] G. Agrawal. *Nonlinear fiber optics - Third Edition*, Academic Press, 2001.

[30] G. Agrawal. *Applications of nonlinear fiber optics*, Academic Press, 2001.

An Investigation of Noise Influences in Optical Transmitters and Receivers in Cable TV Networks

Krasen K. Angelov[1], Kiril R. Koitchev[2], Stanimir M. Sadinov[3]

Abstract – **Optical transmitters and receivers are key elements of optical equipment in optical communication networks Taking into account all specific requirements of CATV networks it is necessary for the optical transmitters and receivers to meet the required specifications for downstream and upstream channel transmission. Basically, the major negative impact in them is caused by the occurrence of noises. This paper analyzes the cause of their origin and evaluates their consequent influence.**

Keywords **– Distributed-Feedback Laser, Relative Intensity Noise, Shot Noise, Thermal Noise, Carrier-to-Noise Ratio.**

I. INTRODUCTION

The subject of noise introduced by optical transmitters and receivers has an important implication for cable TV transmission.

The fundamental noise components in optical transmitters are:

- the relative intensity noise (*RIN*);
- the laser phase noise.

Noise components which are fundamental for receivers are:

- the shot noise;
- the thermal noise;
- laser *RIN* noise.

The performance of these transmitters and receivers is expressed by way of carrier-to-noise ratio (*CNR*).

II. NOISE SOURCES IN OPTICAL TRANSMITTERS

A. RIN of laser transmitter

The output power of laser fluctuates around its steady-state value due to quantum fluctuations in the electron density as well as spontaneous emission events that are converted to intensity noise. The laser relative intensity noise (*RIN*) can be defined as [1]

$$RIN = \frac{\left\langle \left| \delta S(\omega) \right|^2 \right\rangle}{S^2}, \qquad (1)$$

where δS is the output power fluctuation from the average power value S. Then the laser *RIN* can be written as

$$RIN(\omega) = \frac{2R_{sp}\left[\Gamma_N^2 + \omega^2 + G_N^2 S^2\left(1 + \frac{\gamma_N .N}{R_{sp}.S}\right)\right]}{S\left[\left(\omega_R^2 - \omega^2\right)^2 + 2\Gamma_R^2\left(\omega_R^2 + \omega^2\right) + \Gamma_R^4\right]}, \qquad (2)$$

where $R_{sp} = \gamma_{sp} n_{sp} N / \tau_n$ is the spontaneous emission rate (γ_{sp} – the fraction of the spontaneous emission coupled into the cavity mode; n_{sp} – spontaneous emission factor; N – the carrier density; τ_n – the electron lifetimes; ω_R – the laser resonance frequency or relaxation oscillation frequency; $\Gamma_R \equiv (\Gamma_N + \Gamma_S)/2$ – the relaxation oscillation decay rate; $\Gamma_S = R_{sp}/S - G_S S$ – the photon decay rate; S – the photon density; $\Gamma_N = \gamma_N + N(\partial \gamma_N / \partial N) + G_N S$ – the small-signal decay rate; $G_N = \Gamma v_g (\partial g / \partial N)$; $G_S = \Gamma v_g (\partial g / \partial S)$; Γ – the carrier confinement factor in the active layer; v_g – the group velocity; g – the optical gain. At low frequency ($\omega < \omega_R$), the laser *RIN* is almost frequency independent, but it is significantly enhanced in the vicinity of $\omega = \omega_R$. At a given frequency, the *RIN* decreases with the bias current as $(I - I_{th})^{-3}$, where I_{th} is threshold current. As the bias current is increased, the *RIN* decreases more slowly as $1/(I - I_{th})$. The laser *RIN* imposes an upper limit on the maximum achievable *CNR* at the fiber node receiver. Consequently, the *RIN* of DFB laser transmitters, which are used for analog video transmission, are typically equal to *-155db/Hz* or better. It should be pointed out that the laser *RIN* can significantly degraded by multiple optical reflections [2].

B. Laser transmitter phase noise

It is well known that spontaneous emission events in the laser cavity change both the phase and amplitude of the optical field. Coupling of the spontaneous emission into the lasing modes as well as fluctuation in the electron density induce changes in both the real and imaginary parts of the refractive index, and produce phase noise. The spectral linewidth of a laser due to spontaneous emission can be written as [3]

$$\Delta v = \frac{v_g h v n_{sp}\left(\alpha_i + \alpha_m\right)\alpha_m\left(1 + \alpha^2\right)}{8\pi S}, \qquad (3)$$

where α_i and α_m are internal and mirror losses, respectively; α – the linewidth enhancement factor; n_{sp} – the spontaneous emission factor; h – the Plank constant; v – the photon energy. According to Eq. (3), the laser linewidth is inversely

[1]Kiril R. Koitchev is with the Faculty of Electrical Engineering and Electronics, Technical University – Gabrovo, 4 H. Dimitar St., 5300 Gabrovo, Bulgaria, E-mail: koitchev@tugab.bg

[2]Krasen K. Angelov is with the Faculty of Electrical Engineering and Electronics, Technical University – Gabrovo, 4 H. Dimitar St., 5300 Gabrovo, Bulgaria, E-mail: kkangelov@mail.bg

[3]Stanimir M. Sadinov is with the Faculty of Electrical Engineering and Electronics, Technical University – Gabrovo, 4 H. Dimitar St., 5300 Gabrovo, Bulgaria, E-mail: murry@tugab.bg

proportional to the output power and it is enhanced by the factor $(1+\alpha^2)$. The larger DFB laser linewidth is undesirable, particularly for relatively long-distance transmission (>*30km*).

C. Performance of DFB laser transmitters

Generally, the laser transmitter is designed to transport multiple AM/QAM video signals with the maximum possible *CNR* for a given optical received power at the fiber node. To achieve the maximum *CNR* for a given optical link budget, the output optical power as well as the optical modulation index *m* per channel must be as large as possible to overcome the laser *RIN*, fiber losses, and receiver noise. On the other hand, the maximum optical modulation index is limited to avoid unacceptably large nonlinear distortions and a clipping effect [4, 5]. The static clipping behavior increases the laser's nonlinear distortions and imposes an upper limit on the channel capacity. Consider a system with *N* channels, each with modulation index *m* and a total photocurrent at the receiver *I(t)*. For large *N*, the photocurrent *I(t)* can be modeled as a Gaussian random process with mean I_p and standard deviation $\sigma_p = mI_p(N/2)^{i/2}$. The normalized modulation index can be defined as:

$$\mu = \frac{\sigma_p}{I_p} = m\sqrt{\frac{N}{2}} . \qquad (4)$$

It can be shown that the carrier-to-nonlinear distortions (*C/NLD*) caused by clipping for small μ can be approximately given by [6]

$$C/NLD = \frac{1}{\Gamma}\sqrt{\frac{\pi}{2}}\mu^{-3}\left(1+6\mu^2\right)e^{1/2\mu^2} , \qquad (5)$$

where Γ ($\approx 1/2$) is the fraction of the distortion power within the cable TV band. The resultant *CNR* is the sum of the *CNR* and the *C/NLD*.

Digital channels such as QPSK or 16-QAM are transmitted in the return-path portion of the cable TV networks [7]. The return-path laser transmitters are not required to provide the same *CNR* or linearity as the downstream laser transmitters since they are not intended to transport AM video channels. The optical signals can be transmitted at either *1310nm* or *1550nm*. Notice that at room temperature, an error-free QPSK transmission can be achieved over a wide range of optical modulation indices, from about *1%* to *11%*. The operating range of optical modulation indices is reduced by 5dB at a higher temperature ($\approx 80^0C$) indicating an upper limit on an upstream channel capacity. The upstream impairments include a laser RIN, thermal noise, and optical reflections, as well as cumulative ingress noise from all the homes connected to a given fiber node.

III. NOISE SOURCES IN OPTICAL RECEIVERS

A. Shot noise

Shot noise in photodetector is a quantum noise, which is due to the random generation of electron-hole pairs when the photodetector is illuminated by photons. To derive the noise variance of the photocurrent generated in response to an optical signal with constant amplitude, it is make the following assumptions:

– the probability of generating a single electron-hole pair in a very small time interval Δt is proportional to Δt;

– the probability of generating more than a single electron-hole pair in Δt is negligible;

– the electron-hole pair generation events are statistically independent.

Based on these assumptions, the probability of generating exactly *n* electron-hole pairs per unit time can be given by

$$p(n) = \frac{N_0^2 e^{-N_0}}{n!} , \qquad (6)$$

where N_0 is the average number of received photons in the time interval Δt, which is equal to $P_{in}\Delta t / h\nu$ (P_{in} – the incident optical power). Let also assume that every photon generates an electron-hole pair at the receiver (*100%* quantum efficiency). Then, the average photocurrent is simply qN_0, where *q* is the electronic charge. The noise variance of the photocurrent at the receiver per unit frequency bandwidth is given by [8]:

$$\sigma_{shot}^2 = \left\langle \delta i_R^2 \right\rangle - \left\langle \delta i_R \right\rangle^2 = q^2\left[\left\langle n^2 \right\rangle - \left\langle n \right\rangle^2\right] = q^2 N_0 = qI_R . \quad (7)$$

Eq. (7) is also the spectral density of the shot noise, which is frequency independent. The single-sided spectral density becomes $2qI_R$. Under reverse-biased operation, the dark current (I_D), which is the residual photocurrent with no light due, also adds to the photodetector shot noise. Thus, the total photocurrent shot noise variance per unit frequency bandwidth will be:

$$\sigma_{shot}^2 = 2q\left(I_R + I_D\right). \qquad (8)$$

B. Thermal noise

The electrons move randomly in any conductor due to a finite temperature, which manifests itself as random fluctuations in the current even when no electrical voltage is applied. The random photocurrent fluctuations cause random voltage noise over a load-resistor terminal. The thermal noise is also called *Johnson* noise or *Nyquist* noise.

The double-sided spectral density expression is given by

$$S_p(\nu) = \frac{2k_B T}{R_L} , \qquad (9)$$

where R_L is the load resistor, k_B – the Boltzmann constant; *T* – the absolute temperature. The open-circuit single-sided spectral density of the photocurrent is given by:

$$S_p(\nu) = \frac{4k_B T}{R_L} = \left\langle \delta_{th}^2 \right\rangle . \qquad (10)$$

If an RF amplifier with a noise figure *F* is connected directly to the photodetector, then the photocurrent variance per unit frequency interval due to thermal noise is given by:

$$\sigma_{th}^2 = \left\langle \delta i_{th}^2 \right\rangle = \frac{4k_B TF}{R_L} . \qquad (11)$$

At room temperature with a *50Ω* load and a preamplifier with a noise figure of *3*, $\sigma_{th} = 31{,}5\, pA/\sqrt{Hz}$. In particular, at low reverse-bias voltages, the dark current increases by almost three orders of magnitude for a 60^0C temperature increase.

The thermal noise is can be expressed in terms of another useful parameter called noise-equivalent power (*NEP*), which is defined as the minimum optical power per unit of bandwidth that is required to produce *SNR = 1*. Therefore, the *NEP* can be written as:

$$NEP = \frac{hv\left[\left\langle \delta i_{th}^2 \right\rangle\right]^{1/2}}{q\eta} = \frac{hv}{q\eta}\left[\frac{4k_B TF}{R_L}\right]^{1/2}. \tag{12}$$

The *NEP* is useful to estimate the required optical power for a given *SNR* if the noise bandwidth *B* is known. Using the *NEP* definition, the noise-equivalent photocurrent N_R can also be defined as *NERP*, which has typical values in the range $1 \div 10\, pA/\sqrt{Hz}$.

C. Laser receiver RIN noise

The laser relative intensity noise (*RIN*) is due to the laser spontaneous emission and fluctuations in the electron density. The laser *RIN* is defined by

$$RIN = \frac{\left\langle \delta i_{RIN}^2 \right\rangle}{I_R^2}, \tag{13}$$

where $\left\langle \delta i_{RIN}^2 \right\rangle$ is the photocurrent spectral density due to the laser *RIN*. Thus, the photocurrent variance due to the laser *RIN* can be given by:

$$\sigma_{RIN}^2 = \left\langle \delta i_{RIN}^2 \right\rangle = I_R^2 RIN. \tag{14}$$

It should be pointed out that the laser *RIN* has a frequency dependence similar to the small-signal modulation response of a DFB laser. In an optical communication system, the laser *RIN* is replaced by the system *RIN*, which includes the contribution of the various system elements such as the fiber *RIN*, laser *RIN*, and EDFA *RIN*.

D. Performance of optical receivers

Let assume a communication system with modulation index *m* per channel with a DC photocurrent of I_R and an effective noise bandwidth *B* at the receiver. Then, using the *CNR* definition, the *CNR* can be written as

$$CNR = \frac{\left\langle i_R^2 \right\rangle}{\left(\sigma_{shot}^2 + \sigma_{th}^2 + \sigma_{RIN}^2\right)B}, \tag{15}$$

where $\left\langle i_R^2 \right\rangle = \left(mI_R\right)^2/2$ is the mean-square signal photo-current. Substituting Eqs. (7), (11), and (14) for the shot noise, thermal noise, and *RIN* noise, respectively, in Eq. (15) for the *CNR* at the photodetector, will finally obtain:

$$CNR = \frac{\left(mI_R\right)^2}{2B\left[I_R^2 RIN + 2q\left(I_R + I_D\right) + \dfrac{4k_B TF}{R_L}\right]}. \tag{16}$$

To gain further insight into Eq. (16), there should be analyzed the *CNR* in the following three cases: Fig. 1 illustrates the *CNR* behavior versus the received photocurrent according to Eq. (16) as well as the thermal noise, shot noise, and laser *RIN* contributions to the *CNR*. Let assumed that the photocurrent was operating at room temperature with a preamplifier with a noise figure of *3*, a *10kΩ* load resistor, laser *RIN = -155dBc/Hz*, a *4%* modulation index for the transmitted RF channel, a *4MHz* noise bandwidth, and the photodetector dark current was neglected. In most practical cases in witch the incident optical power is very small (*<-10dBm*), the thermal noise dominates over both the shot noise and the laser *RIN* in photodetector as shown in Fig. 1. Therefore, the *CNR* becomes:

$$CNR = \left(\frac{R_L (mR)^2}{8k_B TBF}\right)P_{in}^2. \tag{17}$$

Eq. (17) shows that the *CNR* increases as the square of the input optical power in the thermal noise limit. Furthermore, increasing the load resistor and reducing the noise figure of the amplifier can improve the *CNR*.

Another interesting limit is the shot noise limit, which is where the shot noise dominates over both the thermal noise and the laser *RIN*. In this limit, the *CNR* becomes:

$$CNR = \left(\frac{\eta m^2}{4hvB}\right)P_{in}. \tag{18}$$

Notice that the *CNR* increases linearly with the optical input power.

The third limit is the laser *RIN* limit, which is where the laser *RIN* dominates over both the thermal and the shot noise. In this limit, the *CNR* does not depend on the photocurrent in the receiver and can be improved by reducing the laser *RIN*. This limit plays an important role at high optical-input power levels (*>0dBm*), where the *CNR* at the receiver is upper limited by the laser *RIN*. Consequently, to maximize the *CNR* at the fiber-node receiver, directly modulated DFB laser transmitters with *RIN* of *-155dBc/Hz* or less are typically required.

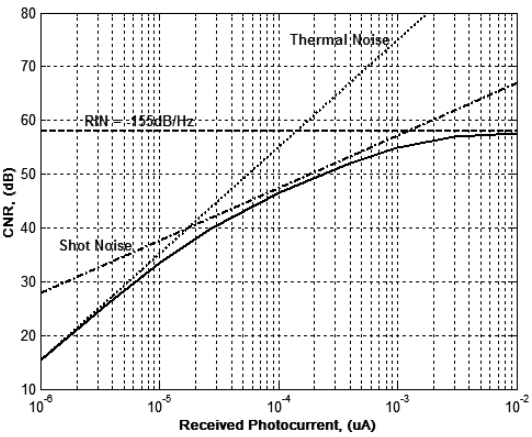


Fig. 1. Calculated *CNR* and its components due to shot noise, thermal noise, and laser *RIN* noise versus the received photocurrent

IV. CONCLUSION

The noise components in optical transmitters and receivers bring about to limitation of maximum achievable *CNR*. While for transmitters' performance it is very important to not allow the ingress of nonlinear distortions by transmission, it is necessary for the receivers to have optimal CNR at their input. The both maximum optical modulation index and the output optical power are limited to avoid unacceptably large nonlinear distortions and a clipping effect. Require performance of laser transmitters for the downstream and upstream channel is different due to the dissimilar nature oflarger temperature range than the downstream DFB laser transmitters. Optical receivers are impacted by a larger number of noise components. Methods of raising laser are typically installed inside the downstream optical receiver at the fiber node, they are required to operate over a much data transmitted along them. Since the return-path transmitters receivers performance vary depending on the prevailing noise component.

REFERENCES

[1] G. Agrawal, N. Dutta, *Long-Wavelength Semiconductor Lasers*, Vabn Nostrand Reinhold Company, New York, 1986.

[2] W. Way, "Multiple-Reflection-Induced Intensity Noise Studies in a Lightwave System for Multichannel AM-VSB Television Signal Distribution", *IEEE Photonics Technology Letters 2*, pp.360-362, 1990.

[3] C. Henry, "Theory of the Phase Noise and Power Spectrum of a Single Mode Injection Laser", *IEEE Journal of Quantum Electronics QE-19*, pp.1391-1397, 1983.

[4] C. Kou, "Fundamental Second-order Nonlinear Distortions in Analog AM CATV transport Systems Based on Single Frequency Semiconductor Lasers", *IEEE Journal of Lightwave Technology 10*, pp.235-243, 1992.

[5] K. Maeda, H. Nakata, K. Fujito, "Analysis of BER of 16QAM Signal in AM/16QAM Optical Transmission System", *IEE Electronics Letters 29*, pp.640-641, 1993.

[6] A. Saleh, "Fundamental Limit on Number of Channels in Subcarrier Multiplexed Lightwave CATV Systems", *IEE Electronics Letters 25*, pp.776-777, 1989.

[7] GR-2853-CORE, *Generic Requirements for AM/Digital Video Laser Transmitters and Receivers*, Issue 2, Bellcore, December 1995.

[8] G. Agrawal, *Fiber-Optic Communication Systems*, John Wiley & Sons, New York, 1992.

An Effectiveness Investigation of Erbium-Doped Fiber Amplifiers for Cable TV Networks in presence of Noise

Kiril R. Koitchev[1], Krasen K. Angelov[2], Stanimir M. Sadinov[3]

Abstract – **Recent growth of modern CATV networks is accompanied by the necessity of ensuring large areas of reception. This brings about to replacement of conventional coaxial cable trunks with optical rings. Despite its major advantages, optical equipment also has its peculiarities and shortcomings. This article makes observations of the impact of various noise components within Erbium-Doped Fiber Amplifiers (EDFAs) and will estimate the carrier-to-noise ratio (*CNR*) in this particular type of amplifiers in transmitting analogue amplitude modulated (AM) video signals.**

Keywords – **Erbium-Doped Fiber Amplifier, Amplifier Spontaneous Emission, Signal-Spontaneous Beat Noise, Spontaneous-Spontaneous Beat Noise, Noise Factor, Carrier-to-Noise Ratio.**

I. INTRODUCTION

Understanding noise generation in optical amplifiers and its negative impact on amplifier performance is a critical issue in optical communication networks. Optical amplifiers serve to amplify TV video signal which are distributed over the network. They are in-line amplifiers, power booster amplifiers (to the laser transmitter), or preamplifiers (to the receiver).

In order to determine the signal *CNR* in optical communication systems it is necessary to analyze noise components in optical amplifiers first. Such analysis would involve the following assumptions:

1) It is assumed a single-stage optical amplifier with gain so that the output power is related to the input power by $P_{out} = GP_{in}$, where P_{out} is the output power, and P_{in} – the input power;
2) It is assumed that there is a 100% coupling efficiency between the amplifier and photodetector;
3) The optical communication network is based on single mode fibers (SMF).

II. OPTICAL FIBER AMPLIFIER NOISE

The photocurrent at the receiver due to the amplifier spontaneous emission (ASE), according to [1], can be expressed as

$$I_{ASE} = RP_{ASE} = 2Rn_{sp}h\nu(G-1)\Delta\nu \,, \tag{1}$$

where P_{ASE} is the amplified spontaneous emission powering the amplifier; R – the load resistor of photodetector; n_{sp} – the spontaneous emission factor; h – the Plank constant; ν – the photon energy; $\Delta\nu$ – the amplifier optical bandwidth.

From Eq. (1), the current variance in the photodetector due to the ASE shot noise can be written as:

$$\sigma^2_{ASE} = \left\langle i^2_{ASE} \right\rangle = 2qI_{ASE} = 4q^2\eta n_{sp}(G-1)\Delta\nu \,. \tag{2}$$

where q is the electronic charge; η – the quantum efficiency of the photodetector.

There is also a shot-noise component due to the photocurrent I_R. Let define a related quantity to the ASE power called the single-sided spectral density of the ASE, which is nearly constant and can be written as:

$$S_{sp(\nu)} = \frac{P_{ASE}}{2\Delta\nu} = n_{sp}(G-1)h\nu \,. \tag{3}$$

The effect of the ASE is to add noise fluctuations to the amplified power, which are converted to current fluctuations during the photodetection process. The primary contributions to the optical receiver noise come from the beating of the spontaneous emission with the signal, which is called a signal-spontaneous beat noise (S-ASE beat noise), and with itself, which is called a spontaneous-spontaneous beat noise (ASE-ASE beat noise) [2].

Using Eq. (3), the variance of the photocurrent due to the S-ASE beat noise can be given by

$$\sigma^2_{S-ASE} = \left\langle i^2_{S-ASE} \right\rangle = 4I_R(RS_{ASE}) = 4\frac{(q\eta)^2}{h\nu}n_{sp}G(G-1)P_{in} \,, \tag{4}$$

where a factor of 4 in Eq. (4) is used because needs to consider a double-sided ASE spectral density, and to include both light polarizations. Notice that the S-ASE beat noise is independent of the amplifier optical bandwidth ($\Delta\nu$). Consequently, the negative impact of ASE cannot be removed without eliminating the desired signal at the same time. In contrast, the ASE-ASE beat noise is proportional to the amplifier bandwidth. This means that a narrow bandpass optical filter can be used to remove this ASE beat noise. The variance of the ASE-ASE beat noise can be given by

$$\sigma^2_{ASE-ASE} = \left\langle i^2_{ASE-ASE} \right\rangle = 4\left(RS_{ASE}\right)^2\Delta\nu = 4\left[q\eta n_{sp}(G-1)\right]^2\Delta\nu \,, \tag{5}$$

where the factor of 4 in Eq. (5) is used for the same considerations as in Eq. (4). When the amplifier gain G is high (i.e., small input power levels), the current variance of both the S-ASE and ASE-ASE beat noise is proportional to G^2, while the ASE shot noise is only proportional to G. Consequently, the ASE shot noise contributions can be neglected

[1]Kiril R. Koitchev is with the Faculty of Electrical Engineering and Electronics, Technical University – Gabrovo, 4 H. Dimitar St., 5300 Gabrovo, Bulgaria, E-mail: koitchev@tugab.bg

[2]Krasen K. Angelov is with the Faculty of Electrical Engineering and Electronics, Technical University – Gabrovo, 4 H. Dimitar St., 5300 Gabrovo, Bulgaria, E-mail: kkangelov@mail.bg

[3]Stanimir M. Sadinov is with the Faculty of Electrical Engineering and Electronics, Technical University – Gabrovo, 4 H. Dimitar St., 5300 Gabrovo, Bulgaria, E-mail: murry@tugab.bg

compared with the ASE-ASE and S-ASE beat noise contributions in the photodetector. This approximation will also be used in the calculation of the *CNR* in the next section.

III. ESTIMATION OF *CNR* AND NOISE FIGURE OF OPTICAL FIBER AMPLIFIER

Here will be made analysis of the *CNR* of fiber-optics communication system, particularly for cable TV applications. It is assumed a communication system with an in-line single-stage optical amplifier, with modulation index m per RF channel, and DC photocurrent of I_R. Then, using the *CNR* definition in [3], the *CNR* can be written as

$$CNR = \frac{\langle i_R^2 \rangle}{\left(\sigma_{shot}^2 + \sigma_{th}^2 + \sigma_{S-ASE}^2 + \sigma_{ASE-ASE}^2\right)B}, \qquad (6)$$

where $\langle i_R^2 \rangle = \left(mI_R^2\right)/2$ is the mean-square signal photocurrent, σ_{shot}^2, σ_{th}^2 – the variance of the photocurrent due to the shot noise and thermal noise; B – the photodetector electrical bandwidth. The photodetector shot-noise variance can be given by the following equation:

$$\sigma_{shot}^2 = \langle i_{shot}^2 \rangle = 2q\left(I_R + I_{ASE}\right) = 2q^2\eta\left[\frac{GP_{in}}{h\nu} + 2n_{sp}(G-1)\Delta\nu\right]. \qquad (7)$$

From a practical aspect, it is convenient to express the *CNR* given by Eq. (6) in terms of the amplifier noise figure. The amplifier noise figure is defined as [4]

$$F = \frac{SNR_{in}}{SNR_{out}} = \frac{CNR_{in}}{CNR_{out}}, \qquad (8)$$

where "*in*" and "*out*" refer to the *SNR* or the *CNR* measured at the output of the photodetector with and without the optical amplifier, respectively). Without the optical amplifier, the photodetector is essentially shot-noise limited. Consequently, the *CNR* can be written as:

$$CNR_{in} = \frac{\left(mI_R\right)^2}{4qI_R} = \frac{m^2\eta P_{in}}{4h\nu}. \qquad (9)$$

In the presence of the optical amplifier, the photodetector is primarily limited by the shot noise and the S-ASE beat noise. The ASE-ASE beat noise and the ASE shot noise contributions can be neglected compared with the S-ASE beat noise contribution. The contribution to the inverse *CNR* caused by the S-ASE beat noise can be written as:

$$CNR_{S-ASE}^{-1} = \frac{8h\nu n_{sp}}{m^2 P_{in}}\left(1 - \frac{1}{G}\right). \qquad (10)$$

The resultant CNR^{-1} at the photodetector with the optical amplifier can be calculated according to:

$$CNR_{OUT}^{-1} = CNR_{Shot}^{-1} + CNR_{S-ASE}^{-1} = \frac{4h\nu}{m^2 P_{in} G}\left[1 + 2n_{sp}(G-1)\right]. \qquad (11)$$

Substituting Eqs. (9) and (11) into Eq. (8):

$$F = 2n_{sp}\left(1 - \frac{1}{G}\right) + \frac{1}{G}. \qquad (12)$$

For a high-gain operation ($G >> 1$), the amplifier noise figure can be approximated by $2n_{sp}$. If the further assume an ideal optical amplifier with a complete population inversion, then $n_{sp} = 1$. This means that in the quantum noise limit, the amplifier has a noise figure of $F = 3dB$ (since $F = 10\log_{10}2$). Neglecting the right-hand side term $1/G$ from Eq. (12), the optical amplifier noise figure can be written as:

$$F = 2n_{sp}\left(1 - \frac{1}{G}\right). \qquad (13)$$

Substituting the noise figure expression given by Eq. (13) in Eq. (6) for the *CNR*, we will finally obtain

$$CNR = \frac{m^2}{2B\left[RIN + \frac{2q}{I_R} + \left(\frac{N_R}{I_R}\right)^2 + 2h\nu + \Delta\nu\left(\frac{Fh\nu}{P_{in}}\right)^2\right]}, \qquad (14)$$

where *RIN* is the overall communication system *RIN* and N_R is the photodectector noise equivalent current.

Drawing from Eq. (14) and assuming the exemplary operating system parameters given in Table I, it is possible to determine the dependency $F = f(P_{in})$ – Fig. 1 or the dependency $CNR_{AM} = f(P_{in})$ – Fig. 2.

TABLE I
SYSTEM PARAMETERS FOR *CNR* CALCULATION

Parameter	**Specification**
Modulation index per channel	*3,2%*
Electrical noise bandwidth of AM channel	*4MHz*
Optical bandwidth of the EDFA	*40nm*
Receiver Photocurrent	*0,85mA*
Receiver noise equivalent current	$7\ pA/\sqrt{Hz}$
RIN	*-164dBc/Hz*

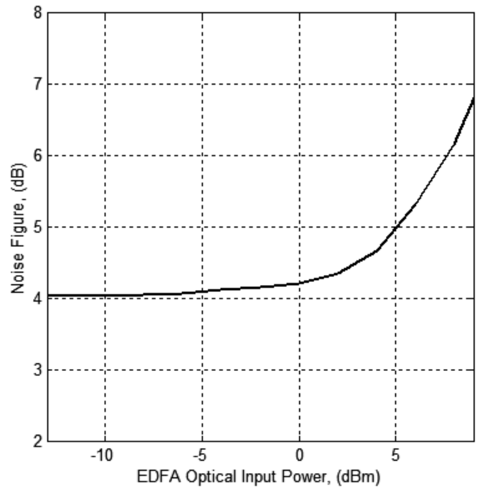


Fig. 1. Calculated noise figure F versus the EDFA optical input power level P_{in}

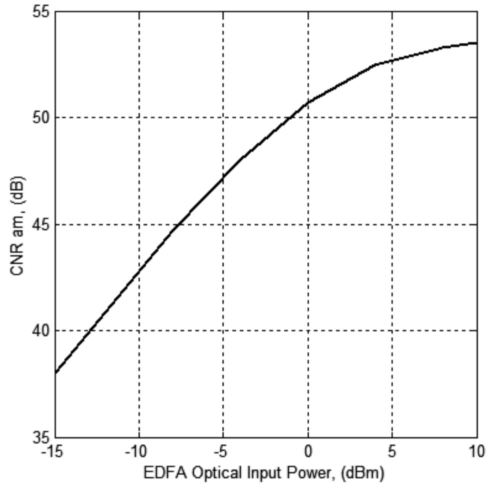


Fig. 2. Calculated AM CNR versus the EDFA optical input power level P_{in}

IV. CONCLUSION

The obtained Eq. (14) is very convenient to use since it describes the *CNR* dependence in terms of two primary measurable parameters, namely, the amplifier noise figure (*F*) and its optical input power. Furthermore, Eq. (14) suggest that the transmitted *CNR* of the RF channel after the in-line EDFA is primarily governed by the S-ASE and ASE-ASE beat noise at low optical input levels (< *-10dBm*). In this input power regime, the amplifier noise figure approaches the quantum noise limit of *3dB*). If the noise figure is nearly unchanged when the EDFA is operating in saturation (optical input power ≥ *0dBm*), the AM *CNR* becomes limited by the receiver's thermal and shot noise at a given detected optical power). However, if the amplifier noise figure increases monotonically with the optical input power, then it sets the upper limit on the AM *CNR* as seen from Eq. (14) and Fig. 2. Therefore, robust transmission of AM video channels often requires the EDFAs to operate in saturation. The result in Fig. 1 suggests that an in-line amplifier with a noise figure of *5dB* or less is needed to achieve *CNR* greater or equal to 50dB (according to the requirements for video signals transmission).

REFERENCES

[1] C. Giles, E. Desurvire, "Propagation of Signal and noise in Concatenated Erbium-Doped Fiber Optical Amplifiers", *IEEE Journal of Lightwave Technology 9*, pp.147-154, 1991.

[2] C. Kou, "Fundamental Second-order Nonlinear Distortions in Analog AM CATV transport Systems Based on Single Frequency Semiconductor Lasers", *IEEE Journal of Lightwave Technology 10*, pp.235-243, 1992.

[3] T. Darcie, "Subcarrier Multiplexing for Lightwave Networks and Video Distribution Systems", *IEEE Journal on Selected Areas in Communications 8*, pp.1240-1248, 1990.

[4] G. Agrawal, *Fiber-Optic Communication Systems*, John Wiley & Sons, New York, 1992.

Architecture for Integrating UMTS and 802.11 WLAN Networks

Toni Janevski, Zaklina Bogoeska

Abstract – **Cellular networks, e.g. UMTS, provide voice and data services to mobile users. In hotspots where users need high speed data services operators can deploy low-cost high speed WLANs, e.g. 802.11, to cover hot-spots. This paper proposes a possible architecture of integrating UMTS and 802.11 WLAN. The architecture allows a mobile node to maintain data (PS) connection through WLAN and voice (CS) connection through UMTS in parallel.**

Keywords – **Architecture, UMTS, Wireless LAN.**

I. INTRODUCTION

The 3G cellular networks, e.g. UMTS [1], support real-time and non-real-time multimedia services. The sustainable data rate per user in reality is up to several hundreds of kbps. On the other hand, clusters of high-speed usage of the mobile network can be found in certain areas, such as Internet cafés, office buildings, apartment buildings etc.

These clusters of high-speed usage areas are called *hot spots*. Fortunately, these areas are scattered within a wireless operator's domain. The operators would like to deploy lowcost high-speed solution to cover the hot spots that is either an extension of UMTS or inter-workable with UMTS so that they can maximally utilize the already deployed infrastructure. Wireless LANs, e.g. 802.11[2], which provide affordable services offers a viable and attractive choice as being high-speed (up to 54Mbps) and low-cost.

Both UMTS and WLANs are already commercially available and become increasingly popular. The number of mobile users is growing rapidly and so does the demand for wireless access to the Internet services, imposing the need for a unified QoS support framework in both UMTS and WLAN.

In particular, the users in the UMTS-WLAN interworking scenario require support for vertical handover (handover between heterogeneous technologies) between WLAN and UMTS. Existing solutions to vertical handover supporting the integration of UMTS and WLAN include network and application layer techniques based on Mobile IP (MIP) [3] and Session Initiation Protocol (SIP) [4], respectively.

In this paper we refer to the architectures for interworking of UMTS and WLAN. The paper is organized as follows. In the next section we give some background for the analysis. Possible integration architectures for UMTS-WLAN are discussed in Section 4. Finally, conclusions are given in Section 5.

[1]Toni Janevski is Associate Professor at the Faculty of Electrical Engineering, Karpos 2bb, Skopje, Macedonia. E-mail:tonij@etf.ukim.edu.mk

[2]Zaklina Bogoeska is graduate student at the Faculty of Electrical Engineering, Karpos 2 bb, Skopje, Macedonia.

II. BACKGROUND

A. UMTS

UMTS provides packet data (PS) service for data applications and circuit-switched (CS) service for telephony voice applications. In UMTS, the network of RNCs and Node Bs constitute the radio access network (RAN), called *UTRAN*.

The network of one RNC and its Node Bs is called *RNS*. Each Node B constitutes a cluster of base stations and a group of Node Bs is connected to a single RNC. The packet core network (CN) is comprised of SGSN and GGSN. In RAN, the RNC receives downlink packets from the SGSN and converts them into radio frames before sending them to Node Bs. On the reverse path the RNC receives radio frames from Node Bs and converts them into IP packets before sending them to the SGSN. The RNC manages the radio resources of Node Bs and sets up Radio Access Bearers (RABs) through them. The core network is connected to the Internet through GGSN. The SGSN manages mobility states of mobile nodes, establishes the data sessions, and controls the RAB set-up through RNCs. The IP packets are transported through GTP tunnels between GGSN and SGSN, and between SGSN and RNCs.

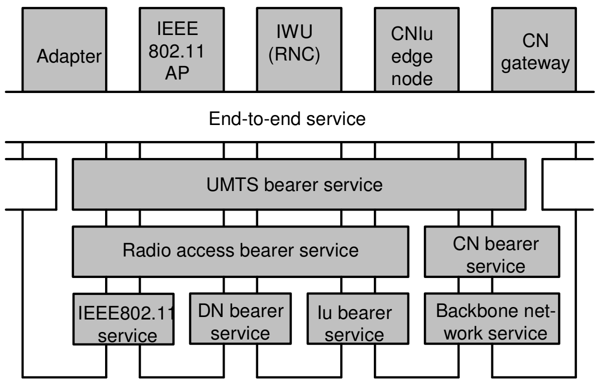


Fig.1. Adapted UMTS bearer concept using IEEE 802.11 bearer

Four different QoS classes are defined for the [5]: conversational class, streaming class, interactive class, and background class. The first two are targeted to real-time services, such as voice-over-IP and video/audio streaming, and the last two are targeted to non-real-time services, such as web-browsing (interactive) or email (background), etc.

B. IEEE 802.11 W LAN

The IEEE standard 802.11 operates in ad-hoc and infrastructure modes. In infrastructure mode, an Access Point (AP) coordinates the transmission among nodes within its radio coverage area, called service set. Only the infrastructure mode is relevant to WLAN integration with cellular network.

A Mobile Node (MN) can only associate with one AP at a time. Roaming across APs is supported in layer-2 through Inter-AP Protocol (IAPP). The APs generate beacons periodically that contain the network id (or Extended Service Set Identifier, ESSID) and cell-id (which is the AP's MAC address) in addition to other information. When the MN moves to a new cell where it receives a beacon with the same network-id but a new cell-id, it associates with the new AP by sending re-associate request frame that includes the MAC address of the old AP. The new AP can communicate with the old AP through IAPP [6] to obtain the context information.

QoS control is required to handle QoS guaranteed services over the WLAN component of the integrated network. In this section, a policy-based QoS architecture is proposed for WLAN as a prelude to the QoS architecture for an integrated UMTS–WLAN system. Typically, a WLAN network comprises a number of WLAN access points (APs) connected to a WLAN router (WR) that provides access to external networks. QoS mechanisms at the network layer in the form of IP packet header marking, policing and conditioning in the differentiated services (DiffServ) architecture, and at the data link layer in the form of WLAN QoS mechanisms (802.11e), media access control (MAC) frame priority indication (802.1p), and virtual LAN (VLAN) tagging (802.1q) are controllable by the QoS architecture. The MAC QoS mechanisms are provided by the WLAN APs, and the IP DiffServ mechanisms are available at the WR. In the following WLAN policy-based QoS architecture, only the DiffServ mechanisms in the WR are available to the policy control process.

There are two approaches [7] proposed for coupling WLAN networks with UMTS: tightly coupled architecture and loosely coupled architecture (Fig. 2).

In a tightly coupled architecture, the WLAN network is connected to the UMTS network as an alternative radio access network. In other words, the WR is connected directly to the serving General Packet Radio Service (GPRS) support node (SGSN) and is treated by the SGSN as a radio network controller (RNC). The data sent by WLAN devices must go through the UMTS PS domain served by the connecting SGSN to reach its destination. To ensure seamless IP QoS services in tightly coupled WLAN–UMTS networks, the UMTS session control entities such as call state control functions (CSCFs) in UMTS IMS are extended into the WLAN network. The session control entities like the CSCFs interact directly with the WLAN devices as if they are normal UMTS user equipment (UE). Thus, a PDF can enforce the network-level policies at the WR directly as if the WLAN network is a part of the UMTS PS domain. In a tightly coupled architecture, the WLAN is an alternative radio access network, so the 3GPP PDF is reused.

In a loosely coupled architecture, the WLAN is connected to a gateway GPRS support node (GGSN) of the UMTS network as a separate network. The WR is treated like a GGSN, and the WLAN network is considered a peer UMTS network. This article proposes that the WLAN constitutes a distinct policy domain with its own PDF, called a WLAN PDF (WPDF). The WPDF acts as the PDP in the WLAN domain, and the WR is the PEP that enforces the policy decisions made by the WPDF. In a loosely coupled architect-ture, the WLAN domain can use session control entities (CSCFs) in the UMTS network.

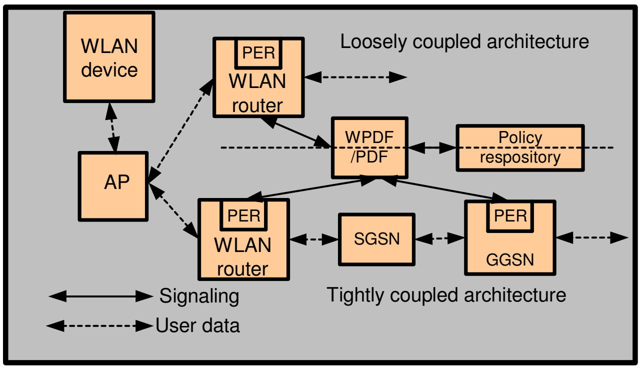


Fig. 2. Policy-based QoS architecture in WLAN

Alternatively, a session control entity can be sited in the WLAN domain to interact with the CSCFs in the UMTS network. This WLAN session control entity is related to the WPDF in the same way as the relationship between the proxy (P)-CSCF and the PDF in UMTS IMS. By adopting an additional session control entity and PDF in the WLAN domain, the distinct interface between the WLAN domain and its interconnected UMTS network is preserved. How the policies implemented in the WLAN domain are related to the policies in the UMTS network is dependent on the interworking scenarios. Consequently, the interaction method between the WPDF and the PDF of the UMTS network is determined by this policy relationship.

The loosely coupled WLAN architecture offers a major advantage over the tightly coupled architecture: integration flexibility offered by the distinct WLAN policy domain. It permits easy integration of the WLAN domain into a multi-operator multidomain environment. This flexibility permits simple extension of the QoS policy control architecture in [8] into the WLAN domain regardless of the UMTS–WLAN integration scenarios.

III. Integration Architecture

Defining an architecture that integrates the UMTS and 802.11 networks would face the following challenges.

1. What is the impact of differences in QoS models of the two radio access networks on the types of applications users can run and consequently on the traffic handled in each network? The UMTS RNS supports QoS for four well defined service classes – interactive, voice, stream, and best effort. In contrast, the QoS support for 802.11 is still under discussion.

2. How to deal with different connection paradigms used in each network? The GPRS is connection oriented, whereas 802.11 is a connection-less wireless LAN network.

3. How to ensure packet routing across the two networks when different mobility management schemes are employed in each network? In GPRS packets are routed through tunnels established between GGSN-SGSN and SGSN-RNC. In contrast, a number of mobility solutions are proposed for the IP network, which can be used for802.11 WLAN (e.g. HMIPv6 [9], CIP [10]).

4. How to select the best integration point when multiple integration points exist each with different cost-performance benefit for different scenarios? For example, the WLAN can be connected to either RNC, or SGSN, or GGSN.

Three scenarios are considered in the following subsections to illustrate the feasibility of the proposed architecture:

- One operator controls the UMTS network and WLANs.
- Different UMTS operators share a WLAN.
- An independent WLAN is interconnected to UMTS.

A. Scenario 1: UMTS and WLAN under one operator

Scenario 1 is a PLMN model where the operator installs and operates an integrated UMTS–WLAN network. The operator fully controls its WLAN sites. For the integrated UMTS–WLAN environment in a single operator's network, the hierarchical policy architecture is used (Fig. 3).

The master policy controller (MPDF) connects to the policy controller of the WLAN network (WPDF) and the UMTS policy controller (PDF). The MPDF of the UMTS network serves as the master policy node of the WPDF so that the policies implemented in the WLAN domain are integrated into the operator's policy hierarchy. The MPDF translates the network-wide policies into domain-specific net- work-level policies on behalf of the WPDF and PDF, and stores them in the policy repository.

The PDFs just retrieve these network-level policies from the repository, translate them into device-level policies, and install these policies in the network devices under their control. Note that these policies are enforced on all new IP multimedia sessions uniformly unless there are policy conflicts regarding the authorization of a new session's QoS requirements. In this case, the policy conflicts must be resolved through the MPDF. The communications protocol between the MPDF and the PDFs is based on the Common Open Policy Services (COPS) [11] protocol.

Note that a two-level hierarchy is shown in Fig. 3 for illustrative purposes only. The network operator may decide to provide intermediate levels of policy conflict resolution for more granular control over its network.

B. Scenario 2: A WLAN network shared by multiple operators

The normal range of a WLAN AP is less than 150 m in open space and just 50 m in closed environments. To provide full WLAN coverage over a built-up area like a city requires deployment of a large number of WLAN APs.

As a result, the cost of providing extensive WLAN services is high for a single operator, and it is desirable for operators to share their WLAN infrastructure to reduce upfront costs. Scenario 2 is such a risk-sharing PLMN model to provide different services in an integrated UMTS–WLAN environment. Multiple operators install and operate shared WLAN sites at different locations that are connected to their own UMTS networks. In this scenario, operator A may usually target business subscribers while operator B targets youth subscribers. The shared WLAN sites are configured to provide different amounts of resources at different times of the day. For example, operator A may be allocated more resources to provide business services during daytime, while operator B may get more resources to provide entertainment services after work hours. To provide end-to-end QoS in this environment, a WPDF is deployed in the shared WLAN domain. Scenario 2, shown in Fig. 4, illustrates how the WPDF interacts with the MPDFs of the cooperating operators' networks.

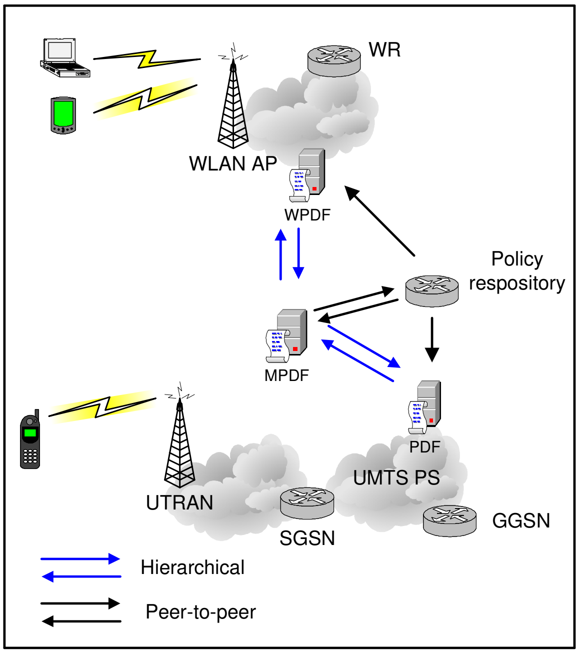


Fig. 3. WLAN and UMTS networks controlled by one operator

In scenario 2, the QoS policies to be applied in the WLAN domain are subjected to the control of operators A and B. For WLAN traffic going into operator A's network, the WPDF applies the policies supplied by operator A's MPDF. Likewise, WLAN traffic going into operator B's network is subjected to the policies supplied by operator B's MPDF. These policies specify how much resources in the WLAN access router (WR) should be provided to the traffic to be carried by the different operators' networks in order to satisfy the QoS requirements contracted by the operators. In terms of policy relationship, the WPDF is a child node in the policy hierarchies of operators A's and B's networks, and it is serving two MPDFs that are peers.

C. Scenario 3: Customer's WLAN network interconnected to operator's UMTS network

This scenario is shown in Fig. 5. Here, the WLAN may belong to an independent Internet service provider (ISP) or an enterprise that is a customer of the UMTS operator. This model allows the UMTS operator to provide wide-area mobile services to customers that have their own WLAN infrastructure.

In contrast to scenario 2, the WPDF in the WLAN domain is a peer of the MPDF in the UMTS network. The WPDF has the sole right to update its policy repository. The network-level policies to be employed by interconnecting the UMTS network and the WLAN network are determined by the service level specifications (SLSs) agreed between the peering WLAN and UMTS operators.

In these SLSs, there are static and dynamic service requirements. The static service requirements can be directly translated into enforceable network-level policies to be retrieved by the WLAN's PDF and the UMTS' PDFs in the respective networks. The purpose of SLS negotiation is to enable the interconnected networks' interdomain policy agents (IPAs) to agree on the specific service requirements that must be supported under the prevailing network states. Once the SLS negotiation is successfully completed, the participating IPAs can translate the agreed on service requirements into enforceable policies in their respective networks. Note that this runtime negotiation may not be initiated on a per-session basis. Instead, SLS negotiation is usually initiated when the IPA detects that the state of its network has changed and the existing policies are no longer enforceable.

IV. CONCLUSIONS

In this paper we have justified the use of the loose coupling approach to UMTS-WLAN integration in most scenarios. Tight coupled architectures is more suitable to situations where the cellular operator owns the WLAN and can reap the benefits of using the already in place infrastructure for billing and authentication. Despite originally being seen as competitors in the telecommunications market it has now become an industry goal on integrating the two different technologies. This integration can lead to significant benefits to service providers and end users. It will allow 3G operators to economically offload data traffic from wide area wireless spectrum to WLANs in indoor locations, hotspots, and other areas with high user density. If the 3G operators can provide a wide range of WLAN hotspots they will stand to increase their customer base and will also increase the services provided to their customers. For WLAN service providers, integration will bring them a larger user base from partner 3G networks, without having to win them through per customer service contracts. Finally the customers will also benefit greatly from such integration will the advent of greater coverage, higher data rates and lower overall cost of such a combined service.

V. REFERENCES

[1] 3GPP General Packet Radio Service (GPRS) Service description; Stage 2 (Release 4), 3GPP TS 23.060-410, www.3gpp.org, June 2001.

[2] IEEE Std. 802.11b, Supplement to ANSI/IEEE Std. 802.11, 1999 Edition, IEEE Standard for Wireless LAN MAC and PHY Specifications, PDF: ISBN 0-7381-1812-5, January 2000.

[3] C.Perkins, "IP mobility support," RFC2002, Oct. 1996.

[4] M. Moh, G. Berquin, and Y. Chen, "Mobile IP telephony: mobility support of SIP," in Proc. IEEE Computer Commun. And Network, Oct.1999.

[5] 3GPP TS 23.107 V5.5.0, "Universal Mobile Telecommunication System (UMTS); QoS Concept and Architecture", available at http://www.3gpp.org.

[6] Matthew Gast, 802.11 Wireless Networks – The Definitive Guide, O'Reilly, 2002.

[7] ETSI, "Requirements and Architectures for Inter-working between HIPERLAN/2 and 3rd Generation Cellular Systems, "Tech. rep. TR101 957, Aug.2001.

[8] W. Zhuang et al., "Multi-Domain Policy Architecture for IP Multimedia Subsystem in UMTS, "Proc. IFIP/IEEE Net-Con 2002-Network Control and Eng. for QoS, Security and Mobility with Focus on Policy-Based Net., Oct.2002, pp.27-38.

[9] H. Soliman, C. Castelluccia, K. Elmalki, L. Bellier, Hierarchical Mobile IPv6, Internet Draft, draft-ietf-mobileip, hmipv6-04.txt, July 2001.

[10] A. Campbell, J. Gomez, S. Kim, A. Valko, C. Wan, Z. Turanyi, "Design, Implementation, and Evaluation of Cellular IP", IEEE Personal Communications, Vol.7, No.4, pp.42-49, August 2000.

[11] T.M.T. Nguyen et al., "COPS Usage for SLS Negotiation (COPS-SLS), "work in progress, IETF, Feb.2002.

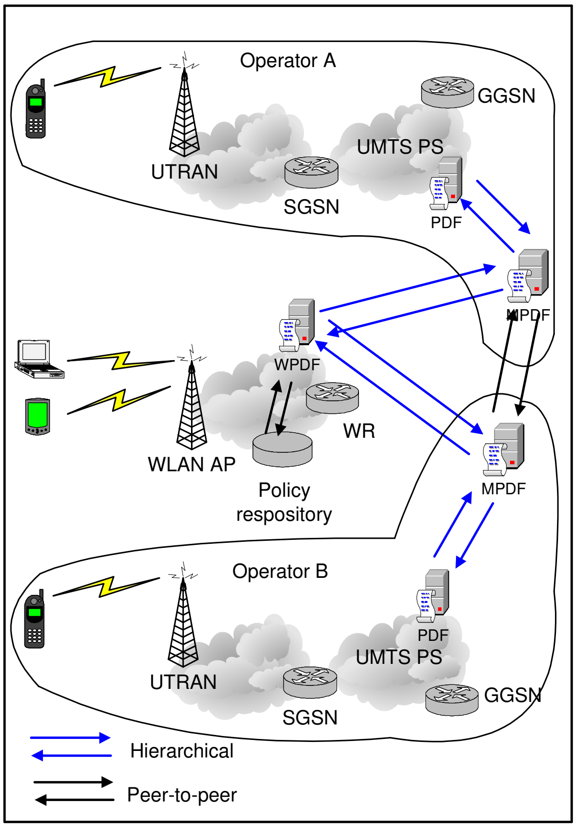


Fig. 4. WLAN shared by different operators

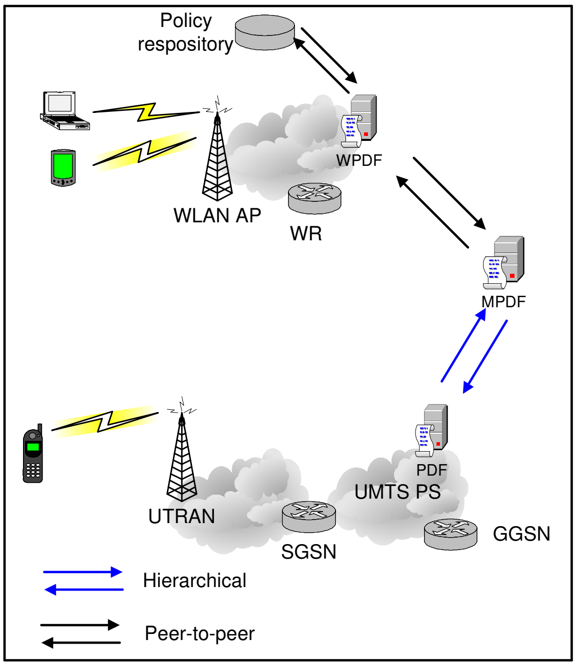


Fig. 5. Interworking of a customer's WLAN network and an operator's UMTS network

Model of OSA / Parlay Gateway For Call Control

Hristo E. Froloshki, Evelina N. Pencheva[1]

Abstract – **Present paper focuses on Parlay/OSA call control functionality. Generic Call Control methods and parameters used in communication with applications were thoroughly analyzed with primary research goal set on definition of OSA compliant generic call agent. Some considerations and guidelines concerning the specialization of the model in supporting different networks are presented as well.**

Keywords – **OSA interfaces, Call Agent, Call Control**

I. Introduction

Parlay/OSA (Open Service Access) architecture is telecommunication industry's response to the challenge of offering flexible and attractive services to customers. It joins IT and telecom efforts in defining a comprehensive set of Application Programming Interfaces (APIs), with final goal set to bring wide developer community in the area of service creation. The approach hides operator network's complexity through strictly defined APIs integrated with well known development environments or coming as software development kits (SDKs) using popular programming language like JAVA. APIs are used by developers to access objects abstracting network resources – these objects are usually run on service platforms [1], directly connected to particular network(s). Although intended for UMTS networks, the principles of Parlay/OSA are applicable in the next generation network domain as well (Figure 1). Some very attractive network capabilities become available (for applications) through Parlay/OSA interfaces: location, mobility, call control, etc. OSA enhances the traditional Intelligent Network (IN) approach of defining building blocks [2] by offering developers objects abstracting network capabilities, enabling them to define the next generation of services for both UMTS and fixed networks. Call control capabilities are split in three Service Capability Features (SCFs): Generic Call Control (GCC), Multiparty Call Control (MPCC), and Multimedia Call Control (MMCC).

Present paper focuses on objects defined in GCC, with the aim to model their behavior in the context of a working Parlay/OSA gateway. OSA specifications [3]-[5] define call control through interfaces of objects, implementing the particular functionality. Implementation of OSA gateway functionality requires generic call agent model, capable of communicating both with application and underlying network.

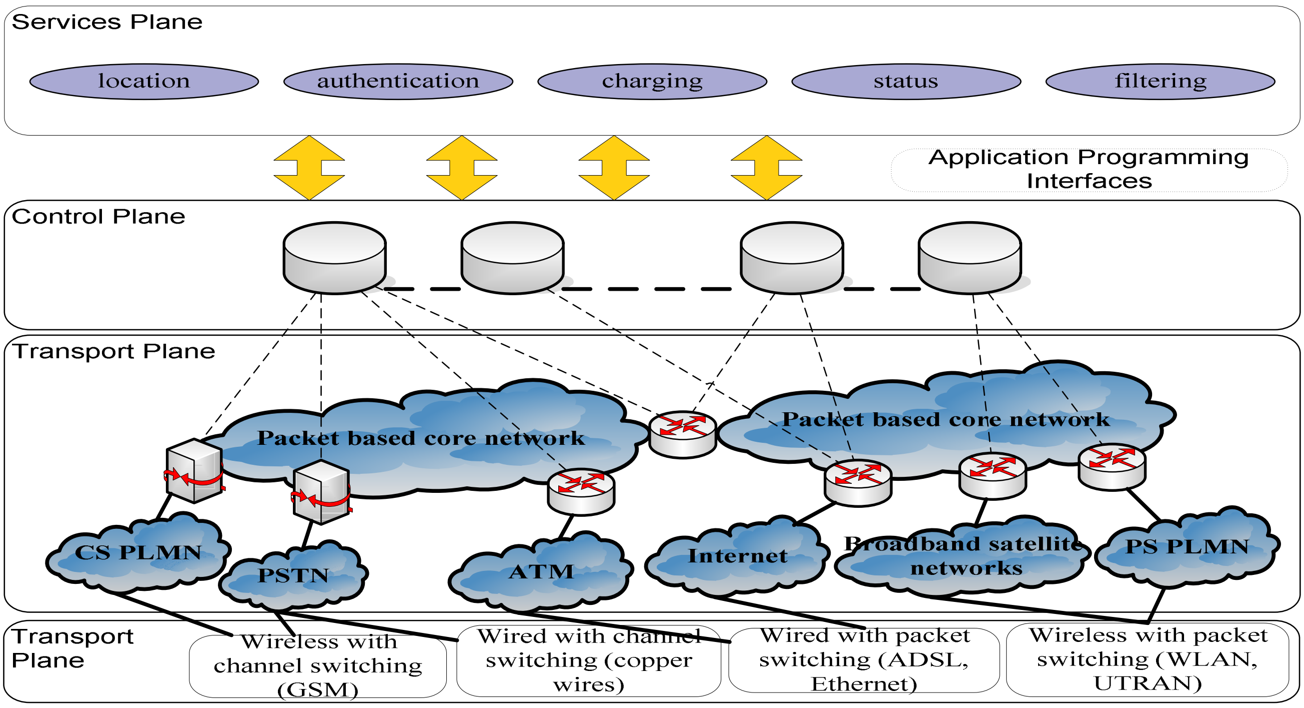


Figure 1. Next Generation Network

[1] Hristo Froloshki and Evelina Pencheva are with Faculty of Communications, Technical University of Sofia, 8, "Kliment Ohridsky" blvd, Sofia 1000, Bulgaria
E-mail:, hef@tu-sofia.bg, enp@tu-sofia.bg

The proposed generic object model is suitable for adaptation to different transport technologies, through the use of generalization and specialization approaches. The paper presents object-oriented call agent model, enabled for application interaction. Possible specializations aiming at compatibility with diverse underlying networks are considered.

II. OSA CALL CONTROL INTERFCES

Generic Call Control SCF provides definition of objects, needed both by entities abstracting underlying network (SCS) and applications [6]. The object responsible for handling application notification is IpCallControlManager. It has the ability to set/remove load control on particular address range but its primary task is the creation of IpCall objects. IpCall is the actual interface, allowing an application to control a call in the underlying network (route/release call, gather charging information, etc.).

Each application should pass successfully authentication and service selection steps, and then the framework instructs service lifecycle manager to create an instance of requested service manager (IpCallControlManager in our case). IpCallControlManager in turn creates IpCall to provide the application with control over a call that matched certain criteria. Most objects in Parlay/OSA utilize asynchronous methods to transfer notifications. An essential step in each object's creation is the setting of a reference to its peer object on application part (i.e. the callback interface).

Functionality described so far is relevant for the service layer of Parlay/OSA architecture. However, a functional call agent model should be able to translate the methods invoked on call abstracting objects into protocol (ISUP, INAP, SIP, MAP, etc.) messages, understandable for the nodes, residing on the resource level.

III. GENERIC CALL AGENT MODEL

IpCallControlManager is the primary interface providing management functions to the generic call control service. It is implemented by Generic Call Control SCF and it must support at least the methods createCall(), enableCallNotification(), disableCallNotification(). An example flow of invoked methods during its operation is presented on Figure 2. Service Instance Lifecycle Manager creates an instance of IpCallControlManager, which enters the "Active" state, where it expects requests from application logic. An application has two options to declare its interest in events associated with a particular call: register its callback interface via "enableCall-Notification()" method or explicitly setting the address of the callback interface on IpCallControlManager. This callback interface will be used to deliver event or state information (party busy, answer, no answer, etc.) to application logic. It is possible for an application to invoke the method "enableCallNotificaiton" several times with different value for the callback interface – each time setting a new callback interface. If the first one fails the next (in order of creation) will be used to deliver event notifications.

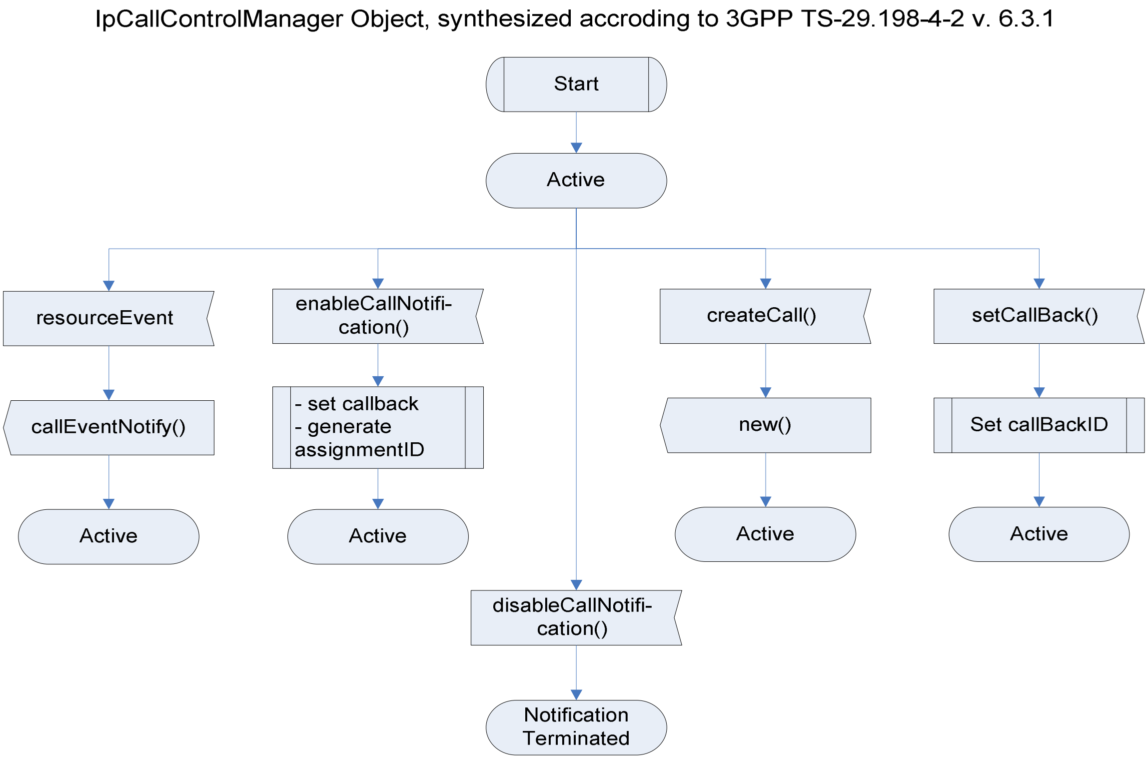


Figure 2. Model of IpCallControlManager

Each “enableCallNotificaiton” request sets an “assignmentID” – to identify the particular event(s) of interest. Having set the callback interface, IpCallControlManager is able to create a call object to represent an actual call taking place in the underlying network. Each IpCall object is assigned unique sessionID value and IpAppCallRef – address of a callback interface to the application. IpCallControlManager then returns to its “Active” state.

In case IpCallControlManager receives “resource Event”, for a call monitored by a certain application, it invokes the method “callEventNotify” on the address pointing the callback reference for IpAppCallControlManager. Exchanged information consists mainly of reference to the object, representing the call and description of the event that occurred in the underlying network.

Application which is no longer interested in certain event invokes “DisableCallNotification()” method with parameter “assignmentID”, in order to terminate monitoring of event in question. IpCallControlManager responds with a transition from “Active” into “NotificaitonTerminated” state.

Minimum requirements for IpCall interface include implementations of the following methods: routeReq(), release() and deassignCall(). IpCall is created by IpCallControlManager, acting on behalf of an application. In its “Active” state IpCall is able to receive some requests (superviseCallReq() and setAdviceOfCharge()), which although executed do not lead to change in state. SuperviseCallReq() method gives an application the opportunity to set a predefined time interval and supervise the call. Important parameters are callSessionID, time (duration) and treatment – defining how the underlying network should process the particular call after timer expiration. SetAdviceOfCharge() method sends charging information to terminals capable of interpreting it. Important parameters for this method are: callSessionID and aOCInfo.

One of the methods causing state transition is “release()” – if IpCall is in “Active” state, an application may invoke the method and send the object in “Application released” state (Figure 3.). If IpCall is subscribed for information regarding the call (previously invoked getCallInfoReq or superviseCallReq) it needs to wait and forward it to IpAppCall object. In case there is no information to for collection the IpCall object is purged.

In another case the underlying network may trigger an event (“Event From Network”), indicating that a call is terminated by one of the calling parties. This makes IpCall invoke “callEnded()” method on IpAppCall to inform the application. If application is subscribed for additional call info, IpCall enters “Network Released State”, and waits for the reports - when they arrive next state is “Finished” and “release()” or “deassignCall()” are legitimate methods for invocation – both lead to the destruction of IpCall and related objects for the particular call. The difference between these two methods is in what happens to the actual call – “deassignCall()” is used when application is no longer interested in controlling it, and frees resources at the OSA/Parlay gateway (call remains in network). Release() causes both call and controlling objects to be released by network and service lifecycle manager respectively. State transition caused by “release()” is possible between “Network released” and “Application released” (not shown on the figure, due to space limitations) – this may happen when call was released in network, but controlling application waits for call-related information to be sent (e.g. for charging purposes). This model reflects the application (service) view of call control functionality accessible through APIs. The main purpose of the model is to hide network protocol complexity from applications.

In order to reflect the specificity of underlying network the model has to be redefined as a specialization of generic functionality. Some guidelines for considering network specificity are given in the next section.

IV. Guidelines for Redefinition of Call Agent Behavior

The idea of a generic call agent model presented encompasses common properties of call control. In terms of SDL this model has to be defined as virtual type that can be redefined in subtypes. The subtypes defined represent the specific network functionality. For example, the part of the model considering call routing is specific for circuit-switched and packet-switched networks. In the model call routing is presented by time delay. Actually the routeReq() transition is virtual type transition that may be redefined and the routing process has to be presented as a procedure. In case of circuit-switched network the procedure includes exchange of Initial Address message (IAM), Address Complete Message (ACM) and Answer Message (ANM) between SCS implementing ISUP and a switch [7]. The sending of routeRes() is triggered by translation of ANM message (indicating that a conversation between parties has begun). Application is informed by the call agent (through routeRes() method) about completing of routing and current call status.

V. Conclusion

We present a generic model of call control agent that can be implemented in NGN. The synthesized model is based on OSA APIs and reflects the call control functionality by the application side. The model describes the call agent behavior by means of methods provided by Call Control interface. The model presents common functionality and hides the protocol complexity of the underlying network. Some guidelines are given to show how the model can be adapted to reflect the network specificity. The full specification of the model including its specializations can be used in implementing OSA gateway.

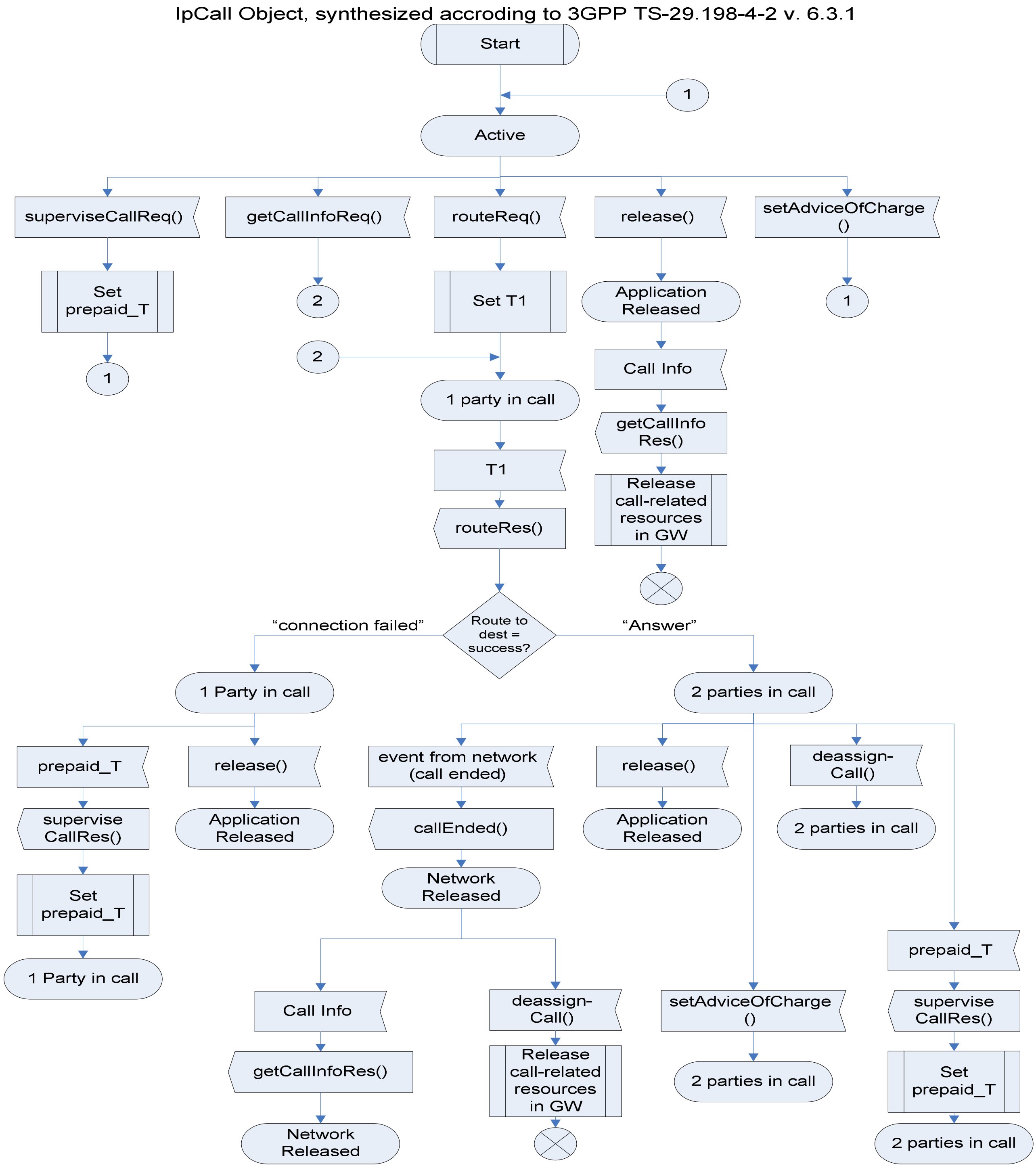


Figure 3. Model of IpCall

REFERENCES

[1] E. S. Chaniotakis, A. E. Papadakis, "Parlay and Mobile Agents in a Homogenized Service Provision", *2nd European Conference on Universal Multi-Service Networks*, Conference Proceedings, pp 150-154, Colmar, France, 2002

[2] ITU-T Rec. Q.1224, Distributed functional plane for intelligent network Capability Set 2

[3] 3GPP TS 29.198-4-2 Open Service Access, Application Programming Interface, Generic Call Control, v6.4.1

[4] 3GPP TS 29.198-4-3 Open Service Access, Application Programming Interface, MultiParty Call Control, v6.4.1

[5] 3GPP TS 29.198-4-4 Open Service Access, Application Programming Interface, MultiMedia Call Control, v6.4.1

[6] J. Zuidweg, *Next Generation Intelligent Networks*, Artech House Inc., 2002

[7] http://www.pt.com/tutorials/ss7/isup.html

Introduction to Telecommunications Network Measurements

Pencho K. Penchev[1]

Abstract – **In this paper, a telecommunications network test and measurements are activities undertaken to characterize the operation of networks and network elements. Measurement determines performance parameters, either on an as-needed basis or continuously via dedicated monitoring equipment. Test adds the comparison of measured parameters to accept/reject thresholds, or the application of controlled stimuli.**

Keywords – **telecommunication, measurements and testing, protocol analyzer, digital performance testing, analog performance testing.**

I. Introduction

Network technology is changing at an increasing rate. In the past, major investments in transmission and switching technology took many years to depreciate. Today, however, the pressures of the market and the advances in technology demand more rapid turnover. The unrelenting rollout of new technology creates challenges for new test equipment and maintenance strategies.

The business climate in telecommunications is changing, too. Because of competition and deregulation, combined with the increasing importance of telecommunications for business activities, network operators are becoming more service- and customer-focused. Network performance is measured in terms of quality of service (QoS). Successful delivery is measured by the highest quality at the lowest prices. Network operators also need to bring new technology into service very quickly to create competitive advantage.

II. Quality of Service

Network quality of service can be characterized by five basic performance measures:

Network availability (low downtime).

Error performance.

Lost calls or transmissions due to network congestion.

Connection setup time.

Speed of fault detection and correction.

Network availability and error performance are usually the parameters that service providers guarantee in terms of QoS. Generally these parameters need to be checked while the network is in service (i.e., carrying traffic), using a network management system.

Lost calls and call setup time are the main criteria for measuring performance in switched networks, often indicating whether network planning is keeping up with traffic growth and changing traffic types. The move to common-channel signaling has greatly reduced call setup time, while also increasing system flexibility for offering new services. The growth of Internet traffic, however, with long holding times on the circuit-switched network, has again called into question network performance.

Some network operators now guarantee that they will fix faults within a specified time or pay compensation to the customer. This requires good processes for troubleshooting, well-trained technicians with access to powerful test equipment, and probably the use of centralized automatic test systems for rapid fault finding.

III. Testing Objectives

An initial reaction to network testing might be that it is something to be avoided if possible because it cost time and money. On reflection, however, effective testing can add value rather than being an expense, and can enhance the network operator's business.

There are three major business problems that are driving operators today:

Time-to-market of new products and services.

Reducing the cost of delivering a service.

Improving and guaranteeing service quality.

Network testing can be divided into three application areas:

Bringing new equipment and systems into service.

Troubleshooting and detecting network degradation.

Monitoring and ensuring quality of service.

Bringing new equipment and systems into service. When new equipment is installed and brought into service, the installer (who may be the equipment manufacturer) makes a comprehensive series of tests. These tests usually are made to more restrictive limits than normal performance expectations. These limits are specified in ITU-T Recommendation M.2100 (formerly M.550), "Performance limits for bringing Into-Service and Maintenance of International PDH Paths, Sections and Transmission Systems".

Once a system is in use, performance standards must be maintained. When service degradation occurs, it must be determined whether the fault exists within a particular vendor's network or elsewhere. This information is determined most effectively by in-service testing or performance monitoring. Many test instruments also provide some degree of nonintrusive testing.

In-service maintenance. Once equipment is in service, long periods of downtime are unacceptable, so maintenance

[1] Pencho K. Penchev is with the Department of Communication Techniques and Technologies, Technical University – Gabrovo, 4 H. Dimitar St., 5300 Gabrovo, Bulgaria, E-mail: pkpen@tugab.bg

strategy has to be carefully thought out. ITU-T Recommendation M.20, “Maintenance Philosophy for telecommunications Networks”, defines three types of maintenance strategy:

Preventive maintenance
Corrective maintenance
Controlled maintenance

Preventive maintenance is carried out at predetermined intervals to reduce the probability of failure or degradation of performance. This method was commonly applied to older analog systems that needed periodic adjustment to compensate for drift.

Corrective maintenance is carried out after a failure or degradation is reported by a monitoring system or user.

Controlled maintenance involves centralized network monitoring and identification of degraded network performance. Centralized monitoring can be supplemented by field maintenance teams using portable test equipment.

Of these methods, controlled maintenance is preferred for maintaining high levels of QoS. It provides early warning of degradations and potential failures, thereby reducing system downtime. Repair work and adjustments can be anticipated and scheduled for quiet periods. In this way, disruption also is minimized.

IV. ANALOG PERFORMANCE TESTING

Figure 1 shows the major measurable elements of an analog transmission system. The simplest analog test is to measure the system gain and signal-to-noise (S/N) ratio between the end-to-end telephone connections. The test is usually made with a portable Transmission Impairment Measuring Set (TIMS). The test operator sends a fixed tone into the system and makes measurements at the opposite end to check for signal level and signal-to-noise ratio (noise with tone). When an analog data modem is to be used on the path, various data-impairment measurements may be specified, such as impulse noise, phase jitter and gain/phase hits.

Although telecommunications networks are now largely digitized, the connection between a telephone and the exchange continues in most cases to be analog. TIMS measurements are therefore still important. In the 1998s, wideband TIMS measurements up to 200 kHz have been used to evaluate local loops for ISDN and digital data transmission.

Similar kinds of measurement can be made in the analog multiplex system using a Selective Level Measuring Set (SLMS). Because this is a frequency Division Multiplex (FDM) system, possibly carrying several thousand separate telephone channels in a bandwidth of up to 65 MHz, the SLMS has to be able to select and measure individual channels as well as the pilot tones inserted to control system levels. FDM systems operate either over coaxial cable or microwave radio.

An FDM multichannel traffic signal resembles while noise; system impairments create degraded signal-to-noise ratio in individual telephone channels due to intermodulation distortion, particularly under heavy traffic loading. To evaluate this, the out-of-service noise-loading test is made using a notched white noise test stimulus. By measuring the noise level in the notch at the receiving end, the equivalent signal-to-noise degradation can be estimated as a function traffic load. In the analog era this was a very important test, particularly for microwave radio, because impairments are additive in analog systems.

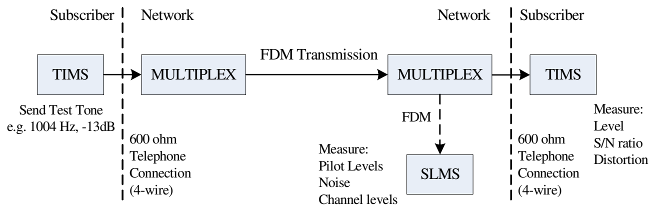


Fig. 1. Analog system performance measurements can be made either at the local loop access voice band frequencies using a TIMS, usually at the 4-wire 600-ohm line, or at the FDM line level is using a SLMS

V. DIGITAL PERFORMANCE TESTING

The tests made on digital transmission systems can be divided into several categories, as shown in Figure 2 and Table 1.

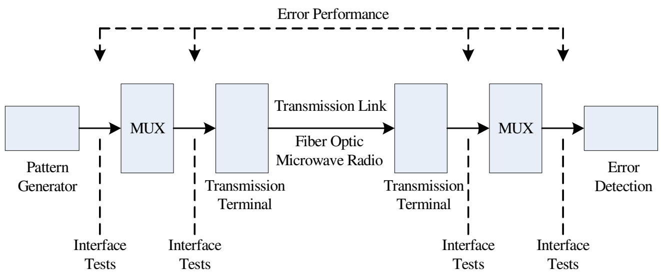


Fig. 2. Digital transmission measurements fall into two main categories. Interface tests check the compatibility of the electrical or optical interfaces of equipment to ensure error-free interconnection.

Interface specifications and tests. Anyone familiar with RF and microwave knows the importance of matching at interfaces so that the performance of cascaded networks equals the sum of the parts. The same is true in digital communications. If instrument parameters do not match equipment parameters, bit error appears when they are connected. This matching is defined in a series of interface specifications contained in ITU-T Recommendation G.703; Recommendations G.823/824 and G.825 address timing jitter.

Electrical interface specifications are usually measured during equipment design and manufacture to ensure compatible interconnection between network elements at a Network Node Interface (NNI) and User Network Interface (UNI).

The ITU-T specifications include pulse height (voltage level); pulse shape (rise time, fall time, overshoot, and duty cycle); and equality of positive and negative pulses in a ternary signal. These measurements usually are made with an oscilloscope to check that the pulse shape falls within a prescribed mask.

TABLE I
CATEGORIES OF DIGITAL TRANSMISSION TESTS
AND APPROPRIATE ITU-T RECOMMENDATIONS

Type of test	Typical tests	Relevant ITU-T standards
Interface tests	PCM Codec	G.712/713/714 (Q.131-133 measurement)
	Pulse shape Clock frequency Voltage/impedance Coding	G.703
	Framing	G.704/706/708
	Jitter wander	G.823/824/825 (Q.171 measurement)
Out-of-service error Performance tests (Installations and commissioning)	BER using PRBS patterns	G.821/826 (O.151 measurement)
In-service error performance tests (maintenance, fault finding, quality of service)	Code error Frame error Parity error	G.821/826 M.2100/2110

The physical interface is usually 75-ohm coaxial cable with a return loss of 15 - 20 dB, although with higher-speed SONET/SDH equipment the physical interface may be fiber optic. In addition, bit rates must be maintained within strict limits, and the tester must check that receiving equipment can operate correctly within this tolerance. Interface coding specifications include algorithms for AMI, HDB3, CMI, B3ZS, etc.

Timing jitter is defined by ITU-T as short-term variations of a digital signal's significant instants from their ideal positions in time. The significant instant might be the rising or falling edge of a pulse.

The simplest way to measure jitter is with an oscilloscope and eye diagram. Jitter appears as a spread or "muzziness" in the vertical transitions. Most telecommunications systems, however, require more precise measurements. In these cases it is essential to know how the level of jitter varies with jitter frequency. This relationship is measured with a jitter test set that demodulates the jitter signal.

Jitter itself must be checked at the input by gradually increasing the level of jitter on a test signal until bit error occurs. In addition, the tester must check that jitter present at the input is not magnified by the equipment; otherwise, problems can arise when several pieces of equipment are cascaded in a network. This measurement is called jitter transfer.

If equipment conforms fully to all the interface specifications, in principle it should be possible to construct any arbitrary network without generating bit errors. Problems still can arise, however, if the live traffic signals have very extreme pattern densities that are not fully simulated by the out-of-service PRBS test.

Error performance tests. Digital network error performance can be measured over a complete end-to-end connection called a path, or over parts of the network called lines and sections. These network segments are illustrated in Figure 3. Path measurements indicate the overall quality of service to the customer. Line and section measurements are used for troubleshooting, installation and maintenance, and for assuring transmission objectives are met.

The fundamental measure of performance or quality in digital systems is the probability of a transmitted bit being received in error. With the latest equipment, the probabilities of this occurrence are very low, on the order of 10-12 or less. It is still necessary to measure the performance of these systems, however, and in particular to analyze the available margins of safety, and to explore potential weaknesses that later could lead to degrades performance.

In-service and out-of-service measurements. In-service error performance measurements rely on checking known bit patterns in an otherwise random data stream of live traffic. Some in-service measurements are more representative than others of the actual error performance of the traffic signal. Furthermore, some are applicable to the path measurement, provided the parameters are not reset at an intermediate network node. Others are only useful at the line or section level. The most commonly used error detection codes (EDCs) are frame word errors, parity errors, or cyclic redundancy checksum errors.

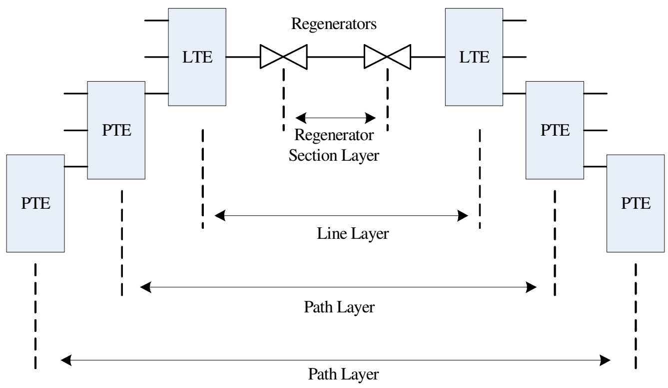


Fig. 3. A digital transmission system can be viewed as an overall end-to-end path terminated by Path Terminating Equipment (PTE).

Out-of-service measurements involve removing live traffic from the link and replacing it with a known test signal, usually a pseudorandom binary sequence (PRBS). These tests are disruptive if applied to working networks, but are ideal for installation and commissioning tests because they give precise performance measurement. Every bit is checked for error. Although the PRBS appears random to the digital system, the error detector (Figure 2) knows exactly what it should receive and so detects every error. The error detector calculates the probability of error as the bit error ratio (BER). BER is defined as the number of errors counted in the measurement period, divided by the total number of bits received in the measurement period.

Thus the bit errors or error events can be detected by out-of-service techniques. These are sometimes referred to as the performance primitives. To be useful for assessing quality of service, however, they must be analyzed statistically as a function of time according to the various error performance standards specified in Table 1. This analysis yields percentages for the availability of a digital communication link, and the portion of time that it exceeds certain performance criteria that are acceptable to the customer. One of the important standards is the ITU-T Recommendation M.2100/2110.

VI. Protocol analysis in the telecommunication network

Up to this point we have discussed the capability of the telecom network to transmit digital bits or analog signals over a path without errors or quality degradation. Testing BER, for example, assumes that the traffic carried by the network is completely random data, or at least that the payload within a frame structure is random. This apparently random traffic signal will, in fact, always have a structure. It might be a PCM voice signal, a data signal, a signaling message for controlling network switching, or possibly an ISDN signal or an ATM cell data stream for broadband services.

When telecom networks were predominantly carrying voice traffic using in-band signaling, there was little interest in checking the information content of traffic or knowing if it conformed to the rules or protocols of data communications, except for the X.25 packet-switched network.

Networks today are much more sophisticated and carry a wide range of different services among many vendors. Rather than being just the transporter of telecommunications traffic, increasingly the network is an integral part of the information structure created by the convergence of computers and communications. The most significant example of this is the common-channel signaling system (SS7), which interconnects the network switches and databases and controls all aspects of service delivery and billing. A large amount of analysis and monitoring is required, not so much of the data transmission itself, but of the messages and transactions taking place. An important example of signaling transactions occurs in a cellular telephone network, when constant reference to databases is necessary for tracking the location of mobile phones during handover from one cell to the next, and for billing and verifying legitimate users.

In order to analyze and troubleshoot these systems, a protocol analyzer is required, in some cases dedicated to the particular application such as SS7 or ATM. Protocol analyzers have been in use for many years in local and wide area networks, predominantly in enterprise networks (see Figure 4). This traditional data communications test tool is now finding its way into the telecom network as traffic becomes more data-intensive.

Often in data communications there is the need to observe and analyze, or even to simulate, the interactions between network devices interconnected by WANs or LANs. The need may be in the context of one or more of the following scenarios:

- The developer of network equipment or the network itself needs to analyze and simulate operation under a number of circumstances.
- The network planner needs to measure current levels of network use and then anticipate future needs as new services and capabilities are added.
- The installer (of computers, communications equipment, and/or networks) needs to commission and test a network system's devices and their interactions.
- Field service personnel for a computer and communications equipment vendor, or for a service provider, are faced with troubleshooting an intermittent problem.
- The network manager of a private network operates with system elements from several vendors and uses multiple service providers in order to get the best performance, reliability, and price. When a problem arises, a tool is needed to determine its source so as to avoid finger-pointing among the vendors.

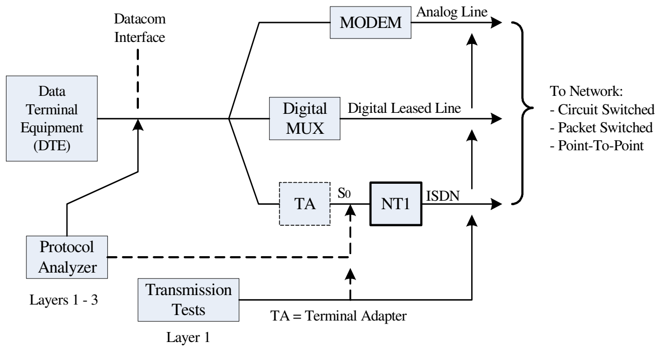


Fig. 4. Protocol analyzers traditionally have been used for testing data communications networks at datacom interfaces, usually at the customer's premises.

In each of these scenarios, there is need for an instrument that can observe nonintrusively and help the user interpret the complex interactions within the data communications protocols that control the behaviors of the devices. In some cases there is need to simulate network elements to test for problems. In other situations, there is an application to measure the performance and utilization of the network and of the devices within it.

These tests and measurements may be made reactively when a problem occurs, or may be made proactively when looking for trends that indicate developing problems. When new services are being introduced, or new equipment is being installed or system software upgraded, it is necessary to emulate specific messages or protocols to confirm correct operation of the network.

VII. Conclusion

In this paper, the telecommunications network measurements introduced the terms of quality of service. Successful delivery is measured by the highest quality at the lowest prices. Network operators also need to bring new technology into service very quickly to create competitive advantage.

References

[1] McDowall, Ron. When SS7 Is Put to the Test, Global Telephony, 1994
[2] Walker, Hugh. Testing and Troubleshooting Datacom Circuits, Evaluation Engineering, 1988.
[3] ITU-T Q.171, Timing Jitter and Wander Measuring Equipment for Digital Systems.
[4] ITU-T G.823, The Control of Jitter and Wander within Digital Networks based on the 2048 kbps Hierarchy.

Evaluation of a Limited Multi-Server Queue with a Generalized Poisson Input Stream*

Seferin T. Mirtchev[1] and Stanimir I. Statev[2]

Abstract: **This paper deals with M(g)/M/n/k/S queue in which we have generalized Poisson arrival process, exponential service time, multiple servers, limited waiting positions and finite number of customers. We use the generalized input Poisson stream that can be peaked, regular or smooth. The idea is based on the analytic continuation of the Poisson distribution and the classic delay systems. We apply techniques based on birth and death processes and state-dependent arrival rates. The influence of the peaked factor on the traffic characteristics is studied. It is shown that the input stream changes significantly the characteristics of the delay systems.**

Keywords: **Queueing system, Poisson process, Peaked and smooth traffic**

I. Introduction

Many studies on traffic measurements from a variety of communication networks, like Ethernet local area networks (LANs) and wide area networks (WANs) with Internet and asynchronous transfer mode (ATM), etc., have shown considerable difference between actual network traffic and assumptions in traditional theoretical traffic models.

Problems with the Poisson modelling are predicted in [10]. The authors are indicating that some arrivals deviate considerably from the Poisson distribution but user-initiated TCP session arrivals, such as remote-login and file-transfer, are well-modelled as Poisson processes with fixed hourly rates. The Internet traffic characteristics are studied by Cao [1]. They have shown that the arrivals tend to Poisson and the packet sizes tend to independence when the number of simultaneous transport connections increase.

Karagianis [6] believed that it is time to re-examine the Poisson traffic assumption in relation to the traffic carried within the Internet core. They have shown that the current network traffic can be well represented by the Poisson model for sub-second time scales.

The traffic flows inside a network are not Poissonion in general [2,4]. For many real teletraffic systems the mean number of events in an interval is not equal to the variance. The offered streams are said to be peaked or smooth according to whether the variance is bigger or smaller than the mean value, respectively.

The MMPP (Markov Modulated Poisson Process) traffic model that accurately approximates the long range dependence characteristics of Internet traffic traces is proposed in [9,11]. The Poisson Pareto Burst Process (PPBP) is presented in [12] as a simple and accurate model for aggregated Internet traffic.

Network analysis really requires a technique that can represent any kind of traffic, peaked or smooth, within the same model [5,8]. Most of the known methods are designed for a particular type of traffic, peaked or smooth, or, if they apply to both, do so using different models for different ranges of peakedness. The presented below method meets the above requirements.

In this paper peaked and smooth input streams are defined. They will be called a generalized Poisson process. A calculation method for the performance measures of a M(g)/M/n/k/S queue in witch we have generalized Poisson arrival process, exponential service time, multiple servers, limited waiting positions and finite number of customers is presented. The idea is based on the analytic continuation of the Poisson distribution and the classic M/M/n system. We apply techniques based on birth and death process and state-dependent arrival rates.

II. Generalized Poisson Process

The Poisson process is a pure birth process with an arrival rate λ independent of the system state. The probability $P_i(t)$ of i arrivals in an interval witch duration is t seconds is given by

$$P_i(t) = \frac{(\lambda t)^i}{i!} e^{-\lambda t} . \qquad (1)$$

Two more parameters, peakedness factor p and number of sources S is introduced for the generalized Poisson process. The process is said to be peaked, regular or smooth according to whether $p > 1$, $p = 1$ or $p < 1$, respectively.

Calls arrive in a generalized Poisson stream at rate λ_i which depends on the number of calls in the system. The time between successive call arrivals is exponentially distributed with different parameter λ_i. This generalized Poisson stream has memoryless property.

The arrival rate is state-dependent

$$\lambda_i = \lambda (i+1)^{1-1/p} . \qquad (2)$$

The state probabilities $P_i\,(t)$ in the case of a generalized Poisson process are

$$P_i(t) = \frac{(\lambda t)^i / (i!)^{1/p}}{\sum_{j=0}^{S} (\lambda t)^j / (j!)^{1/p}} . \qquad (3)$$

The mean value (the average number of arrivals in an interval of length t) is

$$M(t) = \sum_{i=1}^{S} i P_i . \qquad (4)$$

The variance of the number of arrivals in an interval of length t is

$$V(t) = \sum_{i=0}^{S} [i - M(t)]^2 P_i(t) . \qquad (5)$$

[1] Seferin T. Mirtchev, Faculty of Communications Technology, 8 Kl. Ohridsky St., 1000 Sofia, Bulgaria, stm@tu-sofia.bg
[2] Stanimir I. Statev, Faculty of Communications Technology, 8 Kl. Ohridsky St, 1000 Sofia, Bulgaria, stanimir_statev@yahoo.com
* This work was supported by Bulgarian National Science Fund under grant ВУ-ТН-105/2005.

When $p = 1$ and $S\rightarrow\infty$, $M(t) = \lambda t$ and $V(t) = \lambda t$ i.e. it is a regular Poisson process.

III. MULTI-SERVER QUEUE - MODEL DESCRIPTION

Let us consider a multi server queue M(g)/M/n/k/S with a generalized Poisson input stream M(g), exponential service time M, number of servers n, limited waiting room k and number of sources S (S>k). This is a birth and death process and we can use the general solution, as given in [5], for the stationary probability of having j customers in the system

$$P_j = \frac{\prod_{i=0}^{j-1}\lambda_i/\mu_{i+1}}{1+\sum_{v=1}^{k+1}\prod_{i=0}^{v-1}\lambda_i/\mu_{i+1}} \qquad j=0,1,2,...,n+k. \qquad (6)$$

This generalized delay system may be described by selecting the birth-death coefficient as follows

$$\begin{aligned} &\lambda_j = \lambda\,(j+1)^{1-1/p} && j=0,1,2,...,n+k \\ &\mu_j = j\mu && j=1,2,3,...,n \\ &\mu_j = n\mu && j=n,n+1,...,n+k \end{aligned} \qquad (7)$$

The arrival rate is state-dependent and depends on the peakedness factor p. This limited delay system is always ergodic. The finite state-transition diagram is shown in Fig.1.

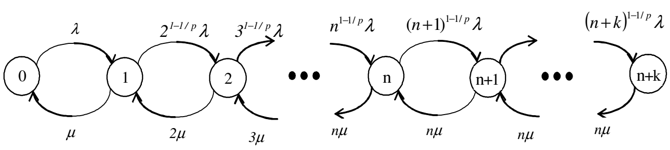


Fig. 1. A state-transition diagram - M(g)/M/n/k/S queue

Applying these coefficients to the general solution of the birth and death process and using traffic intensity $a = \lambda/\mu$ we obtain the steady state probabilities

$$P_j = \frac{a^j}{(j!)^{1/p}}P_0 \qquad 0 \le j \le n$$

$$P_j = \frac{a^j}{(j!)^{1/p}}\frac{j!}{n!n^{j-n}}P_0 \qquad n \le j \le n+k. \qquad (8)$$

$$P_0 = \frac{1}{\sum_{i=0}^{n}\frac{a^i}{(i!)^{1/p}} + \sum_{i=n+1}^{n+k}\frac{a^i}{(i!)^{1/p}}\frac{i!}{n!n^{i-n}}}$$

The offered traffic is calculated by means of the average arrival rate and the mean holding time

$$A = \bar{\lambda}\frac{1}{\mu} = \frac{1}{\mu}\sum_{j=0}^{n+k}\lambda_j P_j = a\sum_{j=0}^{n+k}(j+1)^{1-1/p}P_j . \qquad (9)$$

The carried traffic is equivalent to the average number of busy servers

$$A_c = \sum_{i=0}^{n} i\,P_i + n\sum_{i=n+1}^{n+k}P_i . \qquad (10)$$

IV. GENERALIZED ERLANG DISTRIBUTION

Assume that the number of the servers is equal to the number of the sources. In this case the system has not any losses and delay, the whole offered traffic is carried and it is called the intended traffic load.

The stationary probability of having j customers in the system has generalized Erlang distribution

$$P_j' = \frac{a^j/(j!)^{1/p}}{\sum_{i=0}^{S}a^i/(i!)^{1/p}} \qquad j=0,1,2,...,S. \qquad (11)$$

The intended traffic is the mean number of busy servers

$$A_i = \sum_{j=1}^{S} jP_j' . \qquad (12)$$

The variance of the intended traffic is

$$V(A_i) = \sum_{j=0}^{S}(j-A_i)^2 P_j' . \qquad (13)$$

The peakedness of the intended traffic is the variance to mean ratio

$$z = \frac{V(A_i)}{A_i}. \qquad (14)$$

V. Mg/M/n/k/S – TRAFFIC CHARACTERISTICS

BLOCKING PROBABILITY. The time congestion probability B_t describes the fraction of time that all waiting rooms are busy

$$B_t = P_{n+k} . \qquad (15)$$

The call congestion probability B_C is ratio of lost traffic (offered minus carried traffic) to offered traffic

$$B_c = \frac{A-A_c}{A}. \qquad (16)$$

WAITING PROBABILITY. The waiting probability is denoted by $P(>0)$ which means that the waiting time probability is greater than 0

$$P(>0) = 1-\sum_{i=0}^{n-1}P_i - P_{n+k} . \qquad (17)$$

MEAN NUMBER OF CALLS. The mean number of calls present in the system in steady state by definition is

$$L = \sum_{j=1}^{n+k} jP_j . \qquad (18)$$

MEAN SYSTEM TIME. From the Little formula, we have the mean system time

$$T = \frac{L}{\bar{\lambda}} = L\Big/\sum_{j=0}^{n+k}\lambda_j P_j . \qquad (19)$$

WAITING TIME DISTRIBUTION. Let us assume the first-come-first-out (FIFO) discipline. The waiting time distribution function $P(>t')$ is defined as the probability of waiting time exceeding t'. From the probability theory it is given by

$$P(>t') = \sum_{i=n}^{n+k-1}P_i Q_i(>t') . \qquad (20)$$

An arbitrary call enters service when i calls are in the system P_i. Since the service time is exponentially distributed the probability that i calls terminate in time $(0,t']$ becomes a Poisson distribution with a mean value $n\mu t'$

$$Q_i(t') = \frac{(n\mu\,t')^i}{i!}e^{-n\mu t'} . \qquad (21)$$

The conditional probability $Q_i(>t')$ that the arbitrary call has to wait longer than t', given i calls in the system, is expressed by:

$$Q_i(>t') = \sum_{r=0}^{i-n} \frac{(n\mu t')^r}{r!} e^{-n\mu t'} \,. \qquad (22)$$

VI. State Probability Calculation

The traffic intensity a is not equal to the intended traffic in a case of a generalized Erlang process because we calculate the power of the Erlang unsymmetrical distribution. That is why we have to calculate the intended traffic A_i and the peakedness z when defining the traffic intensity a and peakedness factor p.

From the practical point of view we first define the intended traffic A_i and the peakedness z and after that calculate the traffic intensity a and peakedness factor p.

A fundamental question about the system defined by Eqs. (11), (12) and (14) is whether there exist solutions a, p for an arbitrary A_i, z. Although no formal proof seems to exist, this seems to be the case and the solution appears to be unique.

We can find solutions of the above system with the iterating method of consecutive replacements.

VII. Numerical Results

In this section we give numerical results obtained by a Pascal program on a personal computer. The described methods were tested on a computer over a wide range of arguments.

Figure 2 shows the generalized Erlang distribution where the intended traffic is $A_i = 10$ *erl*, the number of the sources $S = 100$ and the peakedness z is change from 0.6 to 1.5. It will be seen that when the peakedness z increases the probability distribution becomes broad about the mean.

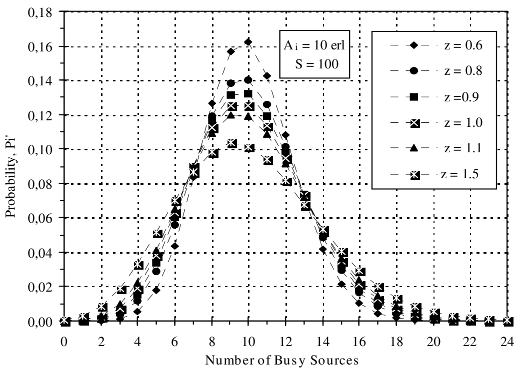


Fig. 2. Generalized Erlang distribution

Figures 3 and 4 illustrate the stationary probability distribution in a multi-server queue M(g)/M/n/k/S with a generalized Poisson input stream, 10 servers, 30 waiting rooms, 100 sources and different intended traffic A_i and peakedness z. It will be seen that when the utilization is from 0.9 to 1 erl and the peakedness is bigger than one the probabilities increase when the number of the customers in the system increases.

Figures 5 and 6 show the call and time congestion probabilities in a multiple delay system with 10 servers, 100 sources, 0.9 erl intended traffic and different peakedness as function of the buffer size. When the utilization is high (0.9 - 1 erl) and the input stream is peaked (z = 1.05 – 1.2) the influence of the buffer size of the congestion probability is negligible. We have to notice that the offered traffic is bigger than intended in the M(g)/M/n/k/S queue when the input stream is peaked.

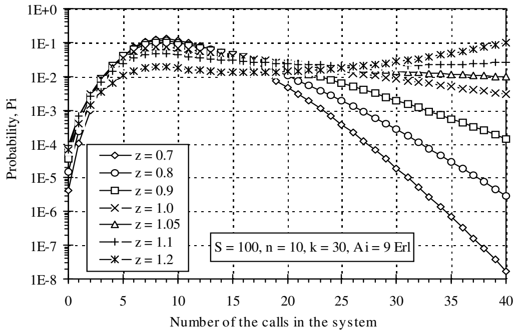


Fig. 3. Stationary probability distribution

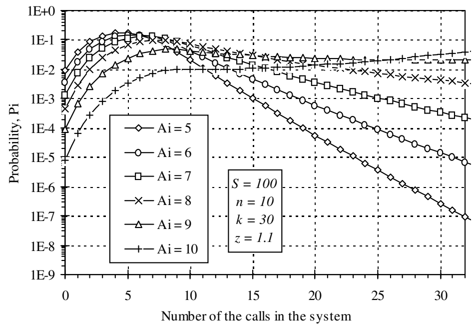


Fig. 4. Stationary probability distribution

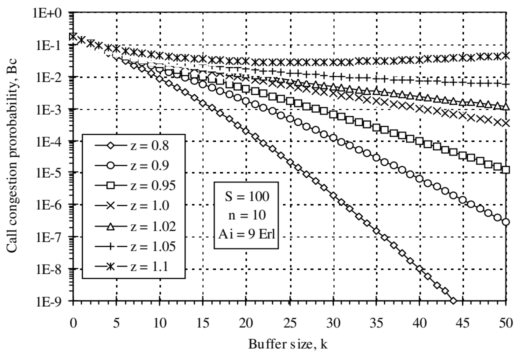


Fig. 5. Call congestion probability

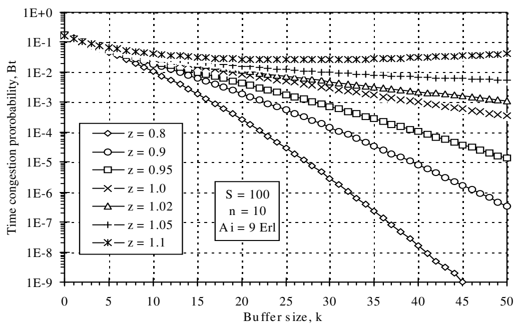


Fig. 6. Time congestion probability

Figure 7 presents the normalized mean system time ($W' = W/\tau$) as function of the intended traffic when the number of servers is 10, the number of sources is 100, the peakedness is 0.9 and 1.1 respectively and different waiting room.

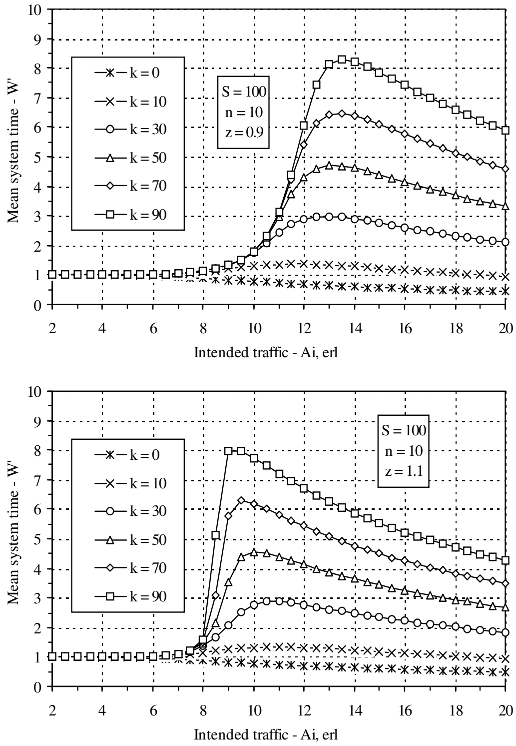


Fig. 7. Normalized mean system time

Figure 8 illustrates the waiting time distribution as function of the normalized waiting time when the number of servers is 10, the number of sources is 100, the buffer size is 30, the peakedness is 0.9 and 1.1 respectively and different intended traffic.

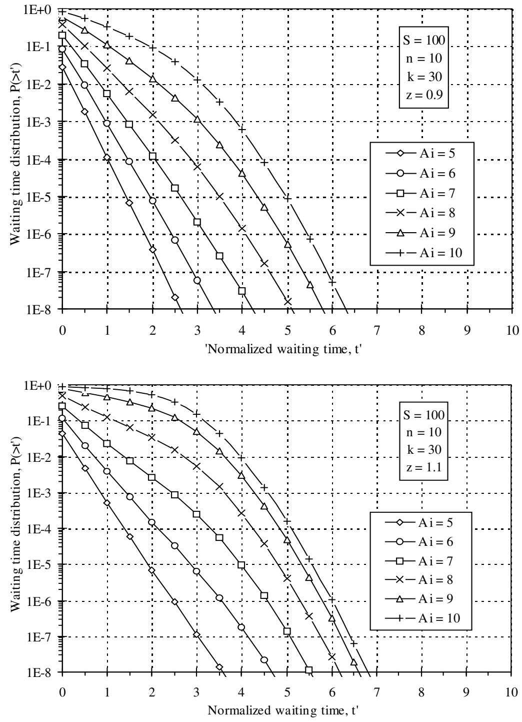


Fig. 8. Waiting time distribution

VIII. CONCLUSIONS

In this paper a generalized Poisson process is introduced and evaluated. A basic model for a queueing system Mg/M/n/k/S is examined in detail.

The proposed method provides a unified framework to model peaked and smooth traffic. Numerical results and subsequent experience have shown that this method is accurate and useful in both analyses and simulations of teletraffic systems.

The importance of a multiple delay system in a case of a generalized Poisson input stream comes from its ability to describe behaviour that is to be found in more complex real queueing systems. It is the case in a general traffic system, which is an important feature in designing telecommunication systems.

It is shown that the influence of the peakedness over the performance measures is significant.

In conclusion, we believe that the presented generalized Poisson process and queueing system will be useful in practice. As part of future work, we plan to analyzed a processor sharing system with a generalized Poisson input stream.

REFERENCES

1. Cao J.,W. Cleveland, D. Lin, and D. Sun. Internet Traffic Tends to Poisson and Independent as the Load Increases. Bell Labs Technical Report, 2001.
2. Delbrouck L.E.N. A Unified Approximate Evaluation of Congestion Functions for Smooth and Peaky Traffics, IEEE Trans. on Commun., Vol.29, No 2, 1981, pp.85-91.
3. Girard A. Routing and Dimensioning in Circuit-Switched Networks, Addison-Wesley, 1990.
4. Iversen, V.B. Teletraffic Engineering Handbook, ITU-D & ITC. 312 pp. Edition spring 2004, http://www.com.dtu.dk/education/34340/.
5. Iversen V. and S. Mirtchev. Generalized Erlang Loss Formula, Electronics Letters, Vol. 32, No: 8, April 1996, pp.712-713.
6. Karagiannis T., M. Molle, M. Faloutsos, A. Broido, A Nonstationary Poisson View of Internet Traffic, IEEE INFOCOM 2004, Vol.23,no.1, March 2004, pp.1559-1570.
7. Kleinrock L. Queueing Systems, Volume I: Theory, John Wiley & Sons, 1975.
8. Mirtchev S. and I. Stanev. Evaluation of a Single Server Delay System with a Generalized Poisson Input Stream, ITC19, Beijing, China, Vol.6a, 2005, pp.553-542.
9. Muscariello L., M.Mellia, M.Meo, M.Ajmone Marsan, R. Lo Cigno, An MMPP-Based Hierarchical Model of Internet Traffic, IEEE ICC 2004, Vol.27, no.1, June 2004, pp.2143-2147.
10. Paxson V. and S. Floyd. Wide Area Traffic: The Failure of Poisson Modelling, IEEE/ACM Transactions on Networking, Vol. 3, no.3, June 1995, pp. 226-244.
11. Radev, D., Stochastic Process Generation for Traffic Flow Simulation. IEEE Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services TELSIKS'2005, Vol. 2, 2005, Nis, p-p 559-562.
12. Zukerman M., T. Neame and R. Addie, Internet Traffic Modelling and Future Technology Implications, IEEE INFOCOM 2003, Vol.22, no.1,March 2003, pp.578-596.

VoIP Traffic Shaping In All IP Networks

R. Iv. Goleva[1], M. At. Goleva[2], D. K. Atamian[3], K. At. Golev[4], P. Iv. Merdjanov[5]

***Abstract* - Traffic shaping is a phenomenon of any change of the traffic characteristics due to the active or passive management. Shaping techniques are considered very important for quality of service support. This paper represents analyses of shaping effect in devices that apply the three mostly used QoS management techniques – IntServ, DiffServ and RSVP. Special attention is paid to the Voice over IP traffic. The results show how VoIP traffic changes its shape starting from the traffic source and passing first access device that apply QoS technique.**

***Keywords* - Voice over IP, Quality of Service, shaping, DiffServ, RSVP.**

I. INTRODUCTION

IP networks are one of the most challenging areas for investigation as they become the only common transmission technology. This increases the demand for Quality of Service (QoS) support. There are plenty of papers and books that try to gather with QoS in IP network [2], [3], [8], [9]. In this paper we demonstrate the shape of the VoIP traffic starting from traffic source and ending at the core network. Furthermore, we show how this shape can be manipulated with the three mostly used QoS management techniques – IntServ, DiffServ, and RSVP. The analyses are made on the basis of the three popular services – VoIP, LAN emulation, transaction exchange. The stress in this paper is on VoIP as the most delay sensitive service. The shaping effect is estimated without implementation of special shaping techniques that can be also supported by softswitches and routers.

The reason is to investigate the effect that can be reached without application of the expensive shaping devices using only queues and priorities specific for the IntServ, DiffSerf, and RSVP. This fractional shaping is important in all IP network and can help estimation of the necessity of special shaping device. The simulation model can be also useful to demonstrate where to apply shaping device in the connection, how many devices to apply and how to make their configuration useful for the overall network performance.

II. TRAFFIC SOURCES

Three types of traffic sources are assumed in an example metropolitan area network – Voice over IP, LAN emulation and transaction exchange. Some assumptions are made for every traffic type. In Voice over IP (VoIP) service silence and talk intervals are exponentially distributed with equal mean values. Some authors use talk to silence ratio of ½. Others do prefer to use on-off model for voice service. The limits for waiting times are calculated under consideration of end-to-end delay limits for every service [4], [5], [6], [7]. The same is valid for queue length. Servicing times per packets are fixed due to the typical transmission line requirements. Table I represents all the parameters for traffic sources in the model. Values are taken from the references or approximated.

TABLE I
TRAFFIC SOURCES PARAMETERS

Parameter	VoIP	LAN	Transac-tions
Pear rate, packets per second	10	164	0
Mean call/ session duration, sec	180	20	10
Mean duration between calls/sessions, sec	360	10	15
Mean talk/ silence duration, sec	20	5	2
Distribution of call/series duration	Expo-nential	Expo-nential	Exponential
Maximal waiting time, sec	0.00071 6	0.6	1
Maximal number of waiting packets	210	1804	2
Traffic sources	5000	500	1500
Priorities	High	Mediu m	Low
Packet length, bytes	800	800	800

LAN emulation is specific with its sessions. Sessions are established for any Internet connections. Packet rate is higher in comparison to the VoIP. Session duration is low. The traffic source is behaving as on-off model with exponential duration of the silence and transmission intervals.

Transaction exchange is specific with few packets exchange. The service is not time demanding. Sessions are short.

Number of traffic sources is taken from the typical image in a business area. Packets are taken to be long. In VoIP traffic 800 bytes carry up to 80 milliseconds voice. This means that quality voice can be transmitted only in the area using up to 2-3 hops. End-to-end delay for VoIP traffic usually should not exceed 150 ms. More precise investigation can be done with 200 bytes voice packets.

III. INTEGRATED SERVICES

Integrated Services (IntServ) is a complex technique often called protocol that ensures end-to-end Quality of Service in IP networks.

[1]Rossitza Goleva is an Assistant-Professor in the Department of Telecommunications at Technical University of Sofia, Bulgaria, email: rig@tu-sofia.bg, Member of IEEE

[2]Maria Goleva is a master student on Communication and Information technology at University of Bremen, Germany, email: mgoleva@gmail.com

[3]Dimitar Atamyan is an Assistant-Professor in the Department of Telecommunications at Technical University of Sofia, Bulgaria, email: dka@tu-sofia.bg

[4]Kostadin Golev is a master student in the Department of Telecommunications at Technical University of Sofia, Bulgaria, Email: kotseto@gmail.com

[5]Pavel Merdjanov is an Associate-Professor in the Department of Telecommunications at Technical University of Sofia, Bulgaria, Email: pim@tu-sofia.bg.

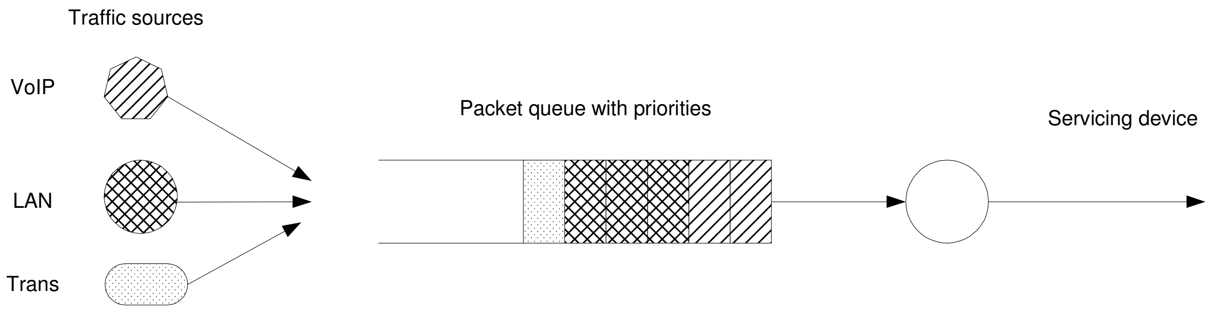


Fig. 1. IntServ model with input data, bounds in waiting times and queue length specific to every service type

IntServ is applied usually in access routers or switches and tried to serve packets from different services in a different ways depending on the quality requirements. IntServ classifies services into three main classes depending on the traffic requirements:

- Elastic applications;
- Tolerant real-time applications;
- Intolerant real-time applications.

Elastic applications are served in a “best effort” way. There is no guarantee for quality level like transaction exchange. Tolerant real-time applications are delay sensitive and usually require high bandwidth. Token bucket model with peak rate control is a good model for such traffic. LAN emulation is usually modeled this way. Some authors propose token bucket with series length and mean rate control as a model for more accuracy. Many others propose the two token buckets to be connected in a cascade configuration.

Intolerant real-time applications require low delays, almost guaranteed bandwidth. The model with two cascaded token bucket is compulsory for such traffic [2]. VoIP service is intolerant to the quality degradation service. IntServ simulation model is based on two cascaded token buckets that bound peak rate, series length and mean rate of the traffic (Fig. 1). The original model is quite complicated and due to this reason it is approximated as a black box that changes the characteristics of the data into output data in specific for IntServ way. As a result after approximation and few calculations it is easy to derive simpler model with one FIFO queue with priorities, fixed rate at the output and different limits for waiting times in the queue [10]. This is the model that has been simulated further. Table II represents main data for model behavior after simply calculations with token bucket formula.

TABLE II
MODEL CHARACTERISTICS IN INTSERV.

Parameter	Value
Queue length, packets	2016
VoIP queue length fraction, packets	210
LAN queue length fraction, packets	1804
Transaction queue length fraction, packets	2
Maximal waiting time for VoIP, sec	0,000716
Maximal waiting time for LAN, sec	0,6
Maximal waiting time for transactions, sec	1
Priority for VoIP	Highest
Priority for LAN	Medium
Priority for transactions	Low

IV. DIFFERENTIATED SERVICES

Differentiated Services (DiffServ) is another quality management technique that is more applicable for core networks. After appropriate marking of the aggregated packets they are gathered in the way that is defined for their class. There are three main types of services we try to highlight in this paper:

- Premium service – low delays, low losses, guaranteed bandwidth like VoIP;
- Assured service – less requirements to the delays and losses in comparison to the premium service like LAN emulation;
- Olympic service – no time requirements at all like transaction exchange.

The model from Fig. 1 that represents IntServ technique is applied again here but with DiffServ procedure in mind. After few calculations again simple FIFO queue with priorities and limits for waiting times per service is derived. Main system parameters are shown on Table III.

TABLE III
MODEL CHARACTERISTICS IN DIFFSERVS.

Parameter	Value
Queue length, packets	1840
VoIP queue length fraction, packets	200
LAN queue length fraction, packets	1640
Transaction queue length fraction, packets	2
Maximal waiting time for VoIP, sec	0,0303
Maximal waiting time for LAN, sec	0,27876
Maximal waiting time for transactions, sec	1
Priority for VoIP	Highest
Priority for LAN	Medium
Priority for transactions	Low

V. RSVP

Resource Reservation Protocol (RSVP) is a technique that is similar to the channel switching reservation. It is especially useful for delay sensitive traffic like VoIP. There are again three main types of services identified for RSVP like:

- Wildcard filter – applied to gather maximal requirements for given interface like LAN emulation;
- Shared explicit – applied to gather maximal requirements for the interface taking into account called address. Transaction exchange is modeled as shared explicit service;
- Fixed filter – full reservation for quality sensitive services like VoIP.

Model simplification with IntServ and DiffServ procedures is applied for RSVP and characteristics of the derived model are shown on Table IV.

TABLE IV
MODEL CHARACTERISTICS IN RSVP.

Parameter	Value
Queue length, packets	1840
VoIP queue length fraction, packets	200
LAN queue length fraction, packets	1640
Transaction queue length fraction, packets	2
Maximal waiting time for VoIP, sec	0,07508
Maximal waiting time for LAN, sec	0,69
Maximal waiting time for transactions, sec	1
Priority for VoIP	Highest
Priority for LAN	Medium
Priority for transactions	Low

VI. RESULTS

Simulation is performed on C++ language. The pseudo exponential pseudo deterministic characteristics of the traffic sources are reached after usage of combination between many random generators. The queue behavior is complex due to the priorities and limits on waiting times. Many parameters have been derived from the model like time and space loss probabilities, probabilities to wait for different types of traffic, probability distribution functions and probability density functions of the packets intervals, queue lengths, waiting times at many interface points in the model like output of the traffic sources, input and output of the queue. Statistical accuracy of the derived results is proven by Student criterion.

IntServ, DiffServ and RSVP have different ways to gather with packets and this influences the way they drop packets and shape them. On Fig. 2 we demonstrate the deterministic character of the traffic at the output of the traffic source. The pdf function of VoIP, LAN and transaction traffic is shown for comparison. The pdf format is chosen for easy comparison with other results.

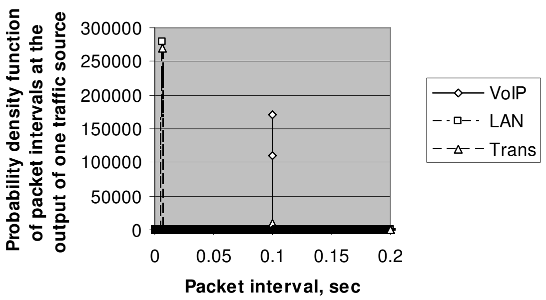


Fig. 2. Probability density function of packet intervals at the output of one traffic source

On Figs. 3, 5, and 7 probability density functions of packet intervals at the input of the queue for IntServ, DiffServ and RSVP is shown. The traffic is accumulated for many deterministic sources like the one shown on Fig. 2. The cumulative traffic has different characteristics and acts as an input traffic for softswitches and routers. It is interesting also for shaping estimation. On Figs. 4, 6, and 8 the pdf of packet intervals at the output of the queue for IntServ, DiffServ and RSVP is shown. The effect of fast servicing in RSVP is visible. The delay variation of the packet intervals is becoming smoother and tends to constant value. Because the VoIP traffic has highest priority and the buffer space for VoIP packets is enough it is visible that the queue acts as a pure multiplexer and shapes the traffic by means of its output packet delivery interval. These interesting results can directly influence interfaces and queue management of the next devices in the network [1]. The three services: VoIP, LAN emulation and transaction exchange are simulated and their queue fraction is also observed and shown in [1].

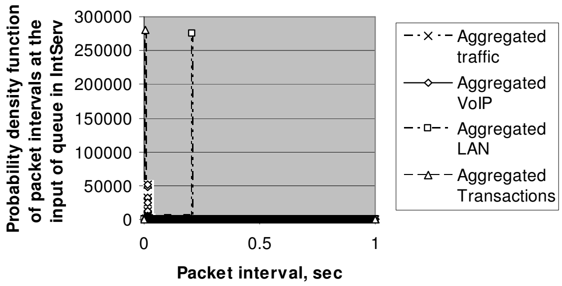


Fig. 3. Probability density function of packet intervals at the input of queue in IntServ

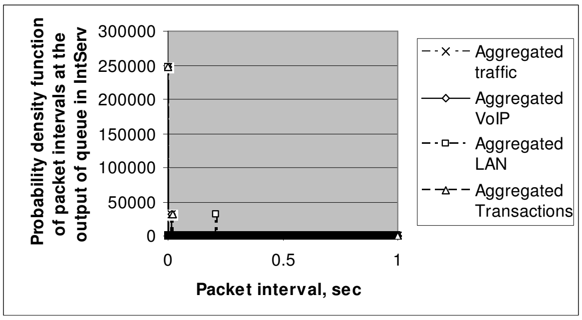


Fig. 4. Probability density function of packet intervals at the output of queue in IntServ

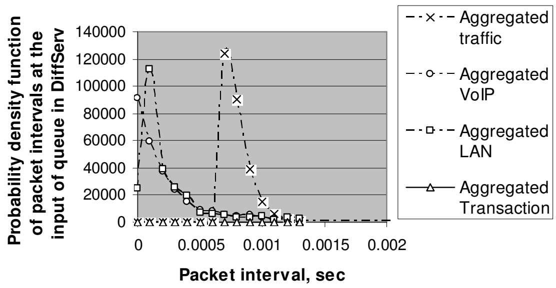


Fig. 5. Probability density function of packet intervals at the input of queue in DiffServ

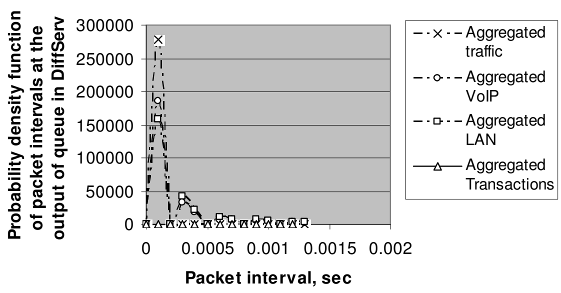


Fig. 6. Probability density function of packet intervals at the output of queue in DiffServ

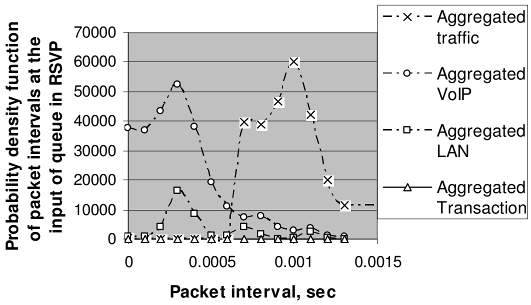


Fig. 7. Probability density function of packet intervals at the input of queue in RSVP

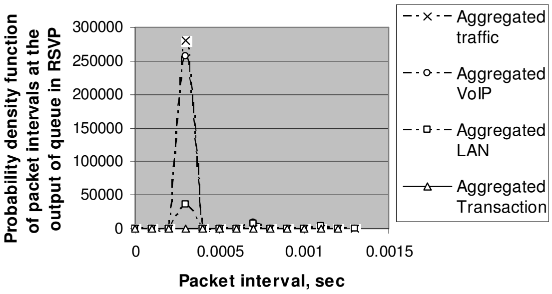


Fig. 8. Probability density function of packet intervals at the output of queue in RSVP

Table V represents mean values of the queue lengths per discipline and per service. They can be used for approximate planning of the time and space limits in the router interfaces.

Because the VoIP traffic has highest priority and the buffer space for VoIP packets is enough it is visible that the queue acts as a pure multiplexer and shapes the traffic by means of its output packet delivery interval. These interesting results can directly influence interfaces and queue management of the next devices in the network [1]. The three services: VoIP, LAN emulation and transaction exchange are simulated and their queue fraction is also observed and shown in [1]. Table V represents mean values of the queue lengths per discipline and per service. They can be used for approximate planning of the time and space limits in the router interfaces.

TABLE V

MEAN QUEUE LENGTH.

Mechanism	Mean queue length, packets	Mean queue length of VoIP fraction, packets	Mean queue length of LAN fraction, packets	Mean queue length of Trans fraction, packets
IntServ	37.98	1	35	2
DiffServ	1586.96	2.24	1583.56	1.98
RSVP	1798.41	156.90	1639.67	2

VII. CONCLUSION

In this paper we show probability density functions of the packet intervals at the output of the traffic sources, at queue input and queue output as well as probability density function of queue length per service type. These results demonstrate the specific characteristics of the queue as a packet shaper in three QoS management algorithms: IntServ, DiffServ, and RSVP. The deterministic nature of the packets streams suppress shaping and increase losses. The statistical multiplexing effect is very limited due to the deterministic streams. Special attention is paid to the VoIP traffic as one of the most delay sensitive traffic.

The results demonstrate the capability of IntServ to define excellent service for VoIP. It is promising in access networks. DiffServ shows excellent resource management and utilization and therefore is better for core services. RSVP is a good counterpart of IntServ in access networks for VoIP. Its future development to NSIS protocol is going to be the most famous and powerful protocol next few years.

The authors refine the simulation model with more traffic sources and more precise generation of the packets from these sources based on the observation of the real traffic. Smaller VoIP packets and variable LAN packet length are under consideration. Limits criteria for queue management are also under investigation.

ACKNOWLEDGEMENTS

This paper is sponsored by the Ministry of Education and Research of the Republic of Bulgaria in the framework of project NFNI BY-TN-105 “Multimedia Packet Switching Networks Planning with Quality of Service and Traffic Management”.

REFERENCES

[1] R. Goleva, M. Goleva, D. Atamyan, K. Golev, P. Merdjanov, *Traffic Shaping Analyses In Metropolitan Area Network*, CSNDSP Conference, University of Patra, Patra, Greece, 2006 (applied paper).

[2] S. Jha, M. Hassan, *Engineering Internet QoS*, Artech House, 2002.

[3] T. Janevski, *Traffic Analysis and Design of Wireless IP Networks*, Artech House, 2003.

[4] L. Kleinrock, *Queueing Systems, Volume I: Theory,* John Wiley and Sons, 1976, ISBN 0-471-49111-X (v. 1).

[5] L. Kleinrock, *Queueing Systems, Volume II: Computer Application*, John Wiley and Sons, 1976, ISBN 0-471-49111-X (v. 2).

[6] V. Iversen, *Teletraffic Engineering Handbook*, ITU-D, 2005.

[7] S. Lavenberg, Editor, *Computer Performance Modeling Handbook*, Academic Press, 1983, ISBN 0-12-438720-9.

[8] J. Pitts, J. Schormans, *Introduction to IP and ATM Design and Performance*, John Wiley&Sons, Ltd., 2000.

[9] V. Ralsanen, *Implementing Service Quality in IP Networks*, John Wiley & Sons, Ltd., 2003.

[10] A. Tanenbaum, *Computer Networks, Second Edition*, Prentice-Hall International, Inc., 1989, ISBN 0-13-166836-6

Formation of Signals Harmonized with a Linear Channel of Connection with Limitation of Their Average Rectified Values

Antonio V. Andonov[1] and Galina P. Cherneva[2]

Abstract: **The paper presents a solution of the problem of synthesizing harmonized signals with the limitation of their average values and a report on the efficiency of their transmission along a linear channel of connection.**
Keywords: **signals harmonized with the channel, functional of synthesis**

I. INTRODUCTION

The problem of forming signals harmonized with the channel of connection is a variation problem connected with finding the conditional extreme of the quality functional of synthesis, i.e. with certain limitations imposed on the signals. They have determined closed area S*, where the signals searched are in its interior. The kind of the functional has been defined by the criteria of quality and the nature of the noise in the channel while the non-functional limitations are defined by the properties of the transmitter. The signals, which have been imposed certain limitation on, are same in structure and united in one class despite the kind of the criteria of quality.

The paper presents a solution of the problem of synthesizing harmonized signals with the limitation of their average values and a report on the efficiency of their transmission along a linear channel of connection.

II. FORMULATION OF THE PROBLEM OF HARMONIZING WITH THE LIMITATION OF THE AVERAGE RECTIFIED VALUES OF THE INPUT SIGNAL FOR THE CHANNEL

The most common kind of the limitations on the signals searched is:

$$m_1\left\{\int_{\tau_H}^{\tau_K}\varphi_j\left[x_j(t), s_j(t), \underline{V}(t), t\right]dt\right\} \le S_j \text{ j=1,2,…m,} \tag{1}$$

where: $s_j(t)$ – input signal;
$x_j(t)$ – output signal;
$\underline{V}$ - vector function of the disturbing signals.

However, the most spread limitation in practice is that of the average value of p-*th* order of the signal being optimized:

$$\left[\frac{1}{\Delta T}\int_{t_H}^{t_K}\left|s_j(t)\right|^p dt\right] \le S_p\text{, } \Delta T=t_K-t_H\text{, j=1,2,…,m} \tag{2}$$

When p=1 is satisfied in (2), it concerns the limitation of the average rectified value of the signal. Such a limitation can be observed with describing movable radio means where the source used to supply the transmitting device is overcharged earlier than the output series of the transmitter.

Then the problem of signal harmonization can be formulated in the following way: within the class of signals $L_1[t_H, t_K]$ determined by limitations of (2), with (p=1), to find those signals $\{s_i(t);\ i=1,…,\ m,\ t\in[t_H, t_K]\}$ that maximize the functional for:

$$I=\varphi^{-1}\sum_{i=1}^{m}\sum_{j=1}^{m}a_{ij}\varphi(r_{ij}) \tag{3}$$

Here r_{ij} is the distance between the i-th and j- th signals at the output of the channel of connection. The distance for a standardized metric linear area with the limitation of the average rectified value of the signal [1] is determined as:

$$r_{ij}=\frac{1}{\Delta\tau}\int_{\tau_H}^{\tau_K}\left|x_i(t)-x_j(t)\right| dt\text{ ,i,j=1,2,…m,} \tag{4}$$

$\Delta\tau=\tau_K-\tau_H$ is the interval of the observation on the output signal.

Substituting (4) into (3) and expressing the output signal by the input one, the functional, which has to be maximized, takes the kind of:

$$I=\int_{t_H}^{t_K}\ldots\int_{t_H}^{t_K}s(t_1)\ldots s(t_v)H_v(t_1\ldots t_v)dt_1\ldots dt_v \tag{5}$$

where H_v is the nucleus of the functional and depends on the channel pulse characteristics.

III. SOLUTION OF THE OPTIMIZATION PROBLEM WITH LIMITING THE AVERAGE RECTIFIED VALUE

From the mathematical point of view, in the case of limiting the average rectified value of the signal, it is most

[1] Antonio V. Andonov is with the Higher School of Transport, Geo Milev Str.158, 1584 Sofia, Bulgaria, E-mail: andonov@vtu.bg
[2] Galina P. Cherneva is with the Higher School of Transport, Geo Milev Str.158, 1584 Sofia, Bulgaria, E-mail: galja_cherneva@abv.bg

convenient to use the inequality of Haldle [2] to work out the conditions necessary and sufficient for the maximum of functional (5). It can be proved, according to him, that in order to reach a maximum (5), signals s(t) have to satisfy the following condition:

$$s(t_1)=\sum_{i=1}^{N} a_i\delta\left(t_i - t_{i\max_x}\right) signF(t_1) \tag{6}$$

where :

$$F(t_1)=\left|\int_{t_H}^{t_K}\ldots\int_{t_H}^{t_K} s(t_2)\ldots s(t_v)H_v(t_2\ldots t_v)dt_2\ldots dt_v\right| \tag{7}$$

– a_i are coefficients that are determined by the condition of the signal norm in $L_1[t_H, t_K]$:

$$\frac{1}{\Delta T}\int_{t_H}^{t_K}\left|s\ (t_1)\right|\ dt = S \tag{8}$$

– $\{t_{i\max}, i=1\ldots N\}\equiv\{\Theta\}$ is a set of moments where function $F(t_1)$ reaches a maximum.

Having substituted (6) into (8) and taking into consideration the filtrating property of δ - function, it is obtained that:

$$\sum_{i=1}^{N} a_i = \Delta T S_1 \tag{9}$$

Hence, the harmonized signals of class $L_1[t_H, t_K]$ with limited average rectified values that are determined by expression (6) present a sequence of δ -pulses supplied at the channel input at moment $\{t_{i\max}\}$, when function (7) depending on the channel pulse characteristics has a maximum.

IV. Determination of the Efficiency of Transmitting Signals Harmonized with a Linear Gauss's Channel and with Limited Average Rectified Values.

The linear cable equivalent to a low-pass filter is an appropriate model of some real channels of connection such as cable and hydro-acoustic ones. For the simplest case of one-unit low-pass filter, the pulse characteristic is in the kind of:

$$h(t)=\alpha . e^{-\alpha . t} \tag{10}$$

where α is the damping coefficient of the channel.

The efficiency of transmitting signals harmonized with the channel will be assessed comparing them to signals in the kind of rectangular pulses transmitted along the same channel. That can be done by the coefficient of efficiency expressed by the ratio of the coefficients with transmitting signals of both types:

$$\gamma^2 = \frac{K^2_{onm}}{K^2_{np}} \tag{11}$$

and showing what is the relation between the power of the signals at the channel output and the characteristic (10) in the both cases.

To deduce the dependency for (11), the ratio of the norms of the input and output signals has been used as an analog to the coefficient of transmitting harmonized signals of limited average values [2]. If the case of Gauss's noise in the channel is examined when power serves to measure the distance between two signals and the pulse characteristic is substituted with expression (10), it will be obtained that after the transformation:

$$K^2_{onm} = \begin{cases} \alpha\dfrac{\Delta T^2}{2\Delta\tau}e^{2\alpha . t}\left(e^{-2\alpha\tau_H} - e^{-2\alpha\tau_K}\right)\ , t\le\tau_H ; \\ \alpha\dfrac{\Delta T^2}{2\Delta\tau}\left(1-e^{-2\alpha(\tau_K - t)}\right)\ , t\succ\tau_H \end{cases} \tag{12}$$

where t is the moment of supplying δ -pulse at the input of the channel of connection.

The dependencies $K^2_{\text{опт}}=f(t)$ in the both time intervals for a linear channel in the kind of a low-pass filter are given in Fig. 1 and Fig. 2.

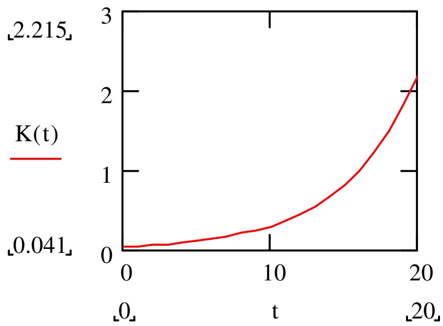


Fig. 1. Dependency $K_{\text{опт}}=f(t)$ for $t \le \tau_H$

To find out the best moment $t= t_{max}$, the dependency $K^2_{\text{опт}}=f(t)$ has been examined for a maximum using "Mathcad". It has been obtained that this moment is $t =\tau_H$.

Substituting (12) in (11), the following expression of the efficiency coefficient has been obtained:

$$\gamma^2 = \frac{K^2_{onm}}{K^2_{np}} = \frac{\alpha\Delta T\left(1-e^{-2\alpha\Delta T}\right)}{2\left(1-\dfrac{3}{2\alpha\Delta T}+\dfrac{2e^{-\alpha\Delta T}}{\alpha\Delta T}-\dfrac{e^{-2\alpha\Delta T}}{2\alpha\Delta T}\right)}$$

The graphic dependency $\gamma^2= f(\alpha\Delta T)$ is given in Fig. 3.

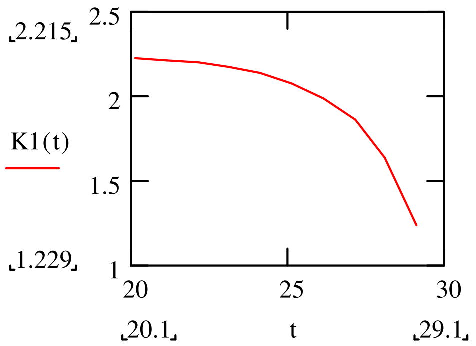


Fig. 2. Dependency $K_{опт}=f(t)$ for $t>\tau_H$

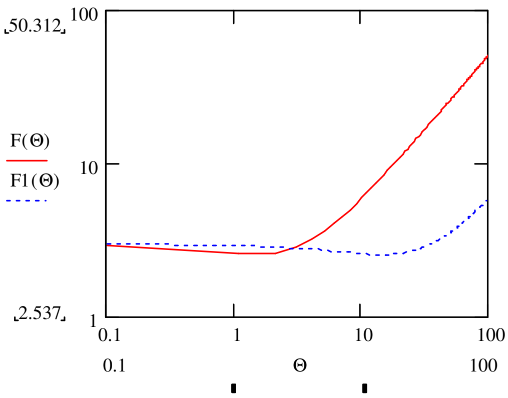


Fig. 3. Dependency $\gamma^2 = f(\alpha\Delta T)$

It is seen that the power of a signal in the kind of a rectangular pulse (dashed line), of duration of $\tau_и=0{,}1\Delta T$, is lower at the channel output than that with transmitting a signal harmonized with the channel and of (dense line). That can be explained by the fact that the limitation of the average rectified value does not result in limiting the power. That is why it is expedient to use those signals in the systems of movable radio connection.

V. Conclusions

A considerable efficiency has been obtained with transmitting signals harmonized with the channel and of a limited average rectified values. It is so because the harmonized signals of the class under examination, $L_1[t_H, t_K]$, are of wide-frequency band (short duration) and therefore their base is B>>1. As it can be seen in Fig.3, the biggest efficiency of transmitting harmonized signals is with the values of $\alpha\Delta T$ commensiorable with their base. Hence the transmission of signals in the kind of δ -pulses, such as the harmonized ones with limited average rectified values, results in the increase of their base and in higher power at the channel output.

References

[1] Cherneva G., Setting the Problem of Synthesis of Signals Coordinated with the Cannel of Connection, Mechanics, Transport, Communications 2/2004

[2] HutsonV.C.L.,J.S.Pym, Applications of Functional Analysis and Operator Theory, London 1980

Tools for Calculating Autocorrelation Spectrum by Using The Wiener-Khinchin Theorem

Miloš M. Radmanović

Abstract — **The paper consider tools for calculating autocorrelation spectrum by using the Wiener-Kinchin theorem. The first tool calculates autocorrelation spectrum implemented over matrices, the second tool use fast Walsh transform over matrices and the third tool calculates autocorrelation spectrum implemented over decision diagrams. Then, I discussed "LGSynth93 benchmark" testing statistics for each tool and tool's efficiency. Also, I analyzed time and space complexity for each tool, especially for tool which use calculations over the decision diagrams.**

Keywords — **Switching functions, Autocorrelation spectrum, Wiener-Kinchin theorem, decision diagrams, software tools..**

I. Introduction

Much work has been performed in applying transforms to switching functions in order to achieve a more global view of the function. Transforms such as the Walsh and their applications in digital logic are well researched [6]. Also, there is a lot of software support. There is far less work, however, on the use of other transforms such as the autocorrelation transform [7].

An alternative view of the switching function is the autocorrelation spectrum (all autocorrelation coefficients) of a function. The autocorrelation coefficients of a function are calculated using the autocorrelation function. They provide the measure of the function's similarity to itself, shifted by given amount, Eq. (1).

The autocorrelation coefficients have been used in various areas including optimization and synthesis of combinational logic [1], [4], [5], variable ordering for decision diagrams [2] and compute the estimate $C(f)$ of a function's complexity [3], [1].

However, their use has been limited, likely due to the fact that until recently, methods for computing the autocorrelation coefficients were exponential in the number of inputs to the function(s). Since new methods for their computation have recently been introduced by Rice [8], [9], and by Stanković [10], I also have performed an investigation into development of software tools for calculating autocorrelation coefficients (spectrum) for Boolean function

In this paper I present the definition, explanation and software tools for calculating autocorrelation spectrum by using the Wiener-Kinchin theorem. The first tool calculates autocorrelation spectrum through matrix multiplication, the second tool use fast Walsh transform and the third tool calculates autocorrelation spectrum through decision diagrams. Then, I discussed "LGSynth93 benchmark" [11] testing space and time complexity statistics for each tool and tool's efficiency.

II. Autocorrelation of Switching Functions

The general cross-correlation (convolution) between two functions f and g at a distance τ is defined as:

$$\sum_{x=0}^{2^n-1} f(x)\cdot g(x\oplus\tau) \tag{1}$$

for two functions $f(X)$ and $g(X)$ where $X = x_n x_{n-1}\dots x_1$. The symbol $\oplus$ represents the bitwise exclusive-or function and the symbol $\cdot$ represents arithmetic multiplication function. When $f = g$, the resulting equation gives the cross-correlation of the function with itself, translated by τ. The resulting coefficients are referred to as the autocorrelation coefficients of the function

The autocorrelation function is defined as follows:

$$B(\tau) = \sum_{x=0}^{2^n-1} f(x)\cdot f(x\oplus\tau) \tag{2}$$

where function is $f(X)$, $X = x_n x_{n-1}\dots x_1$, $x = \sum_{i=1}^{n} x_i 2^{i-1}$, $\tau = \sum_{i=1}^{n} \tau_i 2^{i-1}$ and n is number of inputs [3].

For multiple output functions a second step must be performed to combine the autocorrelation function for each of the individual functions into the total autocorrelation function is defined as:

$$B(\tau) = \sum_{i=0}^{m-1} B_i(m) = \sum_{i=0}^{m-1}\sum_{x=0}^{2^n-1} f(x)\cdot f(x\oplus\tau) \tag{3}$$

where m is the number of outputs and the multiple output function F consists of $f_0 f_1 \dots f_{m-1}$.

III. Wiener-Khinchin Theorem

The Wiener-Khinchin theorem states a relationship between the autocorrelation function and the Walsh (Fourier) coefficients [12].

$$B_f = 2^{-n} W^{-1}(Wf)^2 \tag{4}$$

[1] Miloš M. Radmanović is with the Faculty of Electronics, University of Niš, Aleksandra lmedvedova 14, 18000 Niš, Serbia and Montenegro, e-mail: milos@elfak.ni.ac.yu

where W denote the Walsh transform operator and W^{-1} denote the inverse Walsh transform operator.

An example of calculating the autocorrelation spectrum using Wiener-Kinchin theorem through matrix multiplication for the function $f(x_1, x_2) = x_1 + x_2$ is shown in the following equations:

$$B_f = 2^{-2} W(2)(W(2)F)^2 \tag{5}$$

$$B_f = \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 \end{bmatrix}\left(\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}\right)^2 = \begin{bmatrix} 3 \\ 2 \\ 2 \\ 2 \end{bmatrix} \tag{6}$$

Calculations in determination of the Walsh spectrum can be performed through the flow-graph describing the fast Walsh transform [13]. An example of calculating the autocorrelation spectrum using Wiener-Kinchin theorem through fast Walsh transform for the function $f(x_1, x_2) = x_1 + x_2$ is shown in the following equations:

$$\begin{matrix} 0 \\ 1 \\ 1 \\ 1 \end{matrix} \xrightarrow{\substack{fast \\ Walsh \\ transform}} \begin{matrix} 3 \\ -1 \\ 1 \\ 1 \end{matrix} \xrightarrow{sqare} \begin{matrix} 9 \\ 1 \\ 1 \\ 1 \end{matrix} \xrightarrow{\substack{fast \\ Walsh \\ transform}} \begin{matrix} 12 \\ 8 \\ 8 \\ 8 \end{matrix} \xrightarrow{1/4} \begin{matrix} 3 \\ 2 \\ 2 \\ 2 \end{matrix} \tag{7}$$

Calculations in determination of the Walsh spectrum can be performed through fast Walsh transform over decision diagram [10]. An example of calculating the autocorrelation spectrum using Wiener-Kinchin theorem through fast Walsh transform and decision diagrams for the function $f(x_1, x_2) = x_1 + x_2$ is shown in the following figure:

IV. Software Tools

My applications (software tools) are written in Microsoft Visual Studio [14] and use MFC technology.

The first software tool for calculating the autocorrelation spectrum using Wiener-Kinchin theorem through matrix multiplication uses dynamic arrays of arrays as data structure for storing true vector and Walsh transform matrix. Because autocorrelation coefficient can be number between 0 and $2^n - 1$, it is used 64-bit integer data type for storing autocorrelation spectrum. Using calculating the autocorrelation spectrum by Wiener-Kinchin theorem for a function requires $O(2^{2n})$ operations (multiplication and summing) on dynamic arrays, where n is number of inputs Eq. (5) and (6).

Using another data structures and methods requires significantly less operations for calculating the autocorrelation spectrum

The second software tool for calculating the autocorrelation spectrum using Wiener-Kinchin theorem through fast Walsh transform over vectors uses dynamic arrays as data structure for storing true vector. This tool uses much less memory then first software tool. The tool uses only one 64-bit integer array and it don't produce any temporary arrays. Using calculating the autocorrelation spectrum by Kinchin theorem for a function requires $O(n2^n)$ operations (multiplication and summing) on dynamic arrays, where n is number of inputs Eq. (7). Using decision diagrams as data structure and DD method requires significantly less operations in some cases of calculating the autocorrelation spectrum.

The third software tool for calculating the autocorrelation spectrum using Wiener-Kinchin theorem through fast Walsh transform over decision diagrams uses special case of linked lists as data structure. The tool was develop following basic programming principles for DD packages [15]. Every DD package use an imperative programming language like C++, nodes are class structures that contain a variable identifier and "then" and "else" children pointers; a "next" pointer strings nodes together that belong to the same collision chain in the unique table, recycling of nodes is easily implemented by keeping a reference count for each node. Using calculating the autocorrelation spectrum by Kinchin theorem for a function requires $O(i) + O(j) + O(k) + O(l)$ operations (multiplication and summing linked lists) on special case of linked lists, where i is number of nodes of initial DD, l is number of nodes of temporary DD (after first Walsh transform), k is number of nodes of temporary DD (after squaring) and l is number of nodes of resulting DD (after second Walsh transform), (see figure 1).

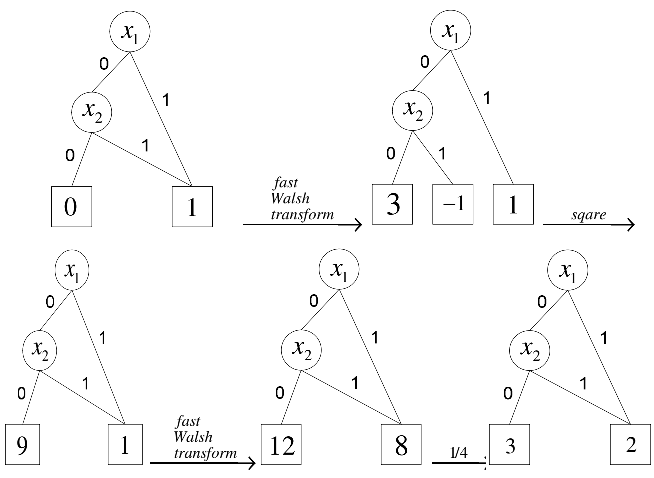


Fig. 1. An example of calculating the autocorrelation spectrum using Wiener-Kinchin theorem through fast Walsh transform and decision diagrams

V. Tests and Results

Below I give a lists and tables of different tools testing statistics. I performed the testing on a PC Pentium IV on 1,4 GHz with 224 MB of RAM (MS Windows XP Professional 2002). The memory usage for all tools was limited to 150 MB, and timing statistics do not include building and storing data structures for functions (arrays of arrays, arrays and linked lists). I have tasted tools with "LGSynth93 benchmark", and I used input files described in ESPRESSO-MV (or pla) format [16].

LGSynth93 benchmark suites is based on the collection of benchmarks from the ISCAS85, ISCAS86 and LGSynth91. testing space and time complexity statistics for each tool and tool's efficiency. The benchmarks then have been categorized in three categories: small benchmarks, medium sized benchmarks (and too-large benchmarks. This allows to judge the quality of tools and gives a better overview of the existing benchmarks. Descriptions and properties of switching functions from LGSynth benchmark suite are given in tables 1, 2 and 3.

Table 4 describes tool efficiency statistics and it shows that most efficient tool is third tool that use fast Walsh transform through array. But, for too-large benchmarks, neither tool is efficient (25% isn't enough).

TABLE I

LGSYNTH93 SMALL BENCHMARKS

Name	Number of inputs	Number. of outputs	Number of cubes
xor5	5	1	16
rd53	5	3	32
squar5	5	8	32
bw	5	28	87
con1	7	2	9
rd73	7	3	141
inc	7	9	34
5xp1	7	10	75
sqrt8	8	4	40
rd84	8	4	256
misex1	8	7	32
9sym	9	1	87
clip	9	5	167
apex4	9	19	438
sao2	10	4	58
ex1010	10	10	1024
alu4	14	8	1028
table3	14	14	175
misex3c	14	14	305
misex3	14	14	1848
b12	15	9	431
t481	16	1	481
pdc	16	40	2810
spla	16	46	2307
table5	17	15	158

Table 5 describes average time over all benchmarks and it shows that best average time for small and medium sized benchmark has third tool. It is expected, because the calculation time is direct proportional to tool's space (memory) request.

TABLE II

LGSYNTH93 MEDIUM SIZED BENCHMARKS

Name	Number of inputs	Number. of outputs	Number of cubes
duke2	22	29	87
cordic	23	2	1206
cps	24	109	654
vg2	25	8	110
misex2	25	18	29

TABLE III

LGSYNTH93 TOO-LARGE BENCHMARKS

Name	Number of inputs	Number. of outputs	Number of cubes
apex2	39	3	1035
seq	41	35	1459
apex1	45	45	206
apex3	54	50	280
e64	65	65	65
apex5	117	88	1227
ex4p	128	28	620
o64	130	1	65

TABLE IV

TOOL EFFICIENCY STATISTICS

Tool	Small benchmarks	Medium sized benchmarks	Too-Large banchmarks
first. tool	87%	0%	0%
second tool	100%	60%	0%
third tool	100%	100%	25%

TABLE V

AVERAGE TIME STATISTICS

Tool	Small benchmarks	Medium sized benchmarks	Too-Large banchmarks
first. tool	46.863	-	-
second tool	1.667	740.797	-
third tool	1.273	0.741	66.561

TABLE VI

MAX SPACE STATISTICS (ELEMENTS OF ARRAY)

Tool	Small benchmarks	Medium sized benchmarks	Too-Large banchmarks
first. tool	~2^34	~2^50	~2^130
second tool	~2^17	~2^25	~2^65

Table 6 describes tool maximal space statistics (number of elements in arrays) and it shows that first tool for medium sized benchmark require almost 2^50*32 bit = 4096 TB and for too-large benchmark 268435456 TB of memory. It is expected; because the space is exponential proportional to function's number of inputs.

Table 7 describes third tool space statistics (number of decision diagram nodes) and shows that we can not calculate space limit and dependency. In most cases, it is shown that $i \sim j \sim k \sim l$, but functions like *cps* ond *apex5* shows that $i << j$. Functions: *alu4*, *table3*, *misex3c* and *misex3* shows that $i << l$. Meanwhile, all functions shown that $j \sim k$. If it is possible. to calculate Walsh spectrum over DD, there is high probability for calculation of autocorrelation spectrum.

TABLE VII

SPACE STATISTICS (NUMBER OF DD NODES) FOR THIRD TOOL

Name	**i** (initial DD.)	**j** (after 1. Walsh T..)	**k** (after . squaring.)	**l** (after 2. Walsh T.)
xor5	9	5	5	9
rd53	23	34	29	29
squar5	38	83	56	58
bw	114	330	236	253
con1	18	76	50	30
rd73	43	57	44	51
inc	89	371	280	209
5xp1	88	297	208	237
sqrt8	42	123	75	126
rd84	59	88	48	55
misex1	47	276	159	88
9sym	33	39	39	37
clip	254	529	326	539
apex4	1021	4836	3979	4595
sao2	154	481	394	472
ex1010	1079	6281	5292	5696
alu4	1352	6195	4351	15364
table3	941	41652	33660	40295
misex3c	847	9189	6851	16077
misex3	1301	18364	13009	55520
b12	91	651	492	151
t481	32	184	60	60
pdc	705	10717	6457	3826
spla	681	8785	5480	3545
table5	873	89589	75130	48959
duke2	976	6581	4590	5273
cordic	80	646	491	463
cps	2318	15420	8292	5339
vg2	1059	3793	3096	6563
misex2	140	1311	844	139
apex2	7102	-	-	-
seq	142321	-	-	-
apex1	28414	-	-	-
apex3	-	-	-	-
e64	1446	6853	5473	2144
apex5	2705	74258	31269	4634
ex4p	1301	-	-	-
o64	-	-	-	-

VI. CONCLUSION

In this paper I present software tools for calculating autocorrelation spectrum by using the Wiener-Kinchin theorem. The first tool calculates autocorrelation spectrum through matrix multiplication, the second tool use fast Walsh transform and the third tool calculates autocorrelation spectrum through decision diagrams. Then, I presented "LGSynth93 benchmark" testing statistics (efficiency, time, space) for each tool. Third DD-based tool has best results over all benchmarks, but for too-large benchmarks, tool require further work to optimize in terms of memory.

ACKNOWLEDGMENT

This research work was based on a scholarship of The German Academic Exchange Service (DAAD) under the Stability Pact for South-East Europe.

REFERENCES

[1] Tomczuk, R., "Autocorrelation and Decomposition Methods in Combinational Logic Design", 1996, PhD dissertation.

[2] Rice, J.E., Serra, M., Muzio, J.C., "The Use of Autocorrelation Coefficient for Variable Ordering for ROBDDs", Proc. 4th Int. Workshop on Applications of Reed-Muller Expansion in Circuit Design, Victoria, Canada, August 20-21, 1999, 185-196.

[3] Karpovsky M.G., "Finite Orthogonal Series in the Design of Digital Devices", John Wiley 1976.

[4] Ruce, J.E., Muzio, J.C., "The Use of Autocorrelation Function in Classification of Switching Functions", In. Euromicro Symposium on Digital System Design: Architectures, Methods and Tools (DSD), pages 244-251, 2002.

[5] Ruce, J.E., Muzio, J.C., "On The Use of Autocorrelation Coefficients in The Identification of Three-level Decompositions", Proc. Int. Workshop on Logic Synthesis, (IWLS 2003), 2003.

[6] S. L. Hurst, D. M. Miller, and J. C. Muzio, "Spectral Techniques in Digital Logic", Orlando, Florida: Academic Press, Inc., 1985.

[7] Rice, J.E., Muzio, J.C., "Properties of Autocorrelation Coefficients", Proc. IEEE Pacific Rim Conf. on Communications, Computers and Signal Processing, 2003.

[8] Rice, J.E., Muzio, J.C., "Methods for Calculating Autocorrelation Coefficients", In Proceedings of the 4th Workshop on Boolean problems, pages 69-76, 2000.

[9] Rice, J.E. "Autocorrelation Coefficients in the Represen-tation and Classification of Switching Functions", PhD thesis, University of Victoria, 2003.

[10] Stanković, R., S., Bhattacharaya, M., Astola, J., T., "Calulation of Dyadic Autocorrelation through Decision Diagrams", Proc. European Conf. Circuit Theory and Design, pp. 337-340, 2001.

[11] "The LGSynth93 Benchmark Suit", Available: http://www.bdd-portal.org/benchmarks/LGSynth93.tar.gz.

[12] Pichler F., "Walsh Functions and Linear System Theory", Proc. Applic. Walsh Functions, Washington, D.C., 1970, 175-182.

[13] Clarke, E., M., McMillan, K.L., Zhao, X., Fujita, M., "Spectral Transforms for Extremely Large Boolean Functions in Kebaschull", U., Schubert, E., Rosenstiel, W., Eds., Proc. IFIP WG 10.5 Workshop on App. of the RM Expression in circuit Design, Humburg, Germany, September 16-17, 1993, 86-90.

[14] Microsoft Corporation, Available: www.microsoft.com

[15] Jansen G. "Design of Pointerless BDD Package" inite", 10th Worshop on Logic and Synthesis Granlibakken, Lake Tahoe, CA, 12-15, 2001.

[16] Lisanke, R., "Logic Synthesis and Optimization Benchmarks User Guide" Version 2.0, 1988.

Dual Channel Quadrature Mirror Filter Bank Containing Sigma-Delta Modulators

Jaroslav Vrána[1] and Tomáš Lukl[2]

Abstract – **Quadrature mirror filter bank is used in digital signal processing. Most often it is realized by FIR filters. The signals are decomposed into some subbands by quadrature mirror filter banks. These subbands are further processed. For example it can be stored with various bit accuracy in subband coding. In this paper a structure are described which can be resistant to distortion of subband signals.**

Keywords – **filter bank, sigma-delta modulator, quantization, noise shaping**

I. INTRODUCTION

In subband coding can be used quadrature mirror filter banks. In this case first the input signal is decomposed into some subbands. In individual subbands the signals are processed. The result of their processing is mostly their distortion. In reconstruction part the output signal is reconstructed from individual subbands. Because the individual subbands can be distorted the output signal also can be distorted. But the distortions of subband signal can be affected in the case of sigma-delta modulator are used in decomposition part. Sigma-delta modulator has different transfer functions for input signal and error signal. Therefore the error signal can be adjusted that it should be more suppressed by passing through reconstruction part.

II. QUADRATURE MIRROR FILTER BANK

Quadrature mirror filter bank decomposes input signal into some subbands. In this paper it is used only dual channel quadrature mirror filter bank. The block scheme of this filter bank can be seen in Fig. 1.

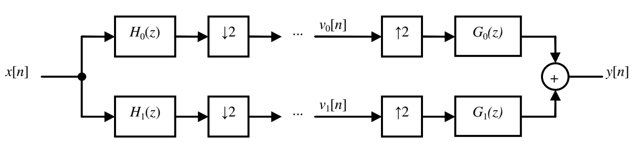


Fig. 1. Block scheme of dual channel filter bank

Input signal is processed in two branches. First branch is low part of half band mirror filter and second branch is high part of half band mirror filter. Input signal enters into low pass filter $H_0(z)$ and high pass filter $H_1(z)$ and thereby it is split in two subbands. Because these parts are frequency limited, the sampling frequency can be decreased. This part is decomposition part. For reconstruction of original input signal at first the sampling frequency must be increased and then the signals are filtered by low pass filter $G_0(z)$ and high pass filter $G_1(z)$ again. This part is reconstruction part. So as the signal is not modified by this structure the Eq. (1) a (2) must be fulfilled.

$$H_0(z)\cdot G_0(z)+H_1(z)\cdot G_1(z)=2z^{-k} \tag{1}$$

$$H_0(-z)\cdot G_0(z)+H_1(-z)\cdot G_1(z)=0 \tag{2}$$

First Eq. (1) describes the passage of input signal through this structure. The signal can be delayed. Second Eq. (2) describes passage of mirror spectrums through this structure, which originates by change of sampling frequency. The most simple filter bank has filters with transfer functions Eqs. (3), (4), (5) a (6).

$$H_0(z)=\frac{1}{2}+\frac{1}{2}z^{-1} \tag{3}$$

$$H_1(z)=-\frac{1}{2}+\frac{1}{2}z^{-1} \tag{4}$$

$$G_0(z)=\frac{1}{2}+\frac{1}{2}z^{-1} \tag{5}$$

$$G_1(z)=\frac{1}{2}-\frac{1}{2}z^{-1}, \tag{6}$$

The processing of signals in individual branches can be expressed as adding error signals to individual subband signals. This situation can be seen in Fig. 2. These error signals pass through filters in reconstruction part of filter bank and they are summed on output. If error signals are white noises with constant power, the power spectrums are summed on output.

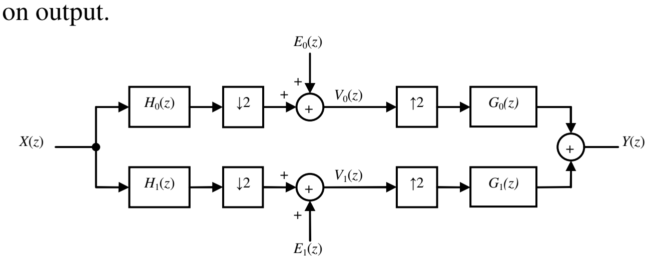


Fig. 2. Block scheme of dual channel filter bank

[1] Jaroslav Vrána is with the Faculty of Electronics and Communication Technologies, Brno Technical University, Purkyňova 118, 612 00 Brno, E-mail:xvrana02@stud.feec.vutbr.cz

[2] Tomáš Lukl is with the Faculty of Electronics and Communication Technologies, Brno Technical University, Purkyňova 118, 612 00 Brno, E-mail: thomlu@eabrno.cz

III. SIGMA-DELTA MODULATION

Sigma-delta modulators are most often used in AD and DA converters. One of their most advantages is quantization noise shaping. It means that quantization noise is not uniform over whole frequency spectrum, but it is more suppressed in lower frequencies and less suppressed in higher frequencies. In detail it is described in [1]. Block scheme of sigma-delta modulator can be seen in Fig. 3.

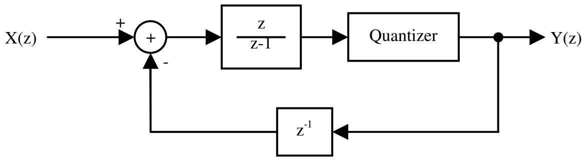


Fig. 3. Sigma-delta modulator

A generalized sigma-delta modulator can be used in filter bank. Digital integrator is replaced by block with transfer function $K(z)$ and block with transfer function $J(z)$ is inserted into feed back. Block scheme of linear model can be seen in Fig. 4.

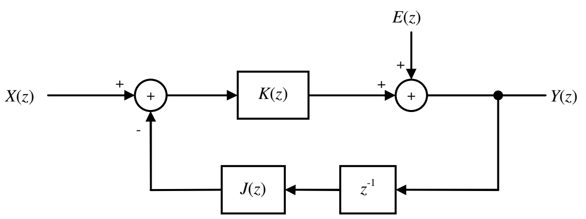


Fig. 4. Linear model of sigma-delta modulator

The input signal enters into summer together with feed back signal. The differential signal pass through block with transfer function $K(z)$. Behind this block a error signal $E(z)$ is added to signal. In AD converter the error signal is quantization noise. Because the quantizer is nonlinear part and the analysis of circuits with this nonlinear part is difficult, the quantizer is replaced by sum with error signal. Thereby the analysis is simpler. The sum of signal and error signal is output signal of generalized sigma-delta modulator. In sigma-delta modulator is also feed back. The output signal is delayed by one sample and then it pass through block with transfer function $J(z)$. This signal is subtracted by input signal. The output signal of this block scheme can be computed by equation:

$$Y(z) = X(z) \cdot \frac{K(z)}{1 + z^{-1} \cdot J(z) \cdot K(z)} + E(z) \cdot \frac{1}{1 + z^{-1} \cdot J(z) \cdot K(z)}. \quad (7)$$

The first term describes passage of useful signal and the second term describes passage of error signal (quantization noise). So the frequency characteristic corresponds to characteristic of filter in filter bank, the characteristics must be same. Therefore Eq. (8) must be fulfilled, which describes equality between transfer function of sigma-delta modulator and required transfer function $H(z)$.

$$H(z) = \frac{K(z)}{1 + z^{-1} \cdot J(z) \cdot K(z)} \quad (8)$$

For realization of sigma-delta modulator it is necessary also to compute transfer function $J(z)$. The transfer function $J(z)$ can be computed by Eq. (8).

$$J(z) = \frac{K(z) - H(z)}{z^{-1} \cdot K(z) \cdot H(z)} \quad (9)$$

By substituting Eqs. (8) and (9) into Eq. (7) the transfer function of sigma-delta modulator is:

$$Y(z) = X(z) \cdot H(z) + E(z) \cdot \frac{H(z)}{K(z)}. \quad (10)$$

From this equation can be seen that the transfer function of input signal is exactly $H(z)$ and sigma-delta modulator has the required transfer function. Transfer function of error signal depends also on transfer function of input signal, but it is divided by transfer function $K(z)$, which can be chosen. By its suitable chose the case can happened that error signal will be more transferred in stop band of reconstruction filter bank and it will be more suppressed in pass band of reconstruction filter bank.

IV. FILTER BANK CONTAINING SIGMA-DELTA MODULATORS

The sigma-delta modulator with general transfer can be used in quadrature mirror filter bank. The advantage of replacement of decomposition filters by sigma-delta modulators is better suppression of subband signals distortion. For example in subband coding the individual coefficients can be stored with lower bit resolution and thereby it produces higher quantization noise into signals which is subsequently suppressed.

Block scheme of dual channel filter bank which use sigma-delta modulators in decomposition part can be seen in Fig. 5. In this structure subsampling is not used yet.

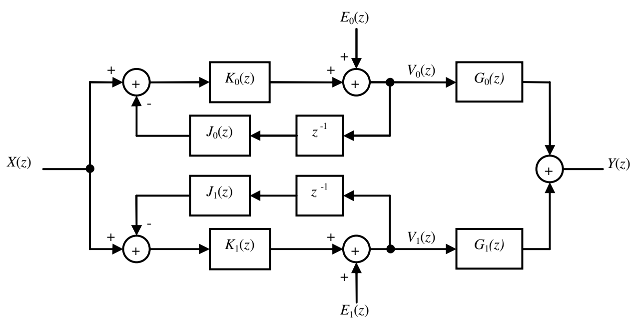


Fig. 5. Dual channel filter bank containing sigma-delta modulators without change of sampling frequency

Both of decomposition filters are realized by sigma-delta modulators and reconstruction part is realized by original filters. The resultant transfer of this structure is:

$$\begin{aligned} Y(z) = X(z) \cdot \frac{K_0(z)}{1 + z^{-1} \cdot J_0(z) \cdot K_0(z)} \cdot G_0(z) + \\ + X(z) \cdot \frac{K_1(z)}{1 + z^{-1} \cdot J_1(z) \cdot K_1(z)} \cdot G_1(z) + \\ + E_0(z) \cdot \frac{1}{1 + z^{-1} \cdot J_0(z) \cdot K_0(z)} \cdot G_0(z) + \\ + E_1(z) \cdot \frac{1}{1 + z^{-1} \cdot J_1(z) \cdot K_1(z)} \cdot G_1(z) \end{aligned} . \tag{11}$$

If transfer function $J(z)$ is expressed from Eq. (8), which describe required transfer of individual decomposition filter bank parts, and the transfer $J(z)$ is substituted into Eq. (11). The resultant form is:

$$\begin{aligned} Y(z) = X(z) \cdot H_0(z) \cdot G_0(z) + X(z) \cdot H_1(z) \cdot G_1(z) + \\ + E_0(z) \cdot \frac{H_0(z)}{K_0(z)} \cdot G_0(z) + E_1(z) \cdot \frac{H_1(z)}{K_1(z)} \cdot G_1(z) \end{aligned} . \tag{12}$$

From this equation can be seen that first two terms describe passage of input signal and last two terms describe passage of distortion subband signals. If Eq. (1) is fulfilled, the input signal is not changed by passage through this structure, only the subband error signals are added to input signal. By choosing transfer function $K_0(z)$ and $K_1(z)$ the transfer of subband error signals to output signal can be influenced. From this transfer functions it is necessary to compute transfer functions $J_0(z)$ and $J_1(z)$ from Eq. (9). Because filters with transfer functions $J_0(z)$ and $J_1(z)$ can not be causal the next condition has to be fulfilled in choosing transfer functions $K_0(z)$ and $K_1(z)$.

If the filters with transfer functions Eqs. (3), (4), (5) and (6) are used in reconstruction part, which are simple FIR filters with two coefficients, then the transfer function $K_0(z)$ is chosen as a product of transfer functions $H_0(z)$ and $G_0(z)$ and the transfer function $K_1(z)$ is chosen as a product of transfer functions $H_1(z)$ and $G_1(z)$. When the transfer functions are chosen by this way the uniform transfer of subband error signal should be happened, because the transfer functions are changed from polynomials to only constants. In choosing it is necessary to fulfill condition for causality thereby multiplication by constant are added to product. The chosen transfer functions are:

$$K_0(z) = 2 \cdot H_0(z) \cdot G_0(z) = \frac{1}{2} + z^{-1} + \frac{1}{2} z^{-2}, \tag{13}$$

$$K_1(z) = 2 \cdot H_1(z) \cdot G_1(z) = -\frac{1}{2} + z^{-1} - \frac{1}{2} z^{-2}. \tag{14}$$

The functions $F_0(z)$ and $F_1(z)$ are established for next computing which describe transfer of individual error signals into output signal. The resultant transfer of subband error signals after substitution $K_0(z)$ and $K_1(z)$ are:

$$F_0(z) = \frac{H_0(z)}{K_0(z)} \cdot G_0(z) = \frac{1}{2}, \tag{15}$$

$$F_1(z) = \frac{H_1(z)}{K_1(z)} \cdot G_1(z) = -\frac{1}{2}. \tag{16}$$

For requirement of sigma-delta modulator the functions $J_0(z)$ and $J_1(z)$ are causal and they are in followed form:

$$J_0(z) = \frac{K_0(z) - H_0(z)}{z^{-1} \cdot K_0(z) \cdot H_0(z)} = \frac{2 + 2z^{-1}}{1 + 3z^{-1} + 3z^{-2} + z^{-3}}, \tag{17}$$

$$J_1(z) = \frac{K_1(z) - H_1(z)}{z^{-1} \cdot K_1(z) \cdot H_1(z)} = \frac{2 - 2z^{-1}}{1 - 3z^{-1} + 3z^{-2} - z^{-3}}. \tag{18}$$

The resultant output error signal is sum of individual subband signals multiplied by individual transfer functions $F_m(z)$. If the subband error signals are white noises with same power in both of branches then output error signal is square root of sum of square of absolute values individual subband error transfers.

$$|F(z)| = \sqrt{|F_0(z)|^2 + |F_1(z)|^2} = \frac{1}{\sqrt{2}} \tag{19}$$

In this case the maximum of total transfer of error signal into output signal is about 0.707. In the case of classic realization of this filter bank, the maximum transfer of error signal into output signal is unitary. Dual channel filter bank with subsampling is in future research.

V. RESULTS OF FILTER BANK SIMULATION

In simulation four simple filters with impulse characteristic by Eqs. (3), (4), (5) and (6) was chosen. The simulation was processed in Matlab. Differential equations (20), (21), (22), (23), (24) and (25) were used in sigma-delta modulator in decomposition part of filter bank. The computing complexity is higher. In classic realization of this filter bank two products and one sum need to be computed for each sample. In the case of sigma-delta modulator is used eleven products and nine sums need to be computer for each sample.

$$v_0[n] = k_0[n] + e_0[n], \tag{20}$$

$$v_1[n] = k_1[n] + e_1[n], \tag{21}$$

$$\begin{aligned} j_0[n] = v_0[n-1] + 2 \cdot v_0[n-2] - 3 \cdot j_0[n-1] - \\ - 3 \cdot j_0[n-2] - j_0[n-3] \end{aligned}, \tag{22}$$

$$\begin{aligned} j_1[n] = v_1[n-1] - 2 \cdot v_1[n-2] + 3 \cdot j_1[n-1] - \\ - 3 \cdot j_1[n-2] + j_1[n-3] \end{aligned}, \tag{23}$$

$$k_0[n] = \frac{1}{2}\cdot x[n] + x[n-1] + \frac{1}{2}\cdot x[n-2] - \\ -\frac{1}{2} j_0[n] - j_0[n-1] - \frac{1}{2} j_0[n-2], \quad (24)$$

$$k_1[n] = -\frac{1}{2}\cdot x[n] + x[n-1] - \frac{1}{2}\cdot x[n-2] + \\ +\frac{1}{2} j_1[n] - j_1[n-1] + \frac{1}{2} j_1[n-2]. \quad (25)$$

A signal $x[n]$ was chosen as sum of two harmonic signals (first with low frequency and second with high frequency) and signals $e_0[n]$ and $e_1[n]$ were chosen as random signals with uniform distributed functions. These signals were input signals to classic filter bank without subsampling and to filter bank with sigma-delta modulators which can be seen in Fig. 5.

Output signals from decomposition part can be seen in Fig. 6. The input signal is decomposed into two parts and one harmonic component is in each subband. Noise is uniform spread over whole spectrum of subband signals.

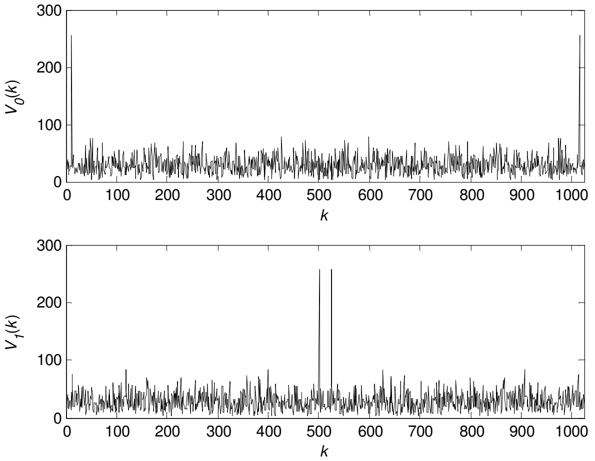


Fig. 6. Spectrum of output signals from original decomposition dual channel filter bank

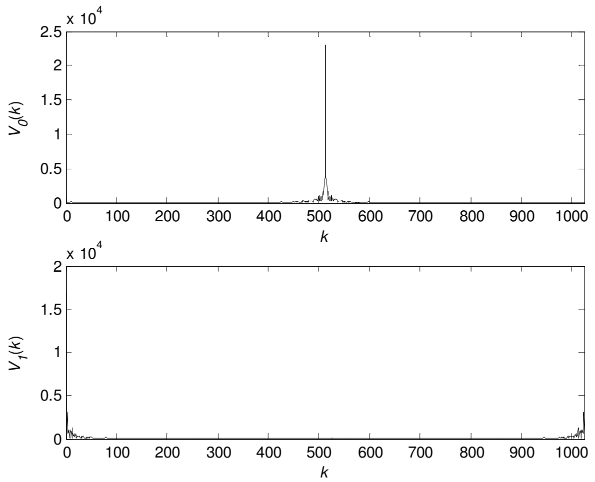


Fig. 7. Spectrum of output signals from decomposition dual channel filter bank with sigma-delta modulators

In the case the sigma-delta modular is used the noise is emphasized in these parts where the signals are suppressed in the reconstruction part. Reconstructed signals can be seen in Fig. 8. In picture can be seen that the noise in output signal is lower if sigma-delta modulator is used in decomposition part of filter bank. This decrement of the noise matches equation (19).

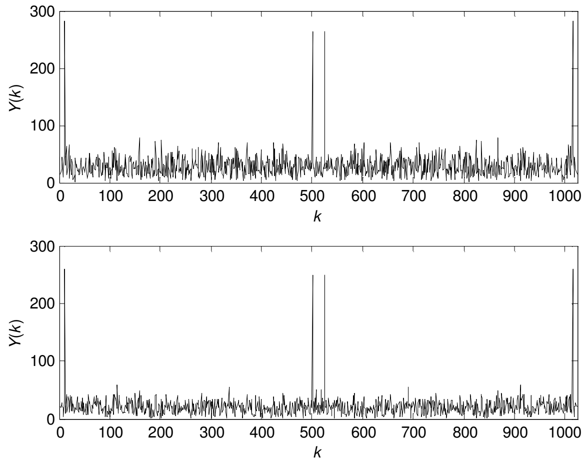


Fig. 8. Spectrum of output signals from original reconstruction dual channel filter bank (upper) and reconstruction dual channel filter bank with sigma-delta modulators (lower)

VI. CONCLUSION

The filter bank with sigma-delta modulators is next realization possibility of filter bank. In proper design of this structure it is more resistant to distortion of individual subband signals. Its advantage is in individual subband error signal shaping so that it is asserted in output signal at least. In comparison with classic realization of filter bank, this structure is more compute complexity.

VII. ACKNOWLEDGEMENT

The paper was prepared within the framework of N° 102/04/1097 project of the Grant Agency of the Czech Republic and the National Research Project “Information Society” No 1ET301710509.

REFERENCES

[1] D. Jarman, *A Brief Introducion to Sigma Delta Conversion*, Aplication Note AN9504, Intersil, 1995

[2] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice hall P T R, Englewood Cliffs, New Jersey, 1993, ISBN 0-13-605718-7

[3] Z. Smékal, R. Vích, *Číslicové filtry*, Academia, Praha, 2000, ISBN 80-200-0761-X.

[4] J. Jan, *Číslicová filtrace, analýza a restaurace signálů*, VUT Brno, 1997, ISBN 80-214-0816-2.

Common theory, approximation method and design of electrical filters based on Hausdorff polynomials

Peter Apostolov[1]

Abstract - **An application of Hausdorff approximation in the modern theory of electrical filters design is shown in the paper. A translated Hausdorff polynomial is defined, which leads to realizable transmission functions of electrical filters. Transmission functions of Hausdorff low-pass filter-prototype and two types of inverse Hausdorff filters are shown. Their frequency characteristics are analyzed.**

Keywords - **Approximation, Polynomial, Design, Inverse filter, Hausdorff, Chebyshev, Magnitude response.**

I. Introduction

In modern filter theory the synthesis is performed by appropriate characteristic function (Fig.1) approximation.

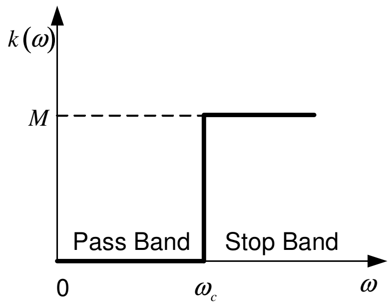


Fig. 1. Characteristic function

The most popular approximations are the canonical – equal-ripple in pass-band (Chebyshev I, Cauer) and maximum-flat in pass-band (Butterworth, Hourglass, Chebyshev II). Another approximations with intermediate properties also exist. Such approximations are of Bessel, Legendre and Gauss.

On Fig. 2 and Fig. 3 amplitude responses and phase responses of filters obtained by mentioned approximations are shown.

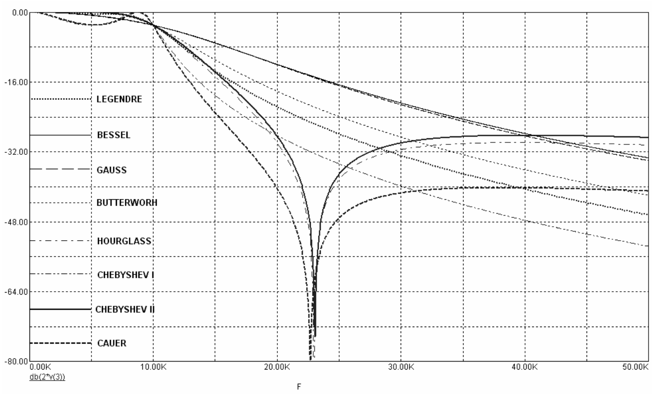


Fig. 2. Comparison of magnitude responses [dB]

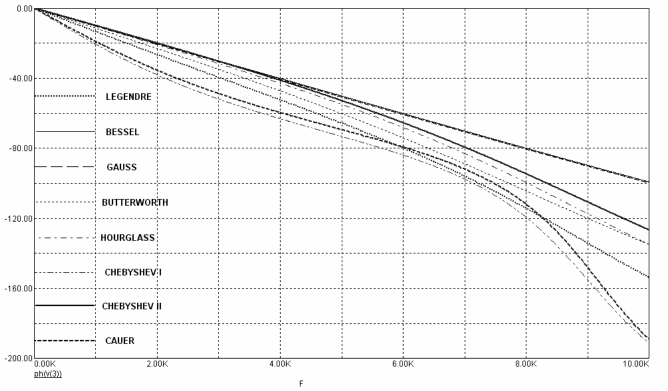


Fig. 3. Comparison of phase responses [deg]

As can be seen, the filters, as linear electrical circuits, have contradictory behaviour. Those with better selectivity (Chebyshev I and Cauer) have non-linear phase-responses and vice versa - those with worse selectivity (Gauss, Bessel, Butterworth) have more linear phase responses. The Chebyshev's II filters combine better selectivity and acceptable linearity of phase response.

The advance of the digitalization in the last years is a motivation for finding new approximations, which lead to decreased distortions of digital signal processing. A step in this direction is the approximation in Hausdorff metric.

II. Approximation in Hausdorff metric. Formation of transmission function of Hausdorff low-pas filter prototype

The Bulgarian Academy of Science (BAS) conducted a research of ε-entropy of the space, a new approximation in Hausdorff metric [5] is offered in the eighties. The theoretical results of this work were successfully applied in an antenna array design for the Technical University of Sofia [2] and in the transmission functions direct synthesis of digital filters in BAS [3]. In the theory, an algebraic polynomial accomplishing the best approximation of "shifted" delta function in Hausdorff metric is offered

$$P_n(x) = \varepsilon T_n\left(\frac{2x+\alpha\varepsilon}{2-\alpha\varepsilon}\right), \tag{1}$$

where ε is Hausdorff distance, T_n is Chebyshev's polynomial of first kind and n degree; α is a parameter, and the factor (product) $\alpha\varepsilon$ determines the bandwidth in which the polynomial approximates the "shifted" delta function in point

[1] Peter St. Apostolov, Ms. Sc., Assoc. Fellow, Institute for Special Technical Equipment-MI, Sofia, e-mail: p_apostolov@abv.bg

1, where it has infinitive steepness (Fig. 4). The relations between polynomial parameters [1] are defined by equation

$$\alpha\varepsilon = 2\frac{\mathrm{ch}\left[\frac{1}{n}\mathrm{Ach}\left(\frac{1}{\varepsilon}\right)\right]-1}{\mathrm{ch}\left[\frac{1}{n}\mathrm{Ach}\left(\frac{1}{\varepsilon}\right)\right]+1}. \tag{2}$$

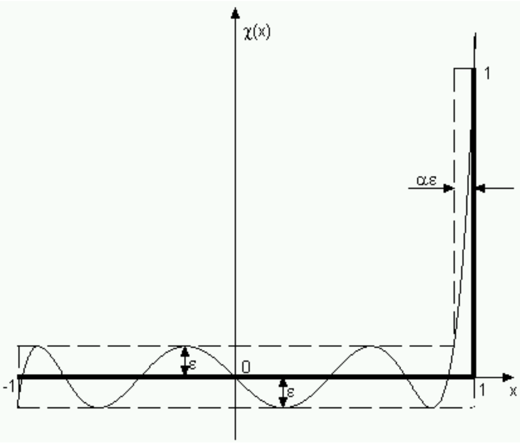


Fig. 4. Approximation with Hausdorff polynomial

As seen on the figure the Hausdorff polynomial approximates function with outline of filter characteristic function (Fig. 1). The pass-band is determined in the interval $[0,1-\alpha\varepsilon]$, stop-band in the interval $(1-\alpha\varepsilon,\infty)$. Considering Eq. (1) and Fig.4 in the defined area the following inequalities (InEq) are valid:

$$0<\varepsilon<1;\ 0<\alpha\varepsilon<1. \tag{3}$$

The representation of the polynomial as a rational function of its argument can be determined by the coefficients of the Chebyshev's polynomial. A fifth-order Hausdorff polynomial will look like

$$P_5(x)=\varepsilon\left[16\left(\frac{2x+\alpha\varepsilon}{2-\alpha\varepsilon}\right)^5-20\left(\frac{2x+\alpha\varepsilon}{2-\alpha\varepsilon}\right)^3+5\frac{2x+\alpha\varepsilon}{2-\alpha\varepsilon}\right]. \tag{4}$$

Fig. 5 shows the graphical representation of the polynomial with values of $\varepsilon=0.07556$ and $\alpha\varepsilon=0.2$.

It can be seen in Eq. (4) that after ε and $\alpha\varepsilon$ are substituted with their values, raising to the power, removing the brackets and reducing, the Hausdorff polynomial can be represented as a rational function of its argument.

$$P_5(x)=2.04738\,x^5+1.02369x^4-1.86824x^3- \\ -0.601419x^2+0.358612x+0.0399253. \tag{5}$$

It is known from the theory that the transmission function of a low-pass filter is a fraction-rational. To achieve a realizable transmission function, the polynomial in the denominator is required to be a strict Hurwitz' polynomial. The formation of the transmission function, using the polynomial in Eq. (4), cannot lead to a realizable transmission function for every $\alpha\varepsilon$ [1]. That is because the Hausdorff polynomial in Eq. (4) is neither even nor odd function of its argument, which can be seen from Eq. (5) and Fig. 5. The polynomial graph is asymmetrical with regards the centre of the coordinate system.

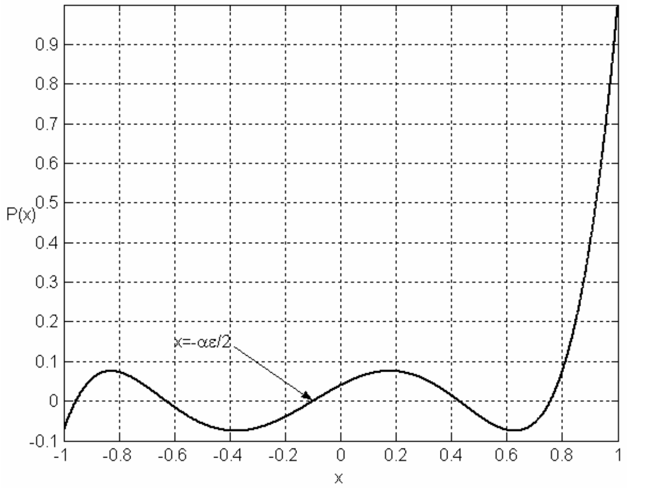


Fig. 5. Hausdorff polynomial, n=5, ε =0.07556, $\alpha\varepsilon$ =0.2

The investigations showed, that for the value of argument $x=-\alpha\varepsilon/2$, the Hausdorff polynomials have a typical point: a root, when the polynomial is even (Fig.5) and a local extremum, when it is odd. If a translation with a value of $\alpha\varepsilon/2$ in the positive direction is performed, the translated polynomials lead to realizable transmission functions.

$$P_n^T(x)=\varepsilon T_n\left(\frac{2x}{2-\alpha\varepsilon}\right)=\varepsilon T_n\left(\frac{x}{1-\alpha\varepsilon/2}\right). \tag{6}$$

It is proved in [1], that the translated polynomial represents the only and the best approximation of "shifted" delta function, translated with $\alpha\varepsilon/2$ in positive direction, in Hausdorff metric.

The square of the transmission function module of low-pass Hausdorff filter-prototype has the form

$$|A_n|^2(\omega)=\frac{1}{1+\varepsilon^2T_n^2\left(\frac{2\omega}{2-\alpha\varepsilon}\right)}=\frac{1}{1+\varepsilon^2T_n^2\left(\frac{\omega}{1-\alpha\varepsilon/2}\right)}, \tag{7}$$

where ω is the angle frequency.

III. Design of Hausdorff Filters

It is known that the module of the transmission function defines the magnitude response of the filter. In Eq. (7) the argument of the Chebyshev's polynomial is divided by the expression $(1-\alpha\varepsilon/2)$. Considering InEq. (3), the expression is positive number less than 1. That leads to a scale "shrinking" of the magnitude response compared to the magnitude response of a Chebyshev's I filter as shown on Fig. 6.

The low-pass Hausdorff filter-prototype design led to the following results:

- The pass-band of the Hausdorff filter-prototype is "shrunk" by coefficient $(1-\alpha\varepsilon/2)$.

- For the same order and ripple in the pass-band, the magnitude response of a low-pass Hausdorff filter-prototype has the same steepness and attenuation in the stop-band as a Chebyshev's I filter.

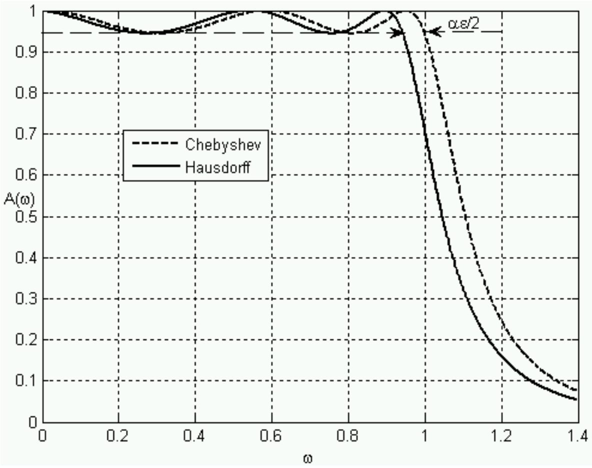


Fig. 6. Magnitude responses of Hausdorff and Chebyshev's I filters

- Hausdorff low-pass filter-prototype has the same linearity in phase response in the pass-band as a Chebyshev's I filter.
- Hausdorff low-pass filter-prototype has the same evenness of GDT in the pass-band as a Chebyshev's I filter.
- Hausdorff low-pass filter-prototype cannot have bigger unevenness in the pass - band than $1/\sqrt{2}$ (-3.01dB), because $\varepsilon < 1$ InEq. (3).
- Poles' quality factors have the same values as a Chebyshev's I filter.

From the above mentioned, a conclusion could be made: *The Hausdorff low-pass filter-prototype cannot find practical application, because of the undesirable "shrinking" of the pass-band.*

The inverse Hausdorff filters are more interesting from a design point of view. Two types are described here – A and B as explained next. The modules of their transmission functions are shown with the following two equations:

$$\left|A_n(\omega)\right|_A = \sqrt{\frac{\varepsilon^2 T_n^2\left[\dfrac{1}{\omega(1-\alpha\varepsilon/2)}\right]}{1+\varepsilon^2 T_n^2\left[\dfrac{1}{\omega(1-\alpha\varepsilon/2)}\right]}}; \tag{8}$$

$$\left|A_n(\omega)\right|_B = \sqrt{\frac{\varepsilon^2 T_n^2\left(\dfrac{1-\alpha\varepsilon/2}{\omega}\right)}{1+\varepsilon^2 T_n^2\left(\dfrac{1-\alpha\varepsilon/2}{\omega}\right)}}. \tag{9}$$

The difference is that for type A the expression $(1-\alpha\varepsilon/2)$ multiplies the argument of Chebyshev's polynomial while for type B it divides the argument. That naturally leads to a scale expanding/shrinking of the magnitude response such as for the low-pass filter-prototype. In contrast to that, the inverse Hausdorff filters have magnitude response poles – real frequencies, where the attenuation is infinitive. This allows the method of concluded equal-ripple approximation [4], [6] to be applied to these filters. Inverse Hausdorff filters could be designed to keep their cut-off frequency, and frequencies in the stop-band could be proportional to the $\alpha\varepsilon$ factor.

In the synthesis the transfer function is represented as relation of three polynomials *e(s)*, *p(s)* and *q(s)* of complex frequency $s = j\omega$. The polynomial *e(s)* is Hurwitz strict polynomial and its zeros ω_i represent the filter own frequencies and these of *p(s)* - the extreme frequencies $\omega_{\infty i}$, for which the transfer function has infinite attenuation. Calculating two of polynomials usually solves the synthesis task and the third is defined by the equation:

$$e(s)e(-s) = p(s)p(-s) + q(s)q(-s) \tag{10}$$

The zeros of *e(s)* and *p(s)* can be found as follows:

For A-type:

$$\omega_i = \frac{1}{\left(1-\dfrac{\alpha\varepsilon}{2}\right)(\sigma_i + j\Omega_i)}; \tag{11}$$

$$\omega_{\infty i} = \frac{j}{\left(1-\dfrac{\alpha\varepsilon}{2}\right)\cos\left(\dfrac{2i-1}{n}\dfrac{\pi}{2}\right)}. \tag{12}$$

For B-type:

$$\omega_i = \frac{1-\dfrac{\alpha\varepsilon}{2}}{\sigma_i + j\Omega_i}; \tag{13}$$

$$\omega_{\infty i} = \frac{j\left(1-\dfrac{\alpha\varepsilon}{2}\right)}{\cos\left(\dfrac{2i-1}{n}\dfrac{\pi}{2}\right)}, \tag{14}$$

where:

$$\sigma_i = -\sin\left(\frac{2i-1}{n}\frac{\pi}{2}\right)\mathrm{sh}\left[\frac{1}{n}\mathrm{Ash}\left(\frac{1}{\varepsilon}\right)\right]; \tag{15}$$

$$\Omega_i = \cos\left(\frac{2i-1}{n}\frac{\pi}{2}\right)\mathrm{ch}\left[\frac{1}{n}\mathrm{Ash}\left(\frac{1}{\varepsilon}\right)\right]. \tag{16}$$

Fig. 7 compares the magnitude responses of the two types inverse Hausdorff filters with Chebyshev's II filter.

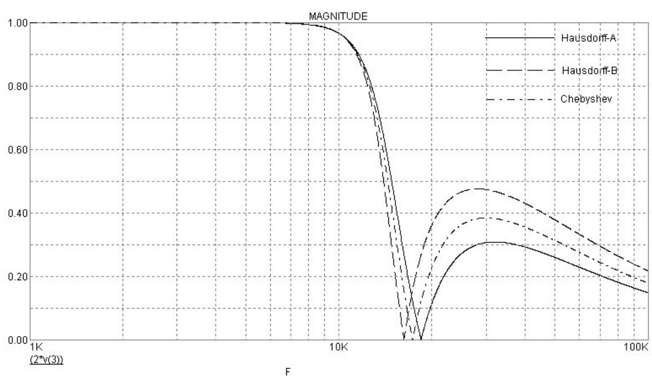


Fig. 7 Comparison of magnitude responses

It can be seen from the figure, that the inverse Hausdorff filters type A (IHF-A) have less magnitude response steepness in the area between the cut-off frequency and the frequency of an infinite attenuation and higher attenuation in the stop-band in comparison with Chebyshev's II filter. As for IHF-B it is the contrary – a bigger steepness and less attenuation in the stop-band.

Fig.8 compares the phase responses. It is seen from the comparison that IHF-A has the best linearity. In this case improvement of the linearity compared with Chebyshev's II filter for the frequency 6kHz is 5.3%. IHF-B has the worse linearity.

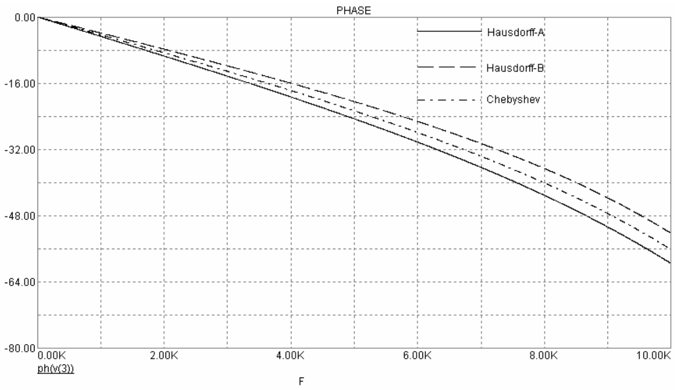


Fig. 8. Comparison of phase responses

Fig. 9 compares the GDT of the filters.

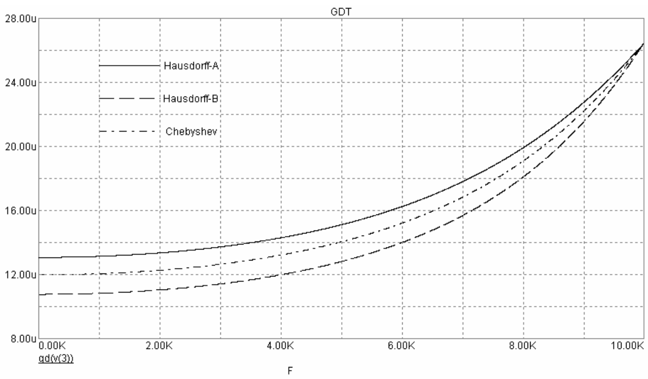


Fig. 9. Comparison of GDT

The improvement of the evenness of GDT of IHF-A compared with Chebyshev's II filter in this case is a little more than 9%. GDT of IHF-B is the most uneven.

Fig.10 compares the poles of the filters in a complex domain.

The poles values distribution could be used to judge the poles' quality factors of the filters. A criterion is the remoteness of the poles with imaginary part from the imaginary axis. The poles of IHF-A are the most remote ones. That means they will have the lowest values of the poles' quality factors and they will have easier values for components realization.

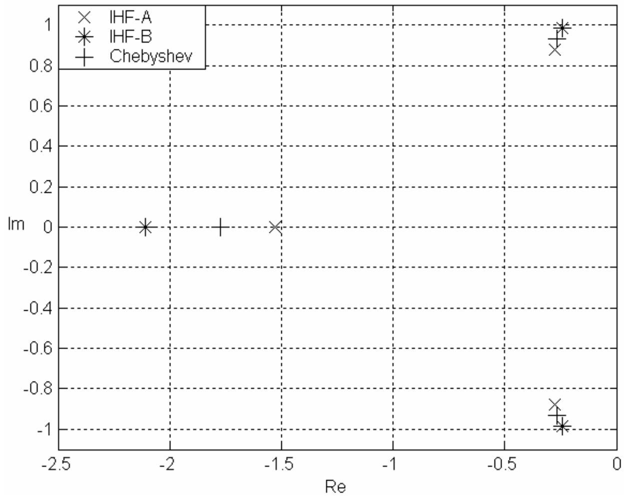


Fig. 10. Poles values distribution

IV. CONCLUSION

The approximation implementation in Hausdorff metric doesn't lead to a "revolution" in the filter design domain. But it could be said, that it is an "evolution" in this domain of investigation. The implementation of the Chebyshev's polynomial in Eq.1, creates filters with characteristics similar to Chebyshev's filters. The difference is defined by the value of the factor $\alpha\varepsilon$.

Low-pass Hausdorff filter cannot find practical application, because there is undesirable "shrinking" of the pass-band, proportional to the factor $\alpha\varepsilon$.

When a suppression of the signals close to the cut-off frequency is required, the IHF-B filters can be used due to their bigger steepness of the magnitude response. They have more nonlinear phase response, more uneven GDT and higher values of poles' quality factors

IHF-A filters are probably the best that can be achieved when applying the Hausdorff metric approximation. In comparison with Chebyshev's II filters they have higher attenuation in the pass–band, more linear phase response, more even GDT and lower poles quality factors. These properties make them appropriate for implementing in filtration of digital signals.

V. REFERENCES

[1] Apostolov, P. Hausdorff filters (analysis and synthesis), Dissertation, Technical University - Sofia, 2005.

[2] Gatchev, M. G. Synthesis of antennas arrays with optional radiation pattern, Dissertation Technical University-Sofia, 1981.

[3] Marinov, P. G. Algorithms for approximation with rationally functions respect to Hausdorff distance. Dissertation, BAS – Sofia, 1991.

[4] Orchard, H. J., G. C. Temes. Filter design using transformed variables. – IEEE trans. On circuit theory, 1968. CT-15, pp.385-408.

[5] Sendov B. Hausdorff approximations, Kluwer Academic Publishers London 1990, ISBN:0792309014.

[6] Stoyanov, G., I. Uzunov, L. Raykovska and R. Bradvarov. Analysis, synthesis and design of electrical filters with personal computers. Sofia, Technica, 1991.

New Complex Orthogonal Narrowband IIR First-Order Filter Sections – an Input Quantization Noise Analysis

Zlatka Nikolova[1], Georgi Stoyanov[2]

Abstract – **In this paper a very low sensitivity first-order orthogonal complex narrow-band band-pass filter section is developed. Then, a noise analysis due to the quantization of the input signal is performed. A comparative study with other well-known complex orthogonal sections has been made on order to show the advantages of the developed filter circuit.**

Keywords – **complex orthogonal digital filters, sensitivity, quantization errors, noise analysis.**

I. Introduction

Quantization of the multiplication products and the input signals in the digital filters is causing parasitic noises usually described as an error signal. Quantization of the multiplier coefficients is deteriorating the filter characteristics. Development of low sensitivity structures is reducing considerably this deterioration and additionally is decreasing the roundoff noises.

Methods of computing the quantization effects in real digital systems are well developed so far. In [1] an evaluating recursive formula for roundoff errors estimation is proposed. The "inners" approach for output noise variance evaluation is suggested in [2]. An alternate method of calculating the noise, suitable both for analytical and numerical computations and useful for higher order filters is presented in [3]. A new algebraic technique proposed by Bomar in [4] provides computationally efficient state-space realizations preserving low roudoff noise, low coefficient sensitivity, and freedom from zero-input cycles.

Although complex coefficients digital filters are gaining popularity in the recent years their quantization noise analysis theory is still not well developed and there are few publications treating the problem. New structures for complex multipliers and their noise analysis are proposed in [5]. In [6] theoretical upper bounds on the amplitude of limit cycles oscillations are determined for direct-form orthogonal second order digital filter. In all these works usually local problems are solved and no general method for roundoff noise estimation is proposed. In this work we present a technique for complex input signal quantization noise analysis. Then we apply it on a newly developed very low sensitivity complex orthogonal first order sections.

The paper is organized as follows. An approach to a complex noise analysis is proposed in section II. The first-order orthogonal complex digital filters are derived in section III. In section IV the resulting error signals at the outputs of these orthogonal complex sections after input signal quantization are examined. In section V, simulation results for quantized input narrow-band complex signals are presented and discussed. Finally, section VI concludes the paper.

[1] Zlatka Nikolova is with the Dept. of Telecommunications, Technical University of Sofia, Bulgaria, e-mail: zvv@tu-sofia.bg

[2] Georgi Stoyanov is with the Dept. of Telecommunications, Technical University of Sofia, Bulgaria, e-mail: styoyanov@ieee.org

II. Quantization Noise Analysis for Complex Input Signals

The effect of quantization of an input signal is equivalent to a set of noise samples added to the actual input. In case of uniformly distributed noise samples, the variance of the input noise is:

$$\sigma_e^2 = \frac{\delta^2}{12} = \frac{2^{-2B}}{12}, \tag{1}$$

where δ^2 is the quantization step and B is the word-length in bits. Then the steady-state (nominal) value of the output noise variance is given by:

$$\sigma_v^2 = \sigma_e^2 \frac{1}{2\pi j} \oint H(z) H\left(z^{-1}\right) z^{-1} dz = \sigma_e^2\, \sigma_{v,n}^2, \tag{2}$$

where $H(z)$ is the transfer function of the digital filter, $\sigma_{v,n}^2$ denotes the noise gain and is called also normalized output noise variance.

Complex coefficient digital filters are capable of processing both real and complex signals. The quantization of a complex input signal presumes a complex error signal as it is shown in Fig. 1. Then the complex output signal $y(n) = y_{\mathrm{Re}}(n) + jy_{\mathrm{Im}}(n)$ will be mixed with complex output noise:

$$v(n) = v_{\mathrm{Re}}(n) + jv_{\mathrm{Im}}(n). \tag{3}$$

Through this model it is possible to examine the complex output noise variance in a similar way as in the real case.

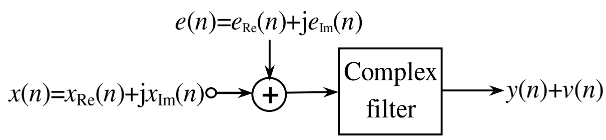


Fig. 1: Noise model for a complex input signal quantization

Evaluating the complex output noise variance presumes all complex quantities to be considered in Eq.(2). Analytic input signals with inphase and quadrature components are processed by special class of complex coefficient digital filters called "orthogonal". The orthogonal complex transfer function can be presented by its real and imaginary parts as follows:

$$H(-jz) = H_{\mathrm{Re}}(z) + jH_{\mathrm{Im}}(z). \tag{4}$$

Realized by real elements, an orthogonal complex structure (Fig. 2) will have two inputs and two outputs (both couples real and imaginary), producing thereby four real coefficient transfer functions two by two equal with ± sign:

$$H_{\mathrm{Re}}(z) = H_{RR}(z) = H_{II}(z);$$
$$H_{\mathrm{Im}}(z) = H_{RI}(z) = -H_{IR}(z). \tag{5}$$

Normally the real and imaginary parts of the complex input signal undergo the same quantization, i.e. $\sigma^2_{e,\mathrm{Re}}=\sigma^2_{e,\mathrm{Im}}=\sigma^2_e$.

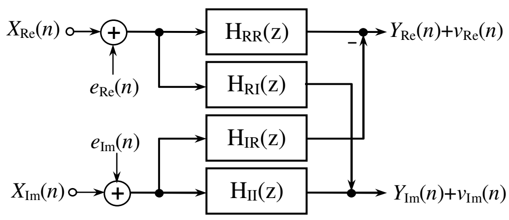


Fig. 2: Block-diagram of complex digital filter structure - noise model for a complex input signal quantization

The structure in Fig.2 shows that the complex output noise variance real and imaginary components should be composed as follow:

$$\sigma^2_{v,\mathrm{Re}} = \sigma^2_e\left(\sigma^2_{v,H_{\mathrm{Re}}} - \sigma^2_{v,H_{\mathrm{Im}}}\right) \tag{6}$$

$$\sigma^2_{v,\mathrm{Im}} = \sigma^2_e\left(\sigma^2_{v,H_{\mathrm{Re}}} + \sigma^2_{v,H_{\mathrm{Im}}}\right) \tag{7}$$

where

$$\sigma^2_{v,H_{\mathrm{Re}}} = \sigma^2_e \frac{1}{2\pi j}\oint H_{\mathrm{Re}}(z)H_{\mathrm{Re}}\left(z^{-1}\right)z^{-1}dz \tag{8}$$

$$\sigma^2_{v,H_{\mathrm{Im}}} = \sigma^2_e \frac{1}{2\pi j}\oint H_{\mathrm{Im}}(z)H_{\mathrm{Im}}\left(z^{-1}\right)z^{-1}\,dz \tag{9}$$

are the output noise variances for the real and imaginary components of the transfer function (4).

III. Complex Orthogonal Digital Filter Circuit Development

One of the best methods of complex digital filter derivation is the method of circuit transformation proposed in [7] and permitting also to obtain orthogonal complex filters with canonical number of elements. We select this method expecting that it will permit the new circuit to inherit good qualities of the real prototype.

According to this expectation and after a careful study of the more often used real-coefficient first-order digital filter structures we selected the best two very low-sensitivity real-prototypes – MHNS and LS1b low-pass sections [8] (Fig. 3).

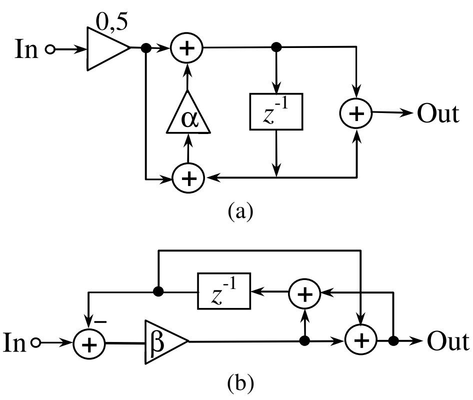


Fig. 3: Real-coefficients first-order low-sensitivity LP filters sections (a) MHNS; (b) LS1b.

The transfer functions these real sections realize are:

$$H^{LP}_{MHNS}(z) = \frac{1-\alpha}{2}\frac{1+z^{-1}}{1-\alpha z^{-1}}; \tag{10}$$

$$H(z)^{LP}_{LS1b} = \beta\frac{1+z^{-1}}{1-(1-2\beta)z^{-1}}. \tag{11}$$

The rotation transformation [7] in its orthogonal case:

$$z^{-1} = jz^{-1} \text{ or } z = -jz, \tag{12}$$

applied on the LP real transfer functions converts them into orthogonal complex coefficients BP transfer functions of doubled order. For the MHNS-based orthogonal complex structure shown in Fig. 4a they are as follow:

$$H^{RR}_{MHNS}(z) = H^{II}_{MHNS}(z) = H_{\mathrm{Re}}(z) = \frac{(1-\alpha)}{2}\frac{1-\alpha z^{-2}}{1+\alpha^2 z^{-2}}, \tag{13}$$

$$H^{RI}_{MHNS}(z) = -H^{IR}_{MHNS}(z) = H_{\mathrm{Im}}(z) = \frac{\left(1-\alpha^2\right)}{2}\frac{z^{-1}}{1+\alpha^2 z^{-2}}, \tag{14}$$

whereas for the LS1b-based orthogonal section (Fig.4b) they are:

$$H^{RR}_{LS1b}(z) = H^{II}_{LS1b}(z) = H_{\mathrm{Re}}(z) = \beta\frac{1+(2\beta-1)z^{-2}}{1+(2\beta-1)^2 z^{-2}}, \tag{15}$$

$$H^{RI}_{LS1b}(z) = -H^{IR}_{LS1b}(z) = H_{\mathrm{Im}}(z) = \beta(1-\beta)\frac{2z^{-1}}{1+(2\beta-1)^2 z^{-2}}. \tag{16}$$

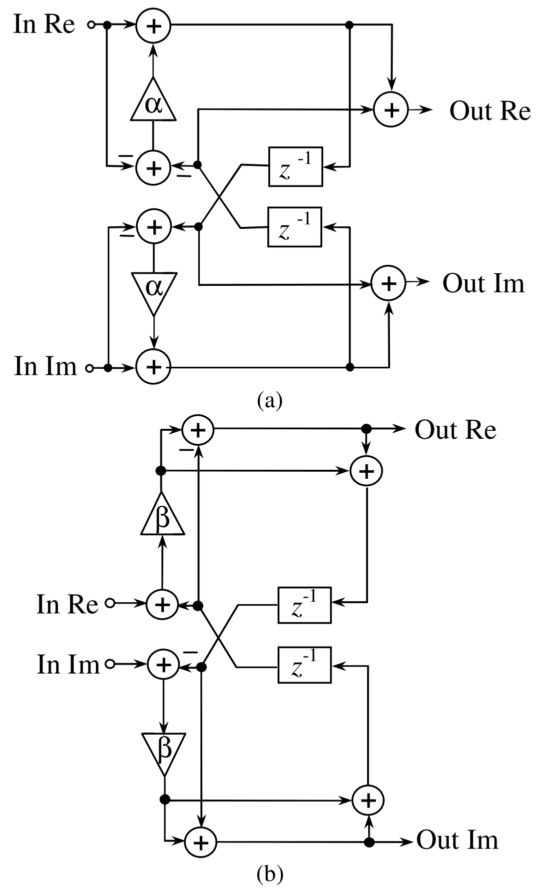


Fig. 4: First-order BP orthogonal structure based on the (a) MHNS; (b) LS1b real sections.

Narrow-band orthogonal BP filters are most often used for practical purposes and they can be derived from narrow-band LP filter-prototypes (pole near *z=1*). Achieving low sensitivity for such pole position is a very difficult task. In [8] the orthogonal structures from Fig.4 have been investigated with respect to their coefficient sensitivity for a given poles-disposition. It was clearly shown that LS1b-based structure preserves its magnitude shape even when the coefficients are quantized to 2 bits, while MHNS-based structure response is considerably changed when the word-length is limited to 3 bits only. The real prototype-sections (Fig. 3) keep the same performance.

IV. Noise Analysis of Complex Input Quantization Errors

In this section both real and orthogonal structures regarding input quantization errors are investigated.

Initially the real input signal for the LS1b and MHNS real sections is quantised to different word-length. First the output noise variance of the real prototype sections, realised with very narrow pass-band (α=0.98 and β=0.01) is calculated. Some experimental results for input signal quantization from 2 to 8 bits are shown in Fig. 5. Apparently, the low sensitivity LS1b section output noise variance is about ten times lower than this of the MHNS-section when the input signal is limited to 2 bits only.

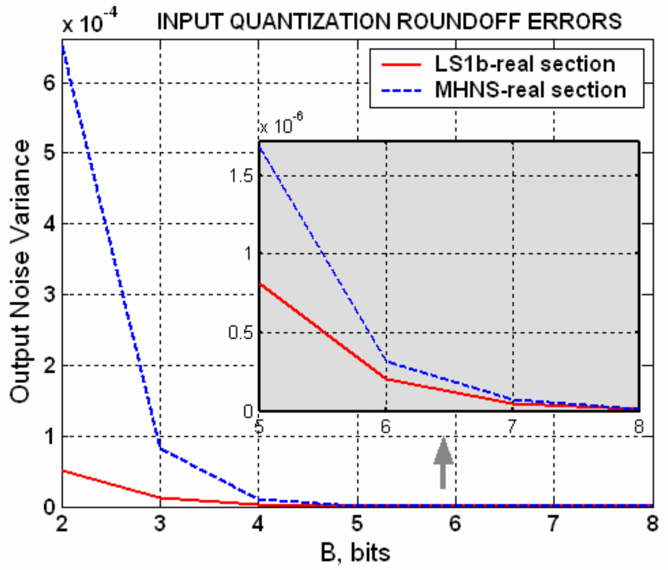


Fig. 5: The output noise variance under input signal roundoff quantization for LS1b and MHNS real sections

Then, the corresponding narrow-band BP orthogonal complex filter sections (Fig. 4) are investigated by using the method from section II.

Following the described in section II method, a complex input signal quantization noise analysis is performed. The calculation results for complex output noise variances for the LS1b and MHNS orthogonal complex sections in different input signal word-length are presented in Tabl. 1.

In order to compare the obtained complex output signal noise variances, their complex modules are shown in Fig. 6.

Tabl. 1

Input signal quantization in bits	Complex output noise variances of the orthogonal complex sections
	MHNS-based ($x\ 10^{-3}$)
2	0.0120638996891700 + j 0.1817906894536900
3	0.00301597492229 + j 0.04544767236342
4	0.00075399373057 + j 0.01136191809086
5	0.00018849843264 + j 0.0028404795227l
6	0.00004712460816 + j 0.00071011988068
7	0.00001178115204 + j 0.00017752997017
8	0.0000294528801 + j 0.00004438249254
	LS1b-based ($x\ 10^{-4}$)
2	0.00514438391361 + j 0.51766175890637
3	0.00128609597840 + j 0.12941543972659
4	0.00032152399460 + j 0.03235385993165
5	0.00008038099865 + j 0.00808846498291
6	0.00002009524966 + j 0.00202211624573
7	0.00000502381242 + j 0.00050552906143
8	0.0000125595310 + j 0.00012638226536

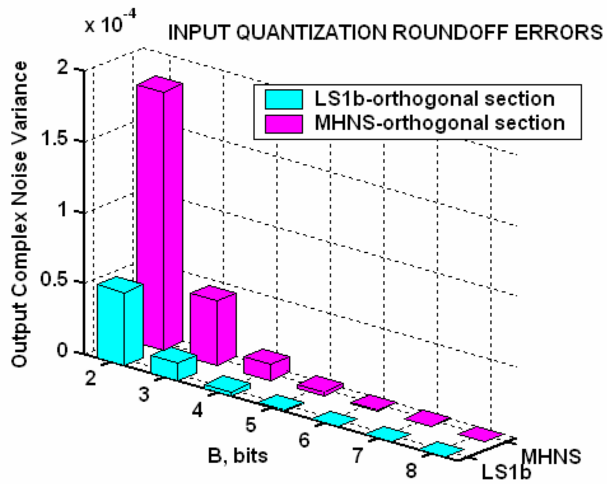


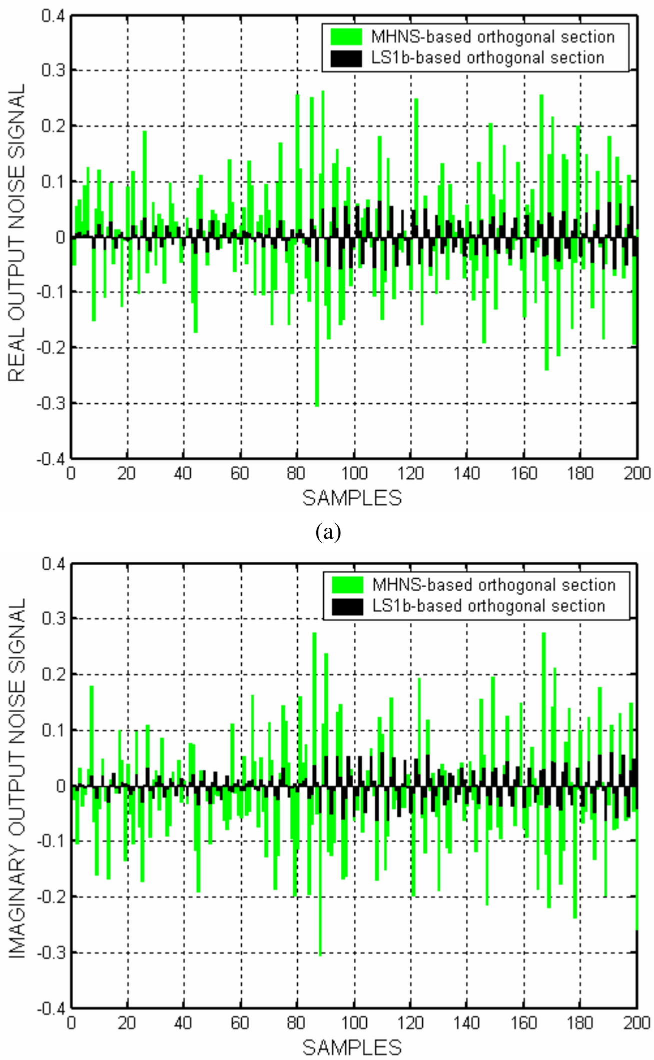
Fig. 6: The output noise variances under the input signal roundoff quantization for LS1b and MHNS -based orthogonal complex first-order sections

It is seen that, the low-sensitivity LS1b-based orthogonal complex section is having more than three times lower output noise after 2 bits input signal quantization. The shorter word-length quantization of the input signal means lower power consumption and faster computation process. For low-sensitivity circuits the resistance against quantization effects provides better signal to noise ratio (SNR), i.e. higher quality digital signal processing.

V. Experiments

The narrowband orthogonal first-order filter sections were investigated in a limited word-length complex signals processsing. The complex input signal is a mixture of white noise and analytic sinusoidal signal. The uniformly distributed white noise samples correspond to the word-length of the input complex signal after the quantization.

In case of 2 bits quantization both real and imaginary parts of the input analytic signal and filter coefficients, some experimental results are shown in Fig. 7. In Fig. 7a the real output noise signals for both orthogonal sections are presented, whilst the imaginary output noise signals are shown in Fig. 7b. Obviously the noise reaching to the complex output of the orthogonal circuits is significantly higher for the MHNS-based than for the LS1b-based section.



(b)

Fig. 7: The output noise signals after input quantization to 2 bits for LS1b and MHNS - based orthogonal complex sections (a) real output; (b) imaginary output.

The output SNR for the LS1b-orthogonal section is about 1,5 times higher in comparison to the MHNS-based circuit. To achieve the same good results as LS1b section demonstrates in 2 bits word-length environment, the MHNS orthogonal filter should be quantized to no less than 6 bits.

It is clear that the famous phrase from the real circuit theory "low sensitivity and low noise go together" is valid also for their complex counterpart.

VI. Conclusions

In this paper an approach to the complex noise analysis is proposed. The resulting error signals at the outputs of orthogonal complex first-order digital filter sections after input signal quantization are examined. The proposed method is general enough to be applied for complex filter sections of higher order. After relevant alterations it could be effectively applied for all other types of roundoff errors estimation in complex coefficient systems like multiplication product quantization.

The expectation that the real prototype properties will be inherited by its complex filter counterpart was confirmed ones again with respect to the quantization noise analysis. It was shown that both real and orthogonal complex LS1b-based filter sections having very low coefficient sensitivity demonstrate low output noise variance due to the input signal quantization – many times lower than that of the MHNS-based circuit.

References

[1] K. J. Astrom, E. I. Jury and R. G. Agniel, "A Numerical Method for the Evaluation of Complex Integrals". *IEEE Trans. Automat. Contr.*, vol. AC-15, pp. 468-471, Aug.. 1970.

[2] E. I. Jury and S. Gutman, "The Inner Formulation for the Total Square Integral", *Proc. IEEE (Lett.)* , vol. 61, pp. 395-397, March.. 1973.

[3] S. K. Mitra, K. Hirano and H. Sakaguchi, "A Simple Method of Computing the Input Quantization and Multiplication Roundoff Errors in a Digital Filter**"**, *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-22, No. 5, pp. 326-329, Oct. 1974.

[4] B. W. Bomar, **"**Computationally Efficient Low Roundoff Noise Second-Order Digital Filter Sections With No Overflow Oscillations", *IEEE Conference Proceedings Southeastcon '88*, pp:606 – 613, 11-13, April 1988.

[5] A. Wenzler and E. Luder, "New Structures for Complex Multipliers and Their Noise Analysis", *IEEE International Symposium on Circuits and Systems, (ISCAS'95)*, Vol. 2, pp. 1432 – 1435, 28 April - 3 May 1995.

[6] P. K. Sim and K. K. Pang, "Quantization Phenomena in a Class of Complex Biquad Recursive Digital Filters**",** IEEE Transaction on Circuit and Systems, vol. CAS-33, No.9, pp. 892-899, Sept. 1986.

[7] E. Watanabe and A. Nishihara, "A Synthesis of a Class of Complex Digital Filters Based on Circuitry Transformations". IEICE Trans., vol. E-74, No.11, Nov. 1991, pp.3622-3624.

[8] G. Stoyanov, M. Kawamata, Zl. Valkova, "New first and second-order very low-sensitivity bandpass/ bandstop complex digital filter sections", Proc. IEEE 1997 Region 10th Annual Conf. "TENCON'97", Brisbane, Australia, vol.1, pp.61-64, Dec. 2-4, 1997.

Some filtering schemes to obtain low aliasing spectrum response

Rumen Arnaudov[1], Rossen Miletiev[2]

Abstract: **Some analog and digital filtering schemes are investigated in the paper to be implemented in the systems with nonuniformly sampled data. The proposed filtering schemes are described regards to the hardware requirements, software realization, conversion speed and application field.**

Keywords: **nonuniformly sampled data, filtering schemes**

I. Introduction

Nonuniformly sampled data occurs in several applications such as geophysics [1], Laser Doppler Anemometry (LDA) [2], oscilloscopes [3] and radar or sonar signal processing [4]-[5]. Such type of data is used by the system designers to avoid aliasing in the signal spectrum or due to the technical problems, it is sometimes impossible to perform regular sampling. But filtering signals, which is nonuniformly sampled, is very difficult task, because the filtering coefficients are time varying [6]. The filtering task is more complicated problem when the signal bandwidth is very wide and the maximum spectrum frequency overcomes the Nyquist limit. In this case several spectrum estimation methods may be used to calculate the signal spectrum [7]-[9], but are distinguished with aliasing effects due to the non – orthogonal basis. The problem solution is concluded in the interpolation spectrum equation utilization [10] with applied filtering schemes to limit the signal bandwidth to the system Nyquist frequency.

The paper considers some possible analog and digital filtering schemes to overcome the aliasing effect problems in the systems, which use nonuniformly distributed grid.

II. Filtering Schemes

The type of the proposed filtering schemes depends basically on the system sampling frequency and signal bandwidth. When the bandwidth is smaller or equal to the equivalent sampling frequency *fs*, then only one analog filter may be used. The filtered signal values are converted to the digital words by analog–to–digital converter (ADC), which sampled the input signal in nonuniformly distributed grid. The ADC output data are analyzed by the microprocessor (μP) by using the digital spectrum shifting according to the equation:

$$u(t_n) = x(t_n)e^{-j\gamma t_n}, \tag{1}$$

where γ – shifting frequency

t_n – time sample points

The spectral shift analytical expression of the nonuniformly sampled data is implemented using the basic spectral shift equation to prove the equation (1):

$$U\left(e^{j2\pi\omega}\right) = X\left(e^{j2\pi(\omega-\gamma)}\right), \tag{2}$$

where $U(j\omega)$ – shifted spectral response

γ – frequency shift.

The nonuniformly sampled data $u(t_n)$ are calculated by well known inverse Fourier transform:

$$u(t_n) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} U(j\omega)e^{-j\omega t_n}d\omega \tag{3}$$

If the equation (2) is substituted at the definition equation (3), the nonuniformly sampled data $u(t_n)$ are estimated according to the equation:

$$u(t_n) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} X(j\Omega)e^{-j(\gamma+\Omega)t_n}d\Omega =$$
$$= \frac{e^{-j\gamma t_n}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} X(j\Omega)e^{-j\Omega t_n}d\Omega = x(t_n)e^{-j\gamma t_n}$$

When the signal bandwidth exceeded the equivalent sampling frequency *fs*, the more complicated filtered schemes have to be used to analyze the input signal. The following sections represent some analog and digital filtering schemes, which may be implemented to estimate the signal spectrum response.

The spectrum estimation procedure consists of the following steps:

1. The input signal is filtered by using the analog or digital filters to select the frequency response of the desired frequency band. The *i*-th filter passband is defined according to the expression $f_p = [(i-1)f_s; if_s], i = 1,\ldots,n$ (Fig.1), so the filter passband is equal to the sampling frequency;
2. If the signal is passed through analog bandpass filter, the input signal is converted to digital words by analog-to-digital converter. The ADC sampling points are nonuniformly spaced and the equivalent sampling frequency *fs* is below Nyquist limit, i.e. it accomplished undersampling process;
3. The spectrum shift procedure is performed by equation (1) to translate the analyzed frequency band to baseband;

[1] Rumen Arnaudov is with the Technical University of Sofia, 8 Kl. Ohridski Blvd, 1000 Sofia, Bulgaria; E-mail: RA@tu-sofia.bg

[2] Rossen Miletiev is with the Technical University of Sofia, 8 Kl. Ohridski Blvd, 1000 Sofia, Bulgaria; E-mail: miletiev@tu-sofia.bg

4. The frequency response is calculated by interpolation spectrum analysis of the nonuniformly sampled data to obtain low aliasing spectrum.

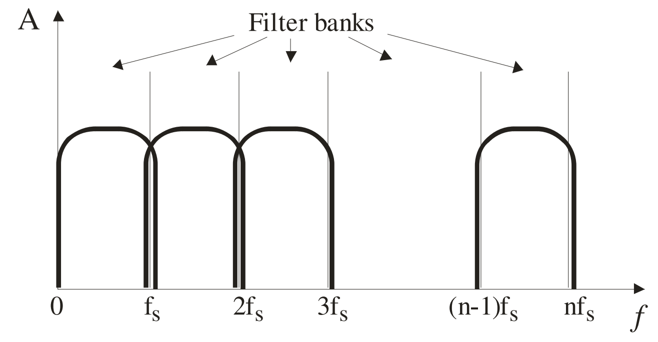


Fig. 1. Filter bank scheme

The next sections discuss the possible filtering schemes to realize the first step of the proposed procedures, while step 2 is analyzed at [11], step 3 is accomplished according to equation (1) and step 4 is discussed in our previous work [10].

2.1. *Analog filtering schemes* – The first proposed analog filtering scheme is defined as a parallel filtering scheme (Fig.2) and it contains multiple bandpass filters (BPFs) and analog-to-digital converters (ADCs). The number of the filter branches is calculated according to the equation:

$$n = \frac{B}{f_s}, \tag{4}$$

where B – signal bandwidth.

The ADC output data are sent to the µP via serial interface (SPI, I^{2}C, etc.). The microprocessor starts the conversion simultaneously for all ADCs to ensure identical time sample points at the circuit branches, but extracts the converted signal values consecutively.

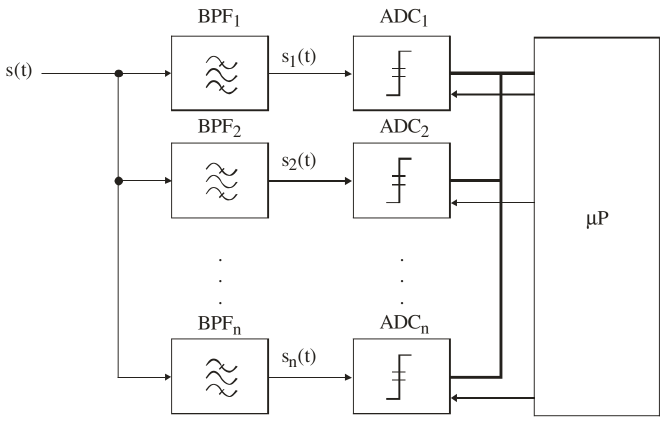


Fig. 2. Parallel filtering scheme

The proposed parallel filtering scheme may produce real time spectrum analysis of the input signal, but requires multiple hardware blocks to accomplish the digital signal processing (*n* BPFs and *n* ADCs). The number of the required hardware may be significantly reduced if multiplexed filtering scheme is used (Fig.3), because it requires only one ADC. The additional block is identified as a multiplexer (MUX), which switches the input signals to its output. At the expense of the reduced number of ADCs, this filtering scheme requires faster ADC regard to the parallel filtering scheme.

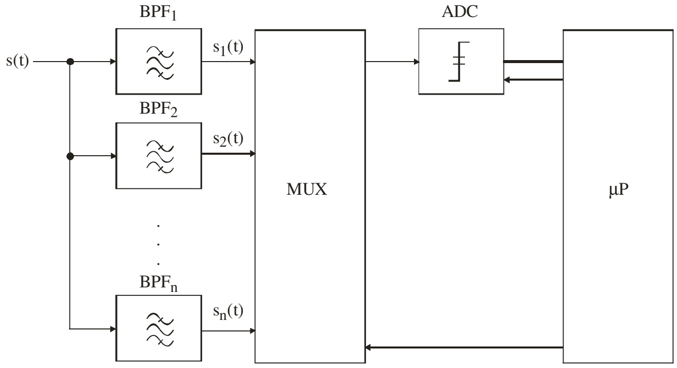


Fig. 3. Multiplexed filtering scheme

When the filter number exceeds the design requirements, then the programmable filtering scheme may be used (Fig.4). It is based on the implementation of the programmable BPF by the included programmable resistors. This scheme is distinguished as the most hardware saving system, but it is the slowest filtering scheme due to the required programming time. Regardless of the selected disadvantage it may be used at the vibration analysis systems (engines, constructions, etc.), where the signal processing time is not a critical parameter.

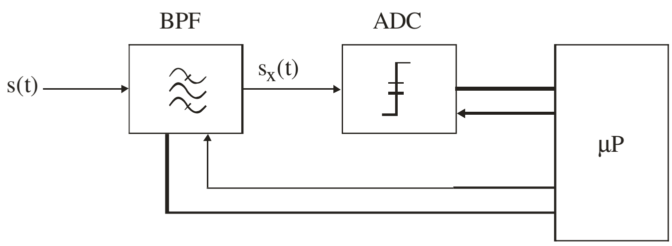


Fig. 4. Programmable filtering scheme

2.2. *Digital filtering schemes* – The digital filtering scheme (Fig.5) is a modern filtering scheme to produce the spectrum analysis of the nonuniform sampled signals. The main problem is concluded in the calculation of the filter coefficients since they have to be time varying [6]:

$$y(\tau) = a_{k\tau}s(t_k) + a_{k-1,\tau}s(t_{k-1}) + \ldots + a_{k-N,\tau}s(t_{k-N}) \tag{5}$$

When the analyzed signals are baseband and bandlimited, Tarczynski *at all* [6] proposed Weighed Least Squares (WLS) approach to solve the problem in the FIR filters. Otherwise, the solution of the selected problem remains unsolved yet. The published filter solution is non-applicable if the signal bandwidth exceeds the sampling frequency, which case is observed in our paper.

Due to its undeniable advantages regards to the analog filtering schemes (higher filter order, bandpass programmable capability, easy software implementation, etc.), the proposed filtering scheme is very suitable for the contemporary digital signal processing solutions.

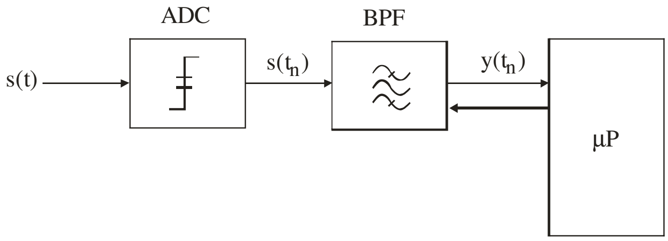


Fig. 4. Digital filtering scheme

Regardless of the selected advantage, the digital filtering scheme theory is not well developed yet, so it remains the unsolved problem at the signal processing field.

III. Conclusion

The nonuniformly sampled data is widely used in the modern digital processing systems, so the questions, connected with their spectrum analysis, are actual problems. The spectrum analysis above Nyquist limit is accomplished at four steps as the first one requires implementation of the filtering schemes to obtain low frequency aliasing spectrum response.

The current paper examines some of the possible analog and digital filtering schemes and indicates their main advantages and disadvantages. Also it describes their application fields to define the scheme enclosures. The contemporary requirements to the digital processing systems are connected to the software implementation of the filtering schemes, their power consumption, low price and integration. The most suitable solution of the shown filtering schemes to meet the selected requirements is identified as a digital filtering scheme. Unfortunately, the unsolved problems, connected with filter coefficients calculation, limit their application.

IV. References

1. P. Yuou, E. Baert and M. F. Loutre, "Spectral analysis of climate data," *Surv. Geophys.*, vol. 17, pp. 619–663, 1996.
2. C. Tropea, "Laser doppler anemometry: Recent developments and future challenges," *Meas. Sci. Technol.*, vol. 6, pp. 605–619, 1995.
3. Y. Rolain, J. Schoukens, and G. Vandersteen, "Signal reconstruction for non-equidistant finite length sample sets: A "KIS" approach," *IEEE Trans. Instrum. Meas.*, vol. 47, pp. 1046–1052, 1998.
4. R.Pribic, "Radar irregular sampling," *ICASSP 2004*, vol.3, pp. 933-936, 2004.
5. J. Koh, T. Sarkar and M. Wicks, 'Spectral analysis of nonuniformly sampled data using a least square method for application in multiple PRF system," *Proceedings of IEEE International Conference on Phased Array Systems and Technology 2000*, pp. 141-144
6. Andrzej Tarczynski, Vesa Valimaki, and Gerald D. Cain, "FIR filtering of nonuniformly sampled signals," *ICASSP 2004,* pp. 2237 – 2240
7. N. Lomb, "Least square frequency analysis unequally sampled data," *Aerophysics and space science*, vol.39, pp.447-462, May, 1975
8. M. Tuszynski and A. Wojtkiewicz, "Application of the Dirichlet transform in analysis of nonuniformly sampled signals," 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing ICASSP-92, May 23-26, 1992, San Francisco, USA, Vol. 5, pp.25-28
9. D. Bland, T. Laakso and A. Tarczynski, "Spectrum estimation of non-uniformly sampled signals," *Proceedings of the IEEE International Symposium on Industrial Electronics,* ISIE '96, Vol.1, 17-20 Jun 1996, pp. 196 -200
10. Miletiev R., "Interpolation method for spectrum analysis of nonuniformly sampled data," *National Military – Scientific forum 2004*, Velico Tyrnovo, 24-27 November 2004, vol.1, pp.83
11. Yin – Chyun Jenq, "Digital spectra of nonuniformly sampled signals: Fundamentals and High – Speed Waveform Digitizers," *IEEE Transactions on Instrumentation and Measurement*, vol.37, No.2, June 1988, pp.245-251

Image Pre-processing for IDP efficiency enhancement

Roumen Kountchev[1] Vladimir Todorov[2] Roumiana Kountcheva[3]

Abstract - **In the paper is presented a new method aimed at image quality and compression ratio enhancement for images processed with IDP decomposition. For this is performed image contrast pre-processing based on segmentation of the areas, containing relatively high density of the image brightness elements. The problem is solved changing the brightness intervals of the obtained segments followed by equalization of the corresponding parts of the histogram. The idea is to smooth slightly the sharp transitions in the image, which will decrease the number of the calculated transform coefficients with big values, retaining the visual quality of the treated image. The method permits easy adaptation of the contrasting algorithm in accordance with the image contents. As a result, for same IDP parameters the compression ratio is higher and the restored image quality – better. The obtained experimental results prove the efficiency of the new method.**

Keywords - **Image Compression, Inverse Difference Pyramid Decomposition, Image Contrast Enhancement.**

I. Introduction

The visual efficiency of an image compression technique depends directly on the amount of visually significant information it retains. One of the disadvantages of the image compression based on image decomposition with Inverse Difference Pyramid (IDP) (and of other compressing algorithms as well) is that for high compression ratios (more than 15) the restored image contains visible artifacts. By "visually significant" is usually meant information to which the human observer is most sensitive. The overall sensitivity depends on the image contrast, color, spatial frequency, etc [1,2,3]. One important aspect is the relationship between contrast sensitivity and spatial frequency, described by the contrast sensitivity function, which quantifies how well the Human Visual System (HVS) perceives the contrast at a given spatial frequency. As it is known, the HVS has nonlinear characteristics and is less sensitive for high frequencies [4,5]. This is particularly important point in the context of image compression: if in a compressed image strong artifacts are present, these artifacts are definitely far above the perception threshold. In order to improve the image quality is necessary to decrease the artifacts visibility to near-visibility-lossless rates.

One approach for solving the problem is to perform an image pre-processing, changing to some degree the image histogram in such a way, that to obtain restored images with better visual quality.

[1] Roumen Kountchev is with the FCTT, Radiocom. Dept., Technical University of Sofia, Bul. Kl. Ohridsky 8, Sofia, Bulgaria, E-mail: rkountch@tu-sofia.bg

[2,3] Vladimir Todorov and Roumiana Kountcheva are with T&K Engineering, Mladost 3, P.O.Box 12, Sofia 1712, Bulgaria
E-mails: todorov_vl@yahoo.com. kountcheva_r@yahoo.com

The new method for image pre-processing is presented below. The paper is arranges as follows: section II contains the method for image pre-processing; section III is a brief presentation of the IDP decomposition; in section IV are given some of the obtained experimental results; section V is the conclusion.

II. Method for Image Pre-processing with Adaptive Contrast Enhancement

The method is aimed at image pre-processing used to prepare the image contents for the IDP decomposition [6]. The offered method concerns adaptive image contrast enhancement, performed in two consecutive stages: (1) brightness segmentation based on image histogram analysis and (2) transformation of the pixels brightness in the so-defined segments in accordance with tables, defined using the segment histograms.

In the first stage is performed the image segmentation using the values k_1 and k_2, in result of which the image histogram is divided in three segments (A,B,C). It is supposed that the second segment (B) contains an image area in the brightness range (k_2-k_1), for which in the histogram exists clearly defined maximum. In case that there is more than one maximum with equal values, the range (k_2-k_1) corresponds to the one, closest to the dark (black) level. The objects, whose brightness values are in the range (k_2-k_1), have relatively low contrast and in order to increase it the sorresponding part of the histogram should be widened (stretched). Its limits k_1 and k_2 are calculated as follows:

- The maximum h(k) of the image histogram is identified:

$$h_{max} = \max\{h(k)\} \text{ for } k = 0, 1, 2, \dots, k_{max}, \qquad (1)$$

- The threshold value $t = \alpha\, h_{max}$, where $\alpha < 1$ (for example $\alpha = 0.9$) is defined;

- The end points of the segment k_1 and k_2 are defined in accordance with:

$$h(k) \le t \text{ for } k = 0, 1, 2, \dots, k_1\text{-}1, \qquad (2)$$

$$h(k) \ge t \text{ for } k = k_2+1, k_2+2, \dots, k_{max}, \qquad (3)$$

where $|k_2 - k_1| \ge \Delta$ (Δ – a value, set in advance).

In case, that the last condition is not satisfied the parameter α is decreased, for example to α=0.8, and with the new threshold value $t=\alpha.h_{max}$ are defined the corresponding values of k_1 and k_2. If the requirement $|k_2 – k_1| \ge \Delta$ is satisfied the calculation cycle is over, else it continues in the already described way.

In the second stage the brightness level k for every pixel in the three segments (A,B,C) is transformed in accordance with an individual table for every segment, as follows:

$$g(k)=\begin{cases} g_A(k) & \text{if} \quad 0\le k<k_1; \\ g_B(k) & \text{if} \quad k_1\le k\le k_2; \\ g_C(k) & \text{if} \quad k_2<k\le k_{max}. \end{cases} \qquad (4)$$

Here $g_A(k)$, $g_B(k)$ and $g_C(k)$ are the corresponding tables for brightness transform in the segments A, B and C. In order to increase the contrast of the low-contrast areas, the segment B is widened, moving the points k_1, k_2 in new positions $(k_1-\delta_1)$ and $(k_2+\delta_2)$, and correspondingly the upper limit point of the segment A is moved down and the lower limit of the segment C – up. In this case δ_1 and δ_2 are parameters, which define the contrast enhancement of the objects in the segment B. Each table for brightness transform calculation is defined in accordance with the histogram equalization relation for the corresponding segment A, B or C with changed (stretched or skewed) brightness range:

$$g_A(k)=(k_1-\delta_1)\sum_{l=0}^{k} h_A(l), \qquad (5)$$

$$g_B(k)=(k_2-k_1+\delta_1+\delta_2)\sum_{l=k_1-\delta_1}^{k} h_B(l)+(k_1-\delta_1), \qquad (6)$$

$$g_C(k)=(k_{max}-k_2-\delta_2)\sum_{l=k_2+\delta_2}^{k} h_C(l)+(k_2+\delta_2). \qquad (7)$$

In particular, when the histogram is uniform, i.e. for

$$h_A(k)=\frac{1}{k_1} \quad \text{for} \quad k=0,1,..,k_1-1; \qquad (8)$$

$$h_B(k)=\frac{1}{k_2-k_1} \quad \text{for} \quad k=k_1,k_1+1,..,k_2; \qquad (9)$$

$$h_C(k)=\frac{1}{k_{max}-k_2} \quad \text{for}\ k=k_2+1,k_2+2,..,k_{max}, \qquad (10)$$

then the table for brightness transform calculation of the pixels in the segments is linear and defined with the relations:

$$g_A(k)=\left(\frac{k_1-\delta_1}{k_1}\right)k; \qquad (11)$$

$$g_B(k)=\left(\frac{k_1-k_2-\delta_1-\delta_2}{k_1-k_2}\right)k+\left(\frac{\delta_2 k_1+\delta_1 k_2}{k_1-k_2}\right); \qquad (12)$$

$$g_C(k)=\left(\frac{k_{max}-k_2-\delta_2}{k_{max}-k_2}\right)k+\left(\frac{\delta_2 k_{max}}{k_{max}-k_2}\right). \qquad (13)$$

In this case the brightness levels in the range (k_1, k_2) are stretched following linear relation and accordingly the levels in the intervals $(0, k_1-1)$ and (k_2+1, k_{max}) are skewed.

In the cases, when $k_1 = 1$ or $k_2 = k_{max}$ the image histogram is divided in two segments A и B only, which are processed with similar way as the already described one for three segments. In this case the dynamic range of the segment A, containing the low-contrast object is stretched, and correspondingly the second segment is skewed. The contrast enhancement of color images in R, G, B format is performed after their transformation in Y, Cr, Cb format, after which the Y component is treated in accordance with the already presented way. After that the three Y, Cr, Cb components are transformed back in R,G, B format.

III. IMAGE COMPRESSION WITH TWO-LEVEL IDP

The essence of the method for grayscale image compression with two-level IDP decomposition is as follows. At first, the image matrix is divided in square sub-images. If every block is presented as a matrix $[B(8)]$ with size 8x8, after lossy compression it is approximated with the matrix $[\hat{B}(8)]$ with same size, presented as:

$$[\hat{B}(8)]=[\hat{B}_0(8)]+[\hat{E}_0(8)], \qquad (14)$$

where $[B_0(8)]$ and $[\hat{E}_0(8)]$ are matrices with size 8×8. The first matrix $[\hat{B}_0(8)]$ is the "zero" approximation of [B(8)], and the second one, $[\hat{E}_0(8)]$ is a difference matrix, representing the approximation error. The matrix $[\hat{B}_0(8)]$ for the "zero" IDP level is calculated using two-dimensional DCT transform:

$$[\hat{B}_0(8)]=[C(8)]^t[\hat{S}_0(8)][C(8)], \qquad (15)$$

where [C(8)] is a DCT matrix with size 8x8 and coefficients:

$$c(i,j)=A(i)\cos[\tfrac{1}{16}(2j+1)]\,i\,\pi], \qquad (16)$$

$$A(i)=\begin{cases}\sqrt{1/8} & \text{for} \quad i=0; \\ 1/2 & \text{for} \quad i=1,2,..,7.\end{cases} \qquad (17)$$

$$\hat{s}_0(u,v)=m_0(u,v)s'_0(u,v) \quad \text{for}\ u,v=0,1,..,7 \qquad (18)$$

are the coefficients of the truncated transform $[\hat{S}_0(8)]$, and

$$m_0(u,v)=\begin{cases}1 \ \text{if} \ (u,v)\in V_0; \\ 0\text{ - in other cases},\end{cases} \qquad (19)$$

are the elements of the binary matrix $[M_0(8)]$ with size 8×8, defining the area V_0 of the retained coefficients of the transform $[\hat{S}_0(8)]$.

In the relation (Eq.18) $s'_0(u,v)$ are the coefficients of the restored transform $[S'_0(8)]$, which is defined from the matrix [B(8)] of the original sub-block with two-dimensional direct DCT and quantization/dequantization of the obtained spectrum coefficients' values:

$$[S_0(8)]=[C(8)][B(8)][C(8)]^t, \qquad (20)$$

$$s_{0,q}(u,v)=Q_0[s_0(u,v)],$$

$$s'_0(u,v)=Q_0^{-1}[s_{0,q}(u,v)], \qquad (21)$$

Here $Q_0[\bullet]$ and $Q_0^{-1}[\bullet]$ are the corresponding operators for quantization/dequantization of the spectrum coefficients calculated for the sub-blocks in the "zero" IDP level.

The calculation of the matrix $[\hat{E}_0(8)]$ in the first IDP level starts with the calculation of the difference matrix:

$$[E_0(8)]=[B(8)]-[\hat{B}_0(8)]=\begin{bmatrix}[E_0^1(4)] & [E_0^2(4)] \\ [E_0^3(4)] & [E_0^4(4)]\end{bmatrix} \qquad (22)$$

The matrix $[E_0(8)]$ is divided in four sub-matrices and for each is calculated the corresponding transform, using two-dimensional direct Wash-Hadamard transform:

$$[S_1^k(4)] = [H(4)][E_0^k(4)][H(4)] \text{ for } k = 1, 2, 3, 4. \quad (23)$$

Here [H(4)] is a Wash-Hadamard matrix with size 4x4:

$$[H(4)] = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}. \quad (24)$$

The coefficients of the truncated approximated transform are calculated in accordance with:

$$\hat{s}_1^k(u,v) = m_1(u,v)s_1'^k(u,v), \text{ for } u,v = 0,1,2,3, \quad (25)$$

where

$$m_1(u,v) = \begin{cases} 1 \text{ if } (u,v) \in V_1; \\ 0 \text{ - in other cases,} \end{cases} \quad (26)$$

are the elements of the binary matrix $[M_1(4)]$ with size 4×4, defining the area V_1 of the retained coefficients in the $[\hat{S}_1^k(4)]$ transform.

In the relation (Eq.25) $s_1'^k(u.v)$ are the coefficients of the restored transform $[S_1'^k(4)]$, obtained after quantization and dequantization:

$$s_{1,q}^k(u,v) = Q_1[s_1^k(u,v)],$$
$$s_1'^k(u,v) = Q_1^{-1}[s_{1,q}^k(u,v)]. \quad (27)$$

Here $Q_1[\bullet]$ and $Q_1^{-1}[\bullet]$ are operators for quantization and dequantization of the spectrum coefficients in the first IDP level. From the transform $[S_1'^k(4)]$ with two-dimensional inverse Walsh-Hadamard transform is defined the matrix:

$$[\hat{E}_0^k(4)] = (1/16)[H(4)][S_1'^k(4)][H(4)]. \quad (28)$$

Then the matrix $[\hat{E}_0(8)]$ for level "one" of the decomposition (Eq.14) is obtained in accordance with (Eq.25) for k = 1, 2, 3, 4:

$$[\hat{E}_0(8)] = \begin{bmatrix} [\hat{E}_0^1(4)] & [\hat{E}_0^2(4)] \\ [\hat{E}_0^3(4)] & [\hat{E}_0^4(4)] \end{bmatrix}. \quad (29)$$

The values of the retained spectrum coefficients from same IDP levels for all sub-images are arranges in corresponding two-dimensional frequency bands:

- For the "zero" level each band with frequency (u,v) in the V_0 area is presented with the matrix $[S_{0,q}(u,v)]$, whose size m×n is defined in accordance with the sub-blocks number;
- For the level "one" each band (u,v) in the V_1 area is presented with the matrix with size m×n:

$$[S_{1,q}(u,v)] = \begin{bmatrix} [S_{1,q}^1(u,v)] & [S_{1,q}^2(u,v)] \\ [S_{1,q}^3(u,v)] & [S_{1,q}^4(u,v)] \end{bmatrix}, \quad (30)$$

where $[S_{1,q}^k(u,v)]$ for k = 1, 2, 3, 4 is sub-matrix with size (m/2)×(n/2).

The calculated spectrum coefficients from all bands in the two IDP levels build an inverse pyramid. The coefficients from each band are arranged sequentially in one-dimensional massif with length (4+64)×mn. The numbers in this massif are treated with adaptive RLE followed by modified Huffman coding and in result is obtained the compressed massif. The data is decompressed performing the already described operations in reverse order. The compression coefficients are controlled by the quantization coefficients set for every pyramid level and with the selected elements of the matrices $[M_0(8)]$ и $[M_1(8)]$.

IV. Experimental results

The software implementation of the presented method confirmed its efficiency. The experiments were performed with significant number of test images. Here are presented the results, obtained for the test image "Lena" (512×512 pixels, 8bpp). The changes in the IDP performance are significant and they concern the compression ratio and the quality of the restored image. The parameters used for the histogram segmentation are as follows: k_1=16, k_2=24 and the corresponding segment was given the value 10, i.e. the segment length was stretched approximately with 7%. The visual image quality is unchanged. The results represent the change of the restored image quality and the increased compression ratio for same IDP quality factors. In Table 1 are presented the results obtained for the compression ratios in the range from 6 to 65. The mean CR increase for the image "Lena" is 0.92; for some images (for example – test image "Fruits" this change is 3,5 and for "Peppers" it is more than 5).

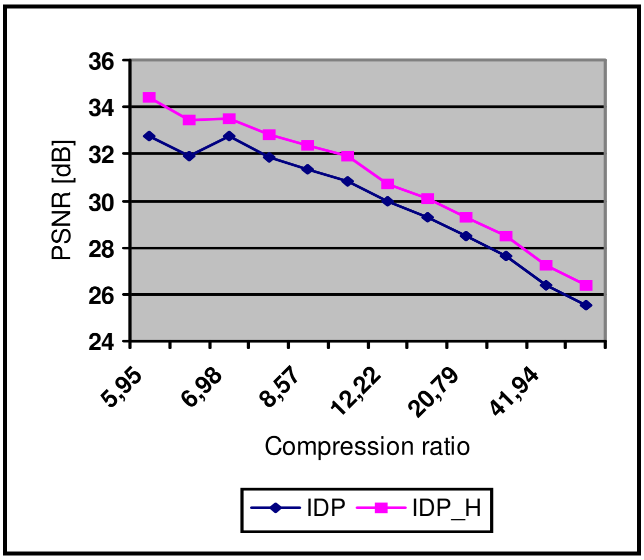


Fig. 1. Comparison of the obtained image quality for same compression ratios (Image "Lena")

For the research was used the 100-stage quality factor (QF) set used for the IDP decomposition research, defined by a number of parameters (QF1 is for best quality and smallest compression, QF100 – correspondingly for worst quality and highest compression ratio). This set of parameters comprises the number of pyramid levels, the approximation transform (DCT or WHT), the quantization values, the used transform coefficients (the binary matrices [$M_0(8)$] and [$M_1(4)$]), the scanning arrangement for same spatial frequencies, etc. In result of the histogram modification the quality of the restored images was improved in average with more than 1 dB. In Fig. 1 and 2 are presented the results obtained for same compression ratios and the improved quality for same quality factors. The characteristics of results, obtained for the other test images are similar with those for “Lena “. The test images « Lena » and « Fruits » are presented in Fig. 3.

TABLE 1. Influence of the Histogram Change on the CR.

QF	CR	CR_HM	Δ CR
40	5,95	6,39	0,44
45	6,35	6,77	0,44
50	6,98	7,46	0,48
55	7,82	8,29	0,47
60	8,57	9,03	0,46
65	9,23	9,72	0,49
70	12,22	12,68	0,46
75	16,21	16,74	0,53
80	20,79	21,58	0,79
85	30,43	31,17	0,74
90	41,94	43,39	1,45
95	65,26	68,00	2,75
100	78,16	80,66	2,50

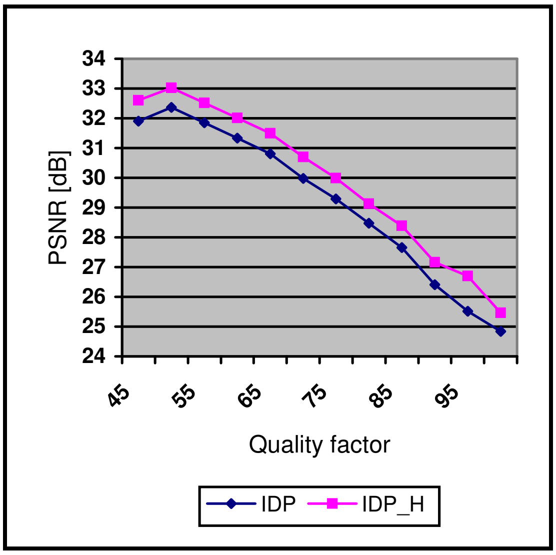


Fig.2. Influence of the histogram modification on the PSNR (Image “Lena”).

Fig. 3. Test images “Lena” and “Fruits”

V. Conclusion

The experimental results proved the efficiency of the proposed approach. In result of the performed research work the following conclusions are defined:

- The histogram modification influences the IDP compression in all range of the settled quality factor (QF) set;
- The compression ratio (CR) for all tested images was increased;
- The quality of the restored after IDP compression image is better;
- The dual influence (CR/image quality) practically moves the QF scale with more than 4 positions, i.e. for example, the performance for QF50 for the test image “Lena” in result of the pre-processing is equivalent to the performance of QF46 regarding the restored image quality and to the performance of QF54 regarding the CR. The results for other test images like “Fruits” and “Peppers” are similar;
- The new method proved its qualities for grayscale images. The results, obtained for color images are good too, but as it is known, the characteristics of the HVS for color images are different from those for grayscale ones and detailed research should be performed specially for such applications.
- The method for histogram modification accordingly the image contents could be used for better contours extraction.

VI. Acknowledgement

This paper was supported by NSF of Bulgarian Ministry of Education and Science (Contract ВУ-МИ-104).

References

[1] P. Barten. “Contrast sensitivity of the human eye and its effects on image quality”. Bellingham, WA: SPIE, 1999.

[2] Y. Lai, C. Kuo. “Wavelet image compression with optimized perceptual quality”. Conference on Applications of Digital Image Processing XXI. San Diego, CA: SPIE, 1998.

[3] M. Nadenau, M. Kunt. “Wavelet-based color image compression: exploiting the contrast sensitivity function”. IEEE Transactions on Image Processing, Vol.12, No.1, January 2003, pp. 58-70.

[4] S. Winkler, P. Vanderdheynst. “Computing isotropic local contrast from oriented pyramid decompositions”. Proc. 6th Int. Conf. Image Processing, Kobe, Japan, Oct. 1999, pp. 420-424.

[5] M. Nadenau, J. Reichel. “Compression of color images with wavelets under consideration of the HVS”. Proc. SPIE, Jan. 1999, vol. 3644.

[6] R. Kountchev, M. Milanova, C. Ford, R. Kountcheva. “Multi-layer Image Transmission with Inverse Pyramidal Decomposition”. In “Computational Intelligence for Modeling and Predictions”, S. Halgamuge, L. Wang (Eds.), Vol. 2, Chapter 13, Springer-Verlag Berlin Heidelberg, 2005, pp. 179–196

Post-Processing for Quality Improvement of IDP-Compressed Images

Roumen Kountchev[1] Vladimir Todorov[2] Roumiana Kountcheva[3]

Abstract - **In the paper is offered new approcah for the reduction of the blocking artifacts in still images, processed with 2-level Inverse Difference Pyramid (IDP) decomposition, based on Walsh-Hadamard and DCT orthogonal transforms. For this purpose was developed a two-dimensional fuzzy digital filter, whose performance changes in accordance with the image contents, framed by the filter window, and depending on the compression ratio and the maximum approximation error, obtained in the second (higher) IDP level. The experimental results show, that the block artifacts in the restored images are reduced without visual deterioration of the image sharpness. The advantages of the new filter are its low computational complexity and adaptation abilities. In the paper are analyzed the abilities for the filter implementation aiming at the quality improvement of decompressed color images with blocking artifacts.**

Keywords – **Deblocking filter, Two-dimensional Adaptive Fuzzy Filter (2DAFF), Image Compression, Inverse Difference Pyramid (IDP), Discrete Cosine Transform (DCT).**

I. Introduction

The block-transformed techniques for image coding, based on orthogonal transforms, usually generate significant distortions, called blocking artifacts [5,16] and this effect is accentuated at high compression ratios. These artifacts display themselves as artificial boundaries between adjacent blocks or around sharp transitions in the processed images. The distortions of the second kind, called ringing artifacts, are a result of the Gibbs effect and are due to the quantization of the used transform coefficients. In order to minimize the artifacts in the decompressed images are already developed significant number of methods for pre- and post – processing [2-13]. The methods in the second group could be classified as follows: Direct linear or non-linear smoothing techniques in the spatial domain [7-10]; Combined techniques employing both edge detection or segmentation for detail classification and spatial adaptive filtering [11]; Iterative techniques based on the theory of projections on to convex set (POCS) [6,12], and Soft threshold approaches in the wavelet domain [13]. The major issues existing in the current post-processing methods can be summarized in brief as: limitation to a certain type of artifacts, concerning the methods in the first group, mentioned above, and such with high computational complexity – included in the remaining three groups.

[1] Roumen Kountchev is with the FCTT, Radiocomunications Dept., Technical University of Sofia, Bul. Kl. Ohridsky 8, Sofia, Bulgaria, E-mail: rkountch@tu-sofia.bg

[2,3] Vladimir Todorov and Roumiana Kountcheva are with T&K Engineering, Mladost 3, P.O.Box 12, Sofia 1712, Bulgaria
E-mails: todorov_vl@yahoo.com. kountcheva_r@yahoo.com

In the paper is offered a relatively simple and efficient technique for removing various blocking artifacts using a two-dimensional fuzzy digital filter, which is quite adequate for the processing of decompressed images, obtained using IDP decomposition, based on DCT and Walsh-Hadamard transforms. In Section II is presented the algorithm for fuzzy filtration, in Section III is described the approach, used for the adaptation of the offered filter, in Section IV are presented the results of the modeling of the filter performance for test images compressed using the IDP decomposition and in the Conclusion are pointed the advantages of the new filter and its possible applications.

II. Post-processing with Fuzzy Adaptive Filter

The offered approach for post-processing of decompressed images is based on the use of fuzzy digital filters [1,3], which became very popular recently. The algorithm for the performance of the two-dimensional fuzzy adaptive filter (2DFAF), using a sliding window with size MxN pixels (M=2R+1 and N=2S+1), is as follows:

$$x_F(i,j)=\begin{cases}\left\lfloor \dfrac{\sum_{r=-R}^{R}\sum_{s=-S}^{S}\mu(i+r,j+s)\,x(i+r,j+s)}{\sum_{r=-R}^{R}\sum_{s=-S}^{S}\mu(i+r,j+s)}\right\rfloor & \text{for } \sum_{r=-R}^{R}\sum_{s=-S}^{S}\mu(i+r,j+s)\ge T,\\ \left\lfloor (1/MN)\sum_{r=-R}^{R}\sum_{s=-S}^{S}x(i+r,j+s)\right\rfloor & \text{- in all other cases,}\end{cases} \tag{1}$$

where $\lfloor\circ\rfloor$ is a rounding operator; T – a noise threshold $x(i,j)$ and $x_F(i,j)$ are correspondingly the pixels of the input and of the filtered output image;

$$\mu[\Delta(i+r,j+s)]=\begin{cases}1 & \text{for } \Delta(i+r,j+s)\le\alpha;\\ \dfrac{\Delta(i+r,j+s)-\alpha}{\alpha-\beta} & \text{for } \alpha\le\Delta(i+r,j+s)\le\beta;\\ 0 & \text{for } \Delta(i+r,j+s)\ge\beta,\end{cases} \tag{2}$$

Here Eq. 2 represents the chosen membership function with parameters α and β (β > α), whose argument Δ is the module of the difference between the central pixel x(i,j) in the filter window and the pixel x(i+r, j+s), moved at the distance (r, s):

$$\Delta(i+r,j+s)=\left|x(i,j)-x(i+r,j+s)\right| \tag{3}$$

for $r=\overline{-R,+R}$ and $s=\overline{-S,+S}$.

The values of parameters α and β are defined in accordance with the image contents and with the kind of the distortions, which should be corrected or suppressed. In the case, when block artifacts obtained in result of the high image compression are concerned, α and β are defined in depending on the compression ratio and the kind of the used compression algorithm. In the presented approach the 2DFAF filter is applied on images with block artifacts obtained in result of lossy compression based on Inverse Difference Pyramid (IDP) decomposition [14] with high compression ratio.

III. Adaptation of the 2DFAF filter

For the suppression of the block artifacts in images obtained in result of the IDP compression parameters α and β of the 2DFAF filter are defined in accordance with the relations:

$$\alpha = \delta - \varepsilon, \quad \beta = \delta + \varepsilon, \tag{4}$$

where δ identifies the center of the filter fuzziness, for which the function $\mu(\Delta) = 0.5$, and ε defines the boundaries of the deviation δ. The value of the parameter δ is defined following the condition:

$$\delta = \frac{1}{2}\left| \hat{E}_0(i,j)_{max} \right| . \tag{5}$$

Here $E_0(i,j)_{max}$ is a pixel of the matrix $[E_0]$ containing the error between the original image and its approximation in the second IDP level. The parameters ε and T of the 2DFAF filter are chosen experimentally in accordance with the compression ratio and the noise level.

IV. Experimental results

The experiments were performed with parameters ε = 1 and T=2. The research on the new filter was performed with the software, implementing the algorithm. The filter treats the brightness component of the image only. For the investigation were used large number (more than 100) grayscale and color images. In the investigation was used the 100-stage quality factor (QF) set used for the IDP decomposition, defined by a number of parameters (QF1 is for the best quality and the smallest compression, QF100 – correspondingly for the worst quality and the highest compression ratio). The set of parameters used for the definition of the QF stages comprises the number of pyramid levels, the approximation transform (DCT or WHT) used in the pyramid levels, the quantization values, the used transform coefficients, the scanning arrangement of the values, calculated for same spatial frequencies coefficients, etc.

Some of the obtained results are presented below. In Fig. 1 are given the mentioned test images “Lena”, “Peppers” and “Fruits”. In Fig. 2 are presented the results obtained for same test images after compression with quality factors QF: 40,50 and 70 and treating the images with filter, whose window is with size 3x5 pixels and whose center of fuzziness changes from 5 to 65. The vertical axis shows the change of the image quality measured as PSNR. One unit in the vertical axis is 0,01 dB.

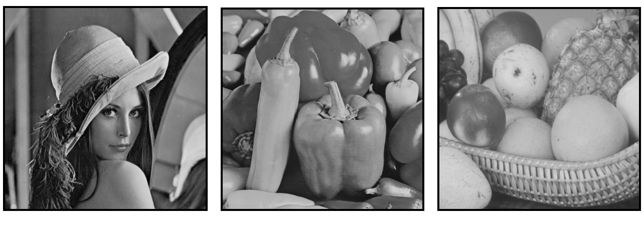


Fig. 1. Test images: “Lena”, “Peppers”, “Fruits”(originals)

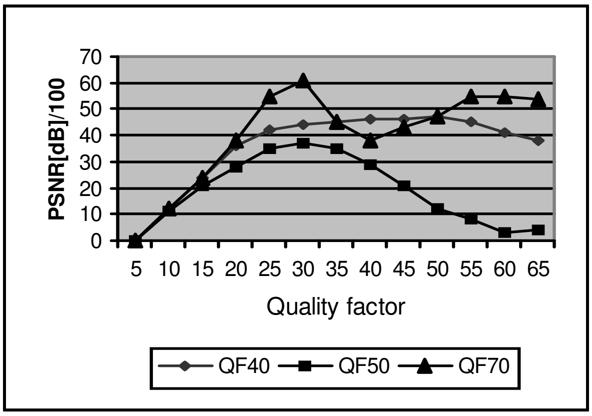


Figure 2a. “Peppers”.

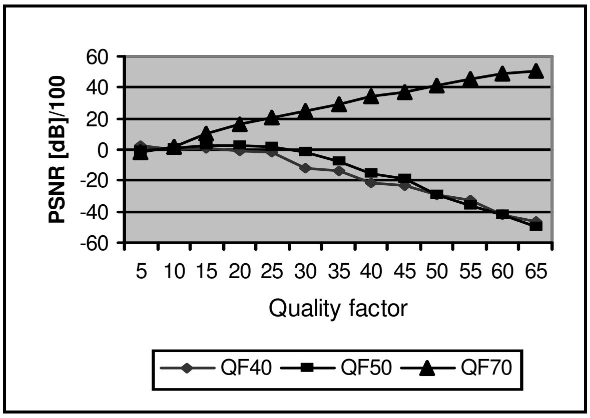


Figure 2b. “Lena”

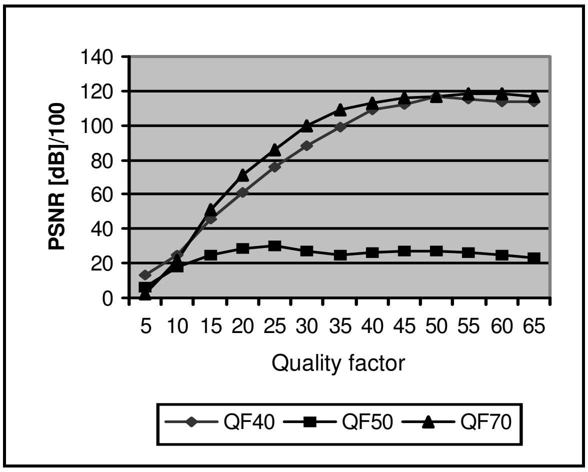


Figure 2. c. “Fruits”

The results show that the results are best for images restored after high IDP compression (corresponding with QF70 for these examples). Together with the increasing of the QF the filter influence increases as well.

In Fig, 3 are presented the results obtained for the test image “Lena” changing the quality factor from 50 to 100 and treating the restored image with filter window with width 3 and 5 pixels (the window height is 5 pixels in both cases). The results obtained for the other test images are similar. Best results are obtained for filter with window width 3 pixels. The vertical axis (for which one unit is equal to 0,01 dB) represents the image quality change.

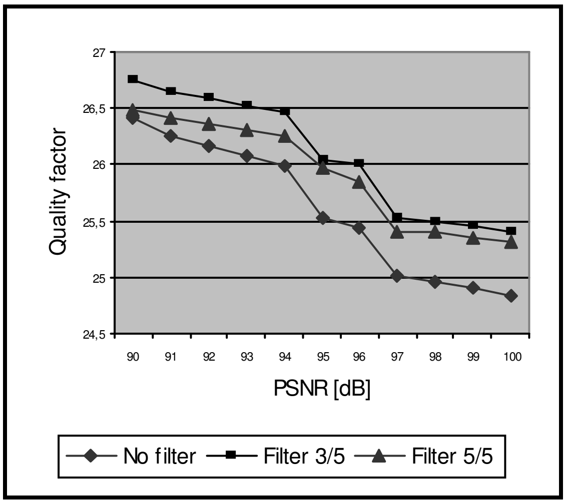


Fig. 3. Influence of the window filter width change (Image "Lena")

These results are in correspondence with the IDP decomposition approach and they suit the size of the sub-image in the higher pyramid level very well (as it was already mentioned, the sub-images in the upper pyramid level are with size 4x4 pixels).

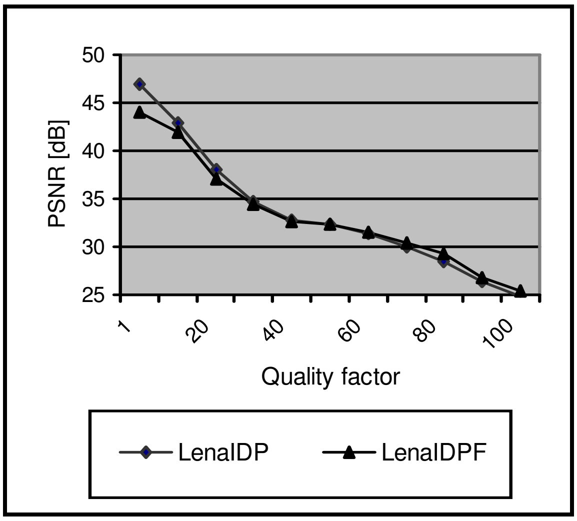


Fig. 4. Image Lena, IDP coding

In Figure 4 are presented the results obtained for the filter performance for the full 100-stage IDP quality factor range. The horizontal axis represents the QF and the vertical – the restored image quality (one unit corresponds to 0,01 dB). The results show that for low QF (i.e. low compression ratio), when the PSNR of the restored image is high (above 45 dB) the filter decreases the image quality. Its performance is useful for QF higher than 50, i.e. when the restored image quality is less than 30 dB.

In Fig. 5 is presented the result of the filtration for part of the image "Pepper" after IDP compression with QF = 86 (Compression ratio = 35). The visual quality improvement is noticeable. Similar results were obtained for the remaining test images. General result for all tested images is that for QF over 90, where the image quality is significantly decreased, better visual results are obtained with filter with window width 7 pixels. The PSNR of the filtered image is lower for window width 7, but the visual quality is much better. These results correspond with the IDP decomposition, because for these QF values the image sub-blocks are with size 8x8 pixels.

Fig. 5a. Decompressed image "Peppers" (QF86, CR=35)

Fig. 5b. The same decompressed image after filtration

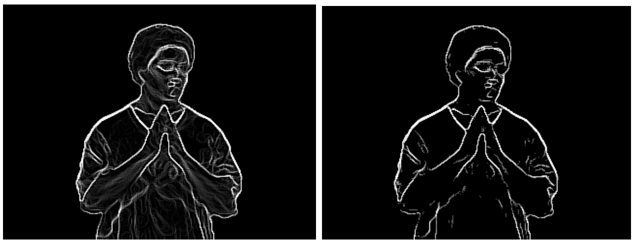
In Fig. 6a,b is presented the decompressed image "Lena" after IDP compression with QF 100 and the same image after filtration.

a b

Fig. 6. Test image "Lena" restored after compression with QF100 and filtration with filter, which center of fuzziness is 65 and the filter window width is 7 (the filter height is 5).

The general result from the performed investigation is that the IDP compression method gives enough information for setting the values of the filter parameters in the process of the image coding. As it was already pointed, the main parameters are its center of fuzziness and the window width. Best results are obtained when the filter center of fuzziness is set to be equal to the half of the maximum difference calculated between the original and the restored image in the coding side. The experiments proved that the filter window width should be 3 for all QF in the range of 50 to 90 and 7 for the range of 90 to 100.

The filter gives very good results for treatment of multi-level contour images, in particular - their adaptive segmentation retaining maximum visual similarity with the original. In result of the filtration the contours extraction is easier and the compression - more efficient. The detailed presentation of the research results on contour images will be presented in another work.



a b

Fig. 7a. Contour image (contours were extracted with Corel Photo Paint 10, Edge detection); b. Same image, after filtration.

Example of the obtained results is presented in Fig. 7. In Fig. 7a is presented the original contour image, and the filtered image is presented in Fig. 7b. It is easy to notice that the number of contours in the filtered image is smaller, retaining the most important details.

V. CONCLUSION

The experimental results proved the efficiency of the proposed approach. In result of the performed research work are defined the following conclusions:

- The two-dimensional adaptive digital filter improves the quality of decompressed images, compressed to high degree with IDP compression;
- The filter parameters suit very well the IDP decomposition, because there is a definite relation between them and the IDP quality factor;
- The filter parameters are defined in the process of the image compression and the corresponding data is included in the obtained compressed data file. This additional information is negligible, because the filter needs two parameters only: the center of fuzziness and the filter window width, i.e. two bytes only.
- The filter improves the quality of the decompressed image in the whole range from QF 50 to QF 100.
- The mean quality improvement for the tested images is 0,4dB, but the visual quality improvement is very good.
- The filter has low computational complexity and its implementation in software solutions based on the IDP image decomposition does not slow-down their performance.

VI. ACKNOWLEDGEMENT

This paper was supported by NSF of Bulgarian Ministry of Education and Science (Contract ВУ-МИ-104).

REFERENCES

[1] M. Nachtegael et all. “Fuzzy Filters for Image Processing”, Springer-Verlag, 2003.

[2] Ch. Tu, T. Tran, J. Liang, “Error Resilient Pre-/Post Filtering for DCT-Based Block Coding Systems”, Proc. of the 2003 IEEE Intern. Conf. on Acoustics, Speech, and Signal Processing, Hong Kong, Vol. III, pp. 153-156, April 2003.

[3] Y. Nie, K. Barner, “Optimized Fuzzy Transformation for Image Deblocking”, IEEE ICME, Vol. I, pp. 541-544, March 2003.

[4] S. Minami, A. Zakhor, “An optimization approach for removing blocking effects in transform coding,” IEEE Trans. on Circuit and System for Video Technology (CSVT), Vol. 5, No. 2, pp. 74–82, Apr. 1995.

[5] T. Meier, K. Ngan, G. Crebbin, “Reduction of coding artifacts at low bit rates”, Proc. SPIE Visual Communications and Image Processing, San Jose, January 1998.

[6] Y. Yang, N. Galatsanos, “Removal of compression artifacts using projections onto convex sets and line modeling”, IEEE Trans. on Image Processing, 6, October 1997.

[7] Y. Yang, N. Galatsanos, A. Katsaggelos, “Projection-Based Spatially-Adaptive Reconstruction of Block Transform Compressed Images", IEEE Trans. on Image Processing, Vol. 4, No. 7, pp. 896-908, July 1995.

[8] S. Kim, J. Yi, H. Kim, J. Ra, “A Deblocking Filter with Two Separate Modes in Block-Based Video Coding,” IEEE Trans. CSVT, Vol. 9, No. 1, Feb. 1999.

[9] J. Chou, M. Crouse, K. Ramchandran, “A Simple Algorithm for Removing Blocking Artifacts in Block-Transform Coded Images,” IEEE Signal Processing Letters, Vol. 5, No. 2, Feb. 1998.

[10] S. Masri, R. Castangno, G. Ramponi, “A Simple Algorithm for the Reduction of Blocking Artifacts in Images and its Implementations”, IEEE Trans. on Consumer Electronics, 44, August, 1998.

[11] A. Al-Fahoum, A. Reza, “Combined Edge Crispiness and Statistical Differencing for Deblocking JPEG compressed Images”, IEEE Trans. on Image Processing, Vol. 10, No. 9, Sept. 2001, pp. 1288-1298.

[12] R. Rosenholtz, A. Zakhor, “Iterative Procedure for Reduction of Blocking Effects in Transform Image Coding”, IEEE Trans. CSVT, Vol. 2, Mar. 1992, pp. 91-95.

[13] S. Wu, H. Yan, Z. Tan, “An Efficient Wavelet-Based Deblocking Algorithm for Highly Compressed Images,” IEEE Trans. CSVT, Vol. 11, No. 11, Nov. 2001, pp. 1193-1198.

[14] R. Kountchev, M. Milanova, C. Ford, R. Kountcheva. “Multi-layer Image Transmission with Inverse Pyramidal Decomposition”. In “Computational Intelligence for Modeling and Predictions”, S. Halgamuge, L. Wang (Eds.), Vol. 2, Springer-Verlag, pp. 179-196, 2005.

[15] R. Kountchev, M. Milanova, C. Ford, S. Rubin, R Kountcheva. Lossless image compression with IDP and Adaptive RLC. Proc. of Intern. Conf. on Imaging Science, Systems and Technology (ISST’04), Las Vegas, USA, June 2004,pp. 608-612.

[16] I. Richardson, “H.264 and MPEG-4 Video Compression - Video Coding for Next-generation Multimedia”, John Wiley & Sons Ltd, 2003.

A Method for Digital Image Compression with IDP Based on Gaussian Radial Basis Function NNs

Noha A. Hikal[1], Roumen Kountchev[2], R.M El-Awady[3], A.A Shaalan[4]

Abstract – **In this paper a new algorithm for still image compression based on Gaussian Radial Basis Function NNs in correspondence with Inverse Difference Pyramid (IDP) decomposition is represented. The new algorithm is well suited to be used in Progressive Image Transmission (PIT). The method advantages relay on the learning process and on the adaptation capability of the NNs to reduce the matrices computation complexity, the total number of pyramid levels required for PIT, and to maximize the PSNR. In addition to, for image reconstruction no interpolation is needed for the improvement of the reconstructed image quality.**
Keywords – **IDP Decomposition, PIT. Neural Networks, Gaussian Radial Basis Function.**

I. INTRODUCTION

The image compression is an important tool to store and transmit visual information used for multiple applications. The image compression refers to a process in which the amount of data used to represent an image is reduced to meet a bit rate requirement (below or at most equal to the maximum available bit rate), while the quality of the reconstructed image satisfies a requirement for a certain application and the complexity of the computation involved is affordable for the application. The progressive Image Transmission (PIT) [1] concept is of particular importance in browsing large image files. Hence, it permits the initial reconstruction of the image approximation and gives the viewer an early impression of image contents, followed by gradual quality improvement of the reconstructed image. The observer may terminate the transmission of an image as soon as its contents are recognized. In order to send the image data progressively, it should be organized hierarchically in the order of importance, from the global image characteristics to its local details. There are two types of data structures for progressive transmission depending on the encoding method employed [2]: (i) Transform-based encoding, (ii) Spatial encoding. In the transform-based encoding, the image is first divided into a set of contiguous non-overlapping blocks, and then each block is transformed into a set of transform coefficients, (e.g, Discrete Cosine Transform (DCT). On the other hand the spatial approach, like the pyramid-level resolutions, the image is successively reduced in spatial resolution and size by sub-sampling or averaging. The image approximation is obtained using a single frame or a combination of frames representing the image. Therefore sending a set of image frames in pyramid form from top to bottom level naturally constitutes a progressive transmission.

[1]N. A. Hikal is with Special Studies Academy Cairo, Eypt, Eng_noha_hikal@yahoo.com
[2]Roumen Kountchev is with FCTT, Radiocomm. Dept. Technical University of Sofia, K1. Ohridsky 8, Sofia, Bulgaria, rkountch@tu-sofia.bg
[3]R. M El-Awady is with Faculty of Engineering, Comm. Dept., El-Mansoura University, Egypt
[4]A. A Shaalan is with Faculty of Engineering, Comm. Dept., El-Zagaaig University, Egypt, drshaalan2006@yahoo.com

II. PYRAMIDAL IMAGE REPRESENTATION

The first pyramidal data structure is the Gaussian-Laplacian pyramid (GP/LP) [3]. The GP can be viewed as a set of low pass filtered copies of the original image, while in the LP each level is a difference between two successive levels of the GP. Various pyramid data structures for PIT have been proposed like: Mean Pyramid, Reduced Sum Pyramid, Reduced Difference Pyramid, Hierarchy Embedded Differential Image, etc [4,5,6]. Many pyramidal decomposition techniques showed improvements over the JPEG standard [7]. The Inverse Pyramidal Decomposition, based on Discrete Cosine Transform (IDP/DCT), was proposed in [8]. The IDP decomposition differs in the way of obtaining the pyramid levels, and the word "inverse" refers to the requirement to compute the pyramid levels from pyramid top (level zero) to the bottom. The coding/decoding processes are simple and flexible. Therefore, it is suitable for real time processing and serves PIT requirements. NNs are interesting alternatives to classical image compression techniques due to their lower matrices computational complexity, adaptation and learning capability. In [6,9] two algorithms for image compression based on IDP using BPLA and 2D-SOM VQ NNs, respectively, were presented. The algorithm showed superiority over IDP/DCT in terms of PSNR, CR, and the total number of levels required for the image reconstructing at the receiver side. In this work, a new technique for image compression based on merging IDP and GRBF NNs is presented, combining the advantages of IDP and NNs methods. The Gaussian networks are a special RBF type, characterized by being highly nonlinear and providing good locality for incremental learning. This paper is organized as follows: Section (3) introduces a definition for the basic GRBF NNs. Section (4) explains the new developed algorithm for image compression based on IDP-GRBF NNs. Sections (5) and (6) represent the performance measurements criteria and the simulation results respectively. Finally Section (7) gives a conclusion on the results of the new algorithm.

III. BASIC GAUSSIAN RBF NNS MODEL

The Gaussian RBFNNs [10] is the most important RBFN class. The Basic model employs for its RBF an un-normalized form of Gaussian density function presented as:

$$\phi(r) = \exp\left(-\left(\frac{r}{\sigma}\right)^2\right) \qquad (1)$$

It has a peak at the center r=0 and decreases monotonically as the distance from the center increases. The basic Gaussian network, Fig.1, consists of a three stages network [11] with input, intermediate and output stage. The intermediate stage consists of an array of nodes c_i that contain center vectors. These nodes calculate the Euclidean distance between the centers and the network input vector and the results are passed through a Gaussian function. The output stage of the neuron is a linear combiner used to obtain the weighted summation of the intermediate outputs. Let X ($x\in R^n$) and Y ($y\in R^m$) be the n-input and the m-output of the network, respectively, and u ($u\in R^l$) be the l outputs of the l Gaussian units. A Gaussian radial basis function Φ_i is assumed to be the same for all units. The connections between the nodes in the hidden layer and the j^{th} output (y_j; here: j=1,2…m) through the linear combiner unit are shown in Figure 1 [11].

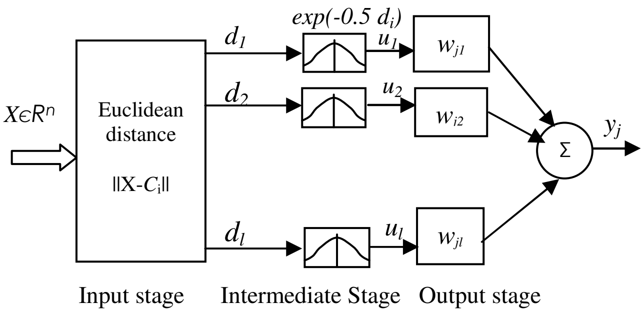


Fig. 1. Block diagram of a Gaussain RBF

Here the term d denotes the difference computed as:

$$d_i = \sum_{k=1}^{l} [(x_i - c_{ik}) / \sigma_{ik}^{\ 2}] \tag{2}$$

and σ^2 is the variance controlling the width of the Gaussian function; u_i is the output of the i^{th} hidden Gaussian neuron, presented with the relation [11]:

$$u_i = \exp\left(-\frac{1}{2}\sum_{k=1}^{n}\left[\frac{x_k - c_{ik}}{\sigma_{ik}}\right]^2\right), \quad 1 \le i \le l \tag{3}$$

Therefore, the input-output relationship of a GRBFNN with multiple outputs y_j ,$1\le j\le m$,can be described mathematically by the following equation:

$$\begin{aligned} y_j &= \sum_{i=1}^{l} w_{ji} u_i \\ &= \sum_{i=1}^{l} w_{ij} \exp\left(-\frac{1}{2}\sum_{k=1}^{n}\left[\frac{x_k - c_{ik}}{\sigma_{ik}}\right]^2\right) \end{aligned} \tag{4}$$

With the centers and variance parameters that are initialized to nearly optimum values, the learning task (which is described by the input/output data pairs $\{x,y_d\}$) will follow the gradient-descent method to form updating equation for the unknown weights. Then the instantaneous value of the cost function, which should be minimized ill be:

$$E = \frac{1}{2}\sum_{j=1}^{m} (y_{dj} - y_j)^2 = \frac{1}{2}\sum_{j=1}^{m} e_j^2 \tag{5}$$

Where:

$$\begin{aligned} e_j &= y_{dj} - y_j \\ &= y_{dj} - \sum_{i=1}^{l} w_{ij} \exp\left(-\frac{1}{2}\sum_{p=1}^{n}\left[\frac{x_k - c_{ip}}{\sigma_{ip}}\right]^2\right) \end{aligned} \tag{6}$$

Using both the above definitions and the network equation (4), the weight updating equation used to minimize the cost function could be derived as follows [11]:

$$w_{ij}^{new} = w_{ij}^{old} + \eta u_i e_j \tag{7}$$

Where η is positive learning rate constant which controls the learning rate, convergence, and stability.

IV. Image Compression Algorithm Using GRBF NNs Based on IDP Decomposition

The basic steps for image compression using GRBF NNs based on IDP coding/decoding process are introduced in the following two subsections.

Coding process

Step 1: For the first pyramid level (P=0), the whole image B(i,j) of size 2^nx2^n, $1\le$ i, j $\le 2^n$, is divided into a number of l blocks, each of them of size (m x m) pixels.

Step 2: A GRBF network will be created with an input vector of length m^2x1; an intermediate stage of the Gaussian network consisting of an array of center vectors c_i (i=1,2,.....l; total number of blocks in the whole image), the RBF center vectors are set to the values of the pixels corresponding to each block. The total l outputs from the intermediate stages feed the Linear Combiner (LC) stage to be multiplied m times by the randomly initialized weight vectors $W_{j\cdot} = [w_{j1}\ w_{j2}.......w_{jl}]^T$; $1\le j \le m$. The output vector y of length mx1 will be compared with the desired output vector y_d (mx1), which is set to the average vector of each block. Fig. 2 shows a block diagram for a GRBF network for block compression.

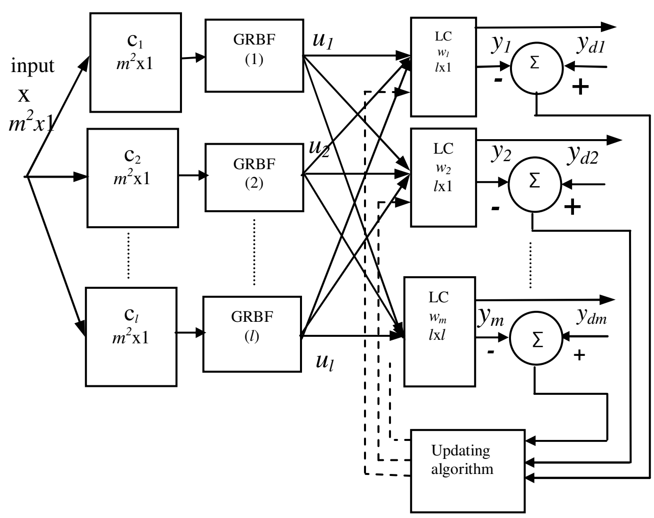


Fig. 2. GRBF network for block compression

Step 3: Applying the training input–output pairs $\{x, y_d\}$ (which are the image blocks themselves) to update the weights of the network. Once the training is complete, the image compression is demonstrated in the recall phase, using row scanning, pixels of each block are arranged as a vector of length (m^2) and these (m^2) components are considered to be the input vector of the GRBF-NN. The obtained weights matrix (lxm) forms the coefficients of the first pyramid level (p=0). The weights matrices will be quantized, encoded and transmitted later. So instead of sending the whole image of size 2^nx2^n, only the weights matrix of size (lxm) per block will be send. For example, if the original image of size 128x128 was divided into 4 blocks each of size 64x64, after compression using the first level of the GRBF-NN, only 4x64 coefficients per block will be sent, this is corresponding to compression ratio of (128x128)/(4x64x4)=16, for further compression encoding techniques may be applied. This form of reduction in the number of pixels is considered a spatial reduction, which is used in PIT.

Step 4: At the transmitter side, the reconstructed image will be recovered from the weights matrices (after inverse quantization and decoding process) in reverse arrangement using input vector of size (lmx1), (m) intermediate stages, output vector of length (m^2) in Fig. 3, the center vectors are set to the weight matrices corresponding to the original blocks and the desired output vector is representing the original pixels of the corresponding original input block. After training, the network reaches stability and a reconstructed image of same size, as the original will be obtained. Pixel by pixel difference is calculated between the original image and the reconstructed image, which results in a difference image $E_0(i, j)$, $1 \le i, j \le 2^n$, of same size as the original one 2^nx2^n.

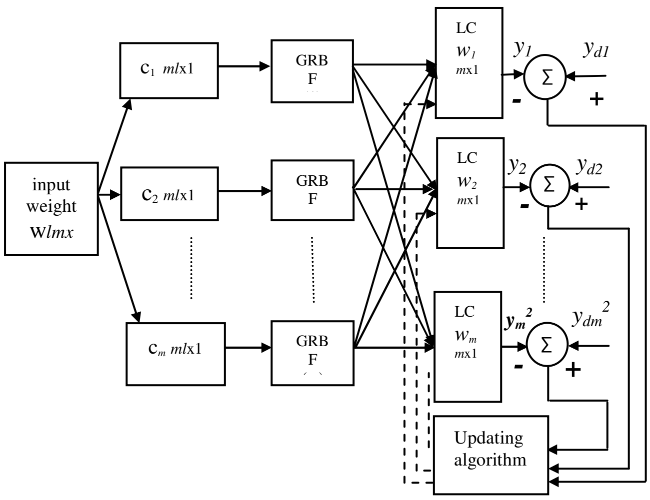


Fig. 3. GRBF NN for Block Reconstruction

Step 5: For the second pyramid level (P=1), starting with the difference image E_0, the whole difference image $E_0(i, j)$ of size 2^nx2^n is divided into a number of ($4l$) blocks, each of them of size ($m/2$x$m/2$) pixels. A GRBF network will be created with an input vector of length $(m/2)^2$x1, an intermediate stage consists of an array of centers c_i (i=1,2,...,4l), the RBF center vectors are set to the values of the pixels corresponding to each block. The output stage consists of $m/2$ weight vectors each weight vector $W_{i\cdot}=[w_{i1}\ w_{i2} .... w_{4il}]^T$. The output vector y of length $m/2$x1 will be compared with the desired output vector y_d ($m/2$x1), which is set to the average vector of each block.

Step 6: Applying the training input–output pairs, and following the same procedure of Step (3), considering the new vectors dimensions. The obtained weight matrices ($4l$x$m/2$) per block form the coefficients of the second pyramid level (p=1). The weights will be quantizated, encoded and transmitted.

Step 7: Following the same reconstruction procedure as Step (4), considering the new dimensions, a reconstructed image of the same size, as the original will be obtained. Pixel by pixel difference is calculated between the difference image E_0 and the reconstructed difference image E_0', results in a new difference image $E_1(i,j)$, $1 \le i, j \le 2^n$, of the same size as the original one 2^nx2^n.

Step 8: For the remaining pyramid levels, starting from Step (5) and following the same procedure to obtain the compressed coefficients, these coefficients are used at the transmitter side to reconstruct the difference image E_p' and to obtain new difference image E_p to be used for the next pyramid level. The stopping criterion here is when the minimum cost function is obtained in accordance with Eq.(6).

Decoding process

Step 1: For each level, the process of decoding the received weight matrices and de-quantization has to be done.

Step 2: The reconstructed difference image is obtained using the same arrangement of block reconstruction as the one, used at the transmitter side for each level.

Step 3: The elements $\hat{B}(i,j)$ of the restored image are calculated in accumulation way [6,8,9], which can be expressed mathematically as follows:

$$\hat{B}(i,j)=\sum_{p=0}^{P-1}\tilde{B}(i,j) \tag{8}$$

Where i, j = 1,2, . . , 2^n and $\tilde{B}(i,j)$ is the reconstructed image using the GRBFNN at each level of the receiver side.

V. Performance Evaluation of the IDP-GRBF

In order to evaluate the performance of the proposed BPNN-IDP algorithm, the commonly known measures will be used:

(i) The Peak Signal to Noise Ratio (PSNR) obtained for the reconstructed image at each level p of the pyramid as:

$$\mathrm{PSNR}(p)=10\log_{10}\frac{B_{max}^2}{\bar{\varepsilon}^2(p)}, \mathrm{dB} \tag{9}$$

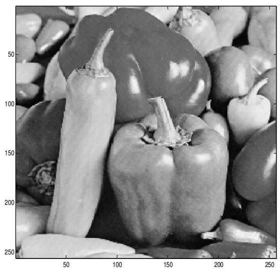
Where $\bar{\varepsilon}^2(p)$ is the mean square error (MSE) at level p:

$$\bar{\varepsilon}^2(p)=4^{-n}\sum_{i=0}^{2^n-1}\sum_{j=0}^{2^n-1}[B(i,j)-\hat{B}_p(i,j)]^2 \tag{10}$$

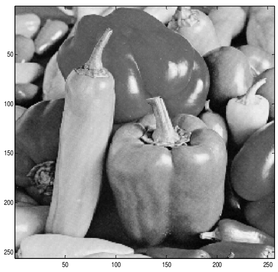
(ii) The image compression is calculated as the ratio of the total number of bits transmitted to the total number of pixels in the original image.

VI. SIMULATION RESULTS

A MATLAB Version 6.5, computer simulation program was designed to simulate the IDP-GRBF depending on the NNs toolbox. Experiments have been performed on grayscale and color images as well. For color images, the described algorithm can be performed applying it on the matrix of every primary color component: R, G, B. In order to obtain higher compression ratio, the R, G, B components of every pixel (i,j) were transformed in Y, Cr, Cb format. For a 512x512 pixel test images, the first level was of size 4x4 blocks (128x128 pixel), while for 256x256 pixels test images the first level was 2x2 blocks (128x128 pixels) and continued as described in section IV. The results for PSNR, CR, and number of levels for a number of test images were reported and compared. Fig. 5a shows two consequent pyramid levels for “Peppers”, Fig.5b represents the first level for “crosses” and “circles” test images. Table 1 shows a comparative result of the developed algorithm with the standard JPEG 2000 for the same compression ratio (results of JPEG are obtained with Lura Smart Compress software); the results obtained from the developed algorithm can be further compressed by applying encoding techniques.



CR=0.125 bpp, PSNR=32.4 dB



CR=1 bpp, PSNR=34.2 dB

CR=0.28 bpp, PSNR=28.9 dB

CR=0.28 bpp, PSNR=28.3 dB

Fig. 5. a) Two consecutive pyramid levels for reconstructed test image “Peppers”, 512x512 pixels; b) Test images “Circles” and “Crosses”, 256x256 pixels

VII. CONCLUSION

The novelty in this study was to develop a new pyramidal scheme based on IDP decomposition lies in modeling each pyramid level using GRBF NNs. This new algorithm can be compared with the most similar pyramid proposed in articles concerned with IDP. It can be underlined that, compared to IDP-DCT decomposition, IDP-GRBF reduces the number of levels, required to reconstruct the image at the receiver side and increases the PSNR, and CR. The coding and decoding of the hidden weights values are relatively simple.

TABLE 1. Comparison between JPEG 2000 standard and IDP-GRBFN for same compression ratio.

Test		CR	JPEG2000 PSNR (dB)	GRBF PSNR (dB)
Lena	7.1	1.13	38.4	33.2
Barbra	7.1	1.13	38.95	31.01
Peppers	7.1	1.13	40.04	34.2
Text	7.1	1.13	28.5	28.09
Circles	28.4	0.28	33.3	28.98
Crosses	28.4	0.28	30.54	28.33

The computation does not contain complexities compared with transformation methods and the process can be speeded-up by using the pre-defined parameters; also the quality of the reconstructed image obtained from the new method after a few number of levels used is considerable. The quality of the image obtained from the IDP-BPLA [10] and the IDP-GRBF is almost the same while the developed algorithm has the superiority in learning time, in addition to no probability to stick in local minima. The main drawback of the developed method is that to reconstruct the image at the receiver side, it requires the receiver to have a prior knowledge about the received image to start the supervised learning algorithm.

VIII. ACKNOWLEDGEMENT

We would like to acknowledge the financial support of the NSF of Bulgaria (Contract ВУ-МИ-104), and the Special Studies Academy at Cairo.

REFERENCES

[1] K.Tzou, "Progressive Image Transmission: A review and comparison of techniques" *Opt. Eng.,Vol 26, pp.581-589, July 1987.*

[2] A. Poularikas. "The Transforms and Applications Handbook". *CRC press 2000.*

[3] P. Burt. "The Laplacian Pyramid as a Compact Image Code". *IEEE Trans. Commun., Vol 31, No. 4, April 1983.*

[4] M. Goldberg. "Comparative Performance of Pyramid Data Structures for Progressive Image Transmission*", IEEE Trans. commun. Vol. 39, No.4, April 1991.*

[5] W. Kim, et al. "Hierarchy Embedded Differential Image for Progressive Transmission using Lossless Compression" *IEEE Trans. Circints and systems for Video Technology. Vol. 5, No.1, Feb. 1995.*

[6] N. Hikal, R. Kountchev, ”2D-SOFM Vector Quantization for Image Compression based on Inverse Difference Pyramidal Decomposition”, *Proceeding of TELSIKS, Sept. 2005.*

[7] W. Kou, ”Image compression Algorithms and Standards”. *Kluwer Academic Publishers, Boston, 1995.*

[8] R. Kountchev, et al. "Inverse Pyramidal Decomposition with Multiple DCT". *Signal Processing. Image communication 17 (2002), pp. 201-218.*

[9] N. A. Hikal, R. Kountchev. ”BPNN for image compression based on IDP Decomposition”. *Proceeding of ICEST, June 2005, pp. 15-18.*

[10] Y. H. Hu, J. N. Hwang. "Handbook of Neural Networks and Signal Processing”, *CRC Press 2002.*

Pixel-Based Searching of Images Stored in a Database

Igor Stojanovic[1] and Momcilo Bogdanov[2]

***Abstract* – In the paper we apply pixel-based searching of images stored in a database. We make use of progressive wavelet correlation along with Fourier methods. The searching consists of three incremental steps, each of which quadruples the number of correlation points. The process can be halted at any stage if the intermediate results indicate that the correlation will not result in a match.**

***Keywords* – JPEG, database, wavelets.**

I. INTRODUCTION

A key tool that helped make the Internet universally useful is the text-search engine. The image-search engines available today are relatively crude. There are several techniques for image searching: descriptor-based search, pixel-based search and image understanding techniques. The fastest methods available today use descriptor-based search techniques. IBM QBIC (www.qbic.almaden.ibm.com) [1] is an example of this type of search engine. Images with higher information content, such as satellite images and medical images, are difficult to encapsulate with descriptors. Queries on images of this type require detailed analysis. Normalized correlation coefficients, an instance of pixel-based search techniques, measure the differences between images and patterns. They can be computed with progressive wavelet correlation using Fourier methods [2]. The images are mapped into the wavelet-frequency domain to take advantage of high-speed correlation.

The paper is organized as follows. Section II contains the brief description of progressive wavelet correlation using Fourier methods [2]. An implementation of progressive wavelet correlation using Fourier methods for searching of images stored in a database is presented in Section III.

II. STONE'S METHOD OF PROGRESSIVE WAVELET CORRELATION

A. Overview of the Method

The basic idea of Stone's method includes: elimination in the earlier phases of searching, correlation in the frequency domain for fast computation, DCT in factorizing form, and wavelet representation of signals for efficient compression [2].

The algorithm consists of three incremental steps:

1. *Coarse correlation* – every eighth point of the correlation is generated.
2. *Medium correlation* – obtain the correlation at indices that are multiples of 4 mod 8 of the full correlation.
3. *Fine correlation* – obtain the correlation at indices that are multiples of 2 mod 8 and 6 mod 8 of the full correlation.
4. *Full correlation* – obtain the correlation at odd indices.

Fig.1 is a flow diagram showing the steps performed for an image search according to the method.

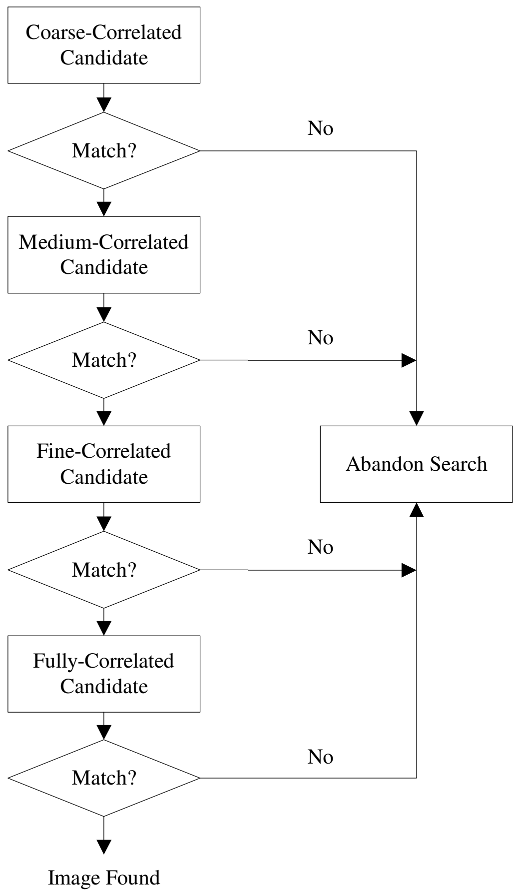


Fig. 1

There are two ways of using the method. The first one is for locating a given image in an image library, and the second one is for locating objects in the JPEG image [5]. Examples of locating a given image, locating an object in a JPEG image and conclusions are given in [3] [4].

---

[1]Igor Stojanovic is with the ICT Sector, Customs Administraton of RM, Lazar Licenoski 13, 1000 Skopje, Macedonia,
E-mail: igor.stojanovic@customs.gov.mk

[2]Momcilo Bogdanov is with the Faculty of Electrical Engineering, Karpos II b.b., P.O. Box 574, 1000 Skopje, Macedonia,
E-mail: bogdanov@ etf.ukim.edu.mk

B. Two-Dimensional Case

We investigate what happens in the two-dimensional case. The assumption is that the image size is N by N. In step 1, we have 64 subbands of length $N^2/64$. We perform one step of the inverse 2D **H** function, and one 2D step of the forward Fourier transform function. Fortunately, these steps are simple generalizations of the 1D functions. Specifically, if in 1D we compute **Hx** (JPEG transform of **x**), where **H** is N by N and **x** is N by 1, then in 2D we compute **H**2**X** conjtrans (**H**2), where **X** is the 2D image to be transformed, **H**2 is N by N, and each row of **H**2 is equal to **H**. This is equivalent to applying **H** to each of the columns of **X** and **H** to each of the rows of **X**. The next step is to add the 64 subbands point by point to create a 2D array of size $N/8$ by $N/8$. If we take its inverse Fourier transform, we will obtain the correlations at points that lie on a grid that is coarser than the original pixel grid by a factor of 8 in each dimension.

In step 2, we obtain 16 subbands of size $N^2/16$ by adding the 16 subbands point by point, and taking the Fourier inverse. We will obtain the correlation values on a grid that is coarser than the original grid by a factor of 4 in each dimension.

In step 3, we obtain 4 subbands of size $N^2/4$. In step 4, we obtain the full resolution.

Formulas for calculating normalized correlation coefficients that measure differences between images and patterns are given in [2]. Normalized correlation coefficients can be computed from the correlations described above. The normalization is very important because it allows for a threshold to be set. Such a threshold is independent of the encoding of the images.

The normalized correlation coefficient has a maximum absolute value of 1. Correlations that have absolute values above 0.9 are excellent, and almost always indicate a match found. Correlations of 0.7 are good matches. Correlations of 0.5 are usually fair or poor. Correlations of 0.3 or less are very poor. There is a tradeoff between the value of the threshold and the likelihood of finding a relevant match. Higher thresholds reduce the probability of finding something that is of interest, but they also reduce the probability of falsely matching something that is not of interest.

III. Adaptation of Stone's Method for Searching in a Database

A. Image Store and Matlab Database Toolbox

The method of Stone provides guidelines on how to locate an image in the image library. To make this method practical, we must first decide how to store the images. The initial choice is to store them in a disk file system. This can be seen as the quickest and simplest approach. Another, better alternative should be looked at and that is the move to store those images in a database. In the past five years, with changes in database technology and improvements in disk performance and storage, the rules have changed and it now makes business sense to use the database to store and manage all of an organizations' digital assets. The following are the strengths a database can offer over traditional file system storage: manageability, security, backup/recovery, extensibility, flexibility.

We use the Oracle Database for investigation purposes. There are two ways of image's storage into Oracle Database. The first one is the use of Large Objects – LOB, and the second one is the use of Oracle *inter*Media.

Unstructured data such as text, graphic images, still video clips, full motion video, and sound waveforms tends to be large in size. A typical employee record may be a few hundred bytes, while even small amounts of multimedia data can be thousands of times larger. Datatypes that are ideal for large amounts of unstructured binary data include the BLOB datatype (Binary Large Object) and the BFILE datatype (Binary File object).

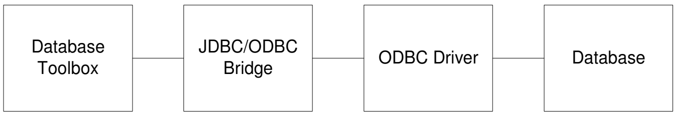
Oracle *inter*Media is a feature that enables Oracle Database to store, manage, and retrieve images, audio, video, or other heterogeneous media data. *inter*Media uses object types, similar to Java or C++ classes, to describe multimedia data. These object types are called ORDAudio, ORDDoc, ORDImage, and ORDVideo. An instance of these object types consists of attributes, including metadata and the media data, and methods.

To store images into database we use the BLOB datatype. After creation of one BLOB column defined table we also create a PL/SQL package with loading of images procedure (load named) included. This procedure is used to store images into the database.

The implementation of the method of Stone into Matlab and connection of the algorithm with the database are next steps. The Database Toolbox is one of an extensive collection of toolboxes for use with Matlab. The Database Toolbox enables one to move data (both importing and exporting) between Matlab and popular relational databases. With the Database Toolbox, one can bring data from an existing database into Matlab, use any of the Matlab computational and analytic tools, and store the results back in the database or in another database. The Database Toolbox connects Matlab to a database using Matlab functions. Data can be retreived from the database and store it in the Matlab workspace. At that point, the extensive set of Matlab tools can be used to work with the data. Database Toolbox functions can be included in Matlab M-files. To export the data from Matlab to a database, Database Toolbox functions can be used. The Visual Query Builder (VQB), which comes with the Database Toolbox, is an easy-to-use graphical user interface (GUI) for exchanging data with your database. The VQB can be used instead of or in addition to Database Toolbox functions.

Before the Database Toolbox is connected to a database, a **data source** must be set. A data source consists of data that you want the toolbox to access, and information about how to find the data, such as driver, directory, server, or network names. Instructions for setting up a data source depend on the type of database driver, ODBC or JDBC.

For Windows platforms, the Database Toolbox supports Open Database Connectivity (ODBC) drivers as well as Java Database Connectivity (JDBC) drivers. For UNIX platforms, the Database Toolbox supports Java Database Connectivity (JDBC) drivers. An ODBC driver is a standard Windows interface that enables communication between database management systems and SQL-based applications. A JDBC driver is a standard interface that enables communication between Java-based applications and database management systems. The Database Toolbox is a Java-based application. To connect the Database Toolbox to a database's ODBC driver, the toolbox uses a JDBC/ODBC bridge, which is supplied and automatically installed as part of the MATLAB JVM. The figure 2 illustrates the use of drivers with the Database Toolbox.



Windows platforms

UNIX and Windows platforms

Fig. 2

If the Windows-based database supports both ODBC and JDBC drivers, the JDBC drivers might provide better performance when accessing the database because the ODBC/JDBC bridge is not used.

The connection definition can be established using either the Oracle ODBC driver or the Microsoft ODBC driver for Oracle. To use these drivers, it is necessary to have Matlab and the Oracle client installed on the same computer. During testing we realized that Microsoft ODBC drivers for Oracle cannot be used for tables with columns of data type LOB.

For testing purposes JDBC drivers were usually used. If you have an Oracle database system that includes the JDBC drivers in the classes12.zip file, the following steps should be performed:

1. Create a directory $matlabroot/dbtools.
2. Put the classes12.zip file into dbtools.
3. Add the classes12.zip file to the classpath.txt file, by including this line in classpath.txt
4. $matlabroot/dbtools/classes12.zip
5. Restart MATLAB.

After setting up the data source for connecting to and importing data from a database we have used several standard functions of the Matlab Database Toolbox.

We can retrieve BINARY or OTHER Java SQL data types. However, the data might require additional processing once retrieved. For example, data can be retreived from a MAT-file or from an image file. Matlab cannot process these data types directly. One needs knowledge of the content and might need to massage the data in order to work with it in Matlab, such as stripping off leading entries added by the driver during data retrieval.

For purposes of saving of the extracted data into file "testfile", we created the Matlab file parsebin.m. Using the “imread” function, we stored the file date into a two-dimensional output variable x.

In working with the Microsoft Access database, ODBC drivers are used for extracting of the database data. The extracting process must contain adjustment in cutting of the header created by the driver. Let "m" be quantity of bytes attached to the beginning of the data package by the ODBC driver. The quantity of bytes depends on the file type (file extension).

To discovering the value of "m" we created a procedure in Matlab. That procedure helps us find the adequate value for "m" when linking with the image format.

When working with Oracle databases and extracting data with JDBC and ODBC drivers there is no need for adjustment, so "m" has always the value "1" (m="1").

B. HTTP Application

The last step in adaptation is to create Matlab applications that use the capabilities of the World Wide Web to send data to Matlab for computation and to display the results in a Web browser. The Matlab Web Server depends upon TCP/IP networking for transmission of data between the client system and Matlab.

In the simplest configuration, a Web browser runs on your client workstation, while Matlab, the Matlab Web Server (matlabserver), and the Web server daemon (httpd) run on another machine as shown in figure 3.

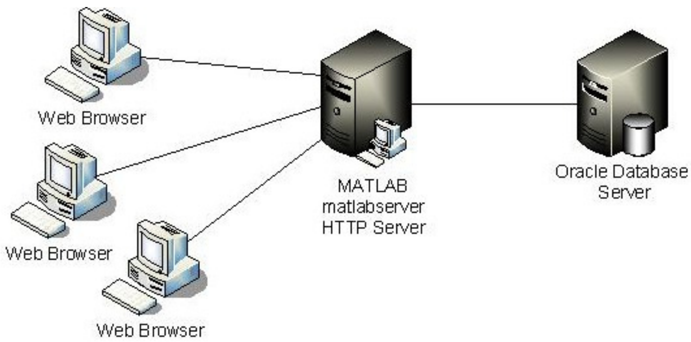


Fig. 3

In a more complex network, the Web server daemon can run on a machine apart from the others (Fig. 4).

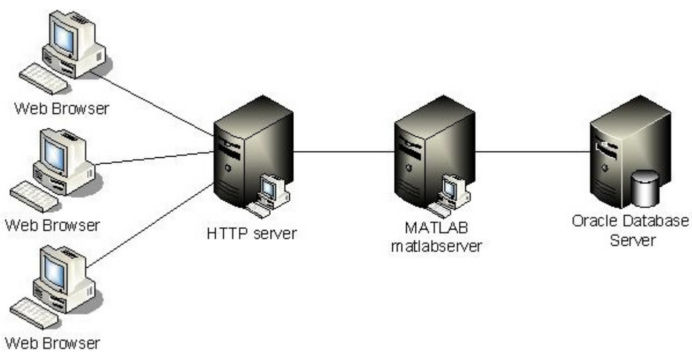


Fig. 4

Matlab Web Server applications are a combination of M-files, Hypertext Markup Language (HTML), and graphics.

The application development process requires a small number of simple steps:

1. Create the HTML documents for collection of the input data from users and display of output. You can code the input documents using a text editor to input HTML directly, or you can use one of the commercially available HTML authoring systems, such as Front Page from Microsoft etc.
2. List the application name and associated configuration data in the configuration file matweb.conf.
3. Write a Matlab M-file that:
   3.1. Receives the data entered in the HTML input form.
   3.2. Analyzes the data and generates any requested graphics.
   3.3. Places the output data into a Matlab structure.

The input mask of our application consists of three parameters: the image size *N*, the threshold lim, and the name of the image that we are looking for (Fig. 5).

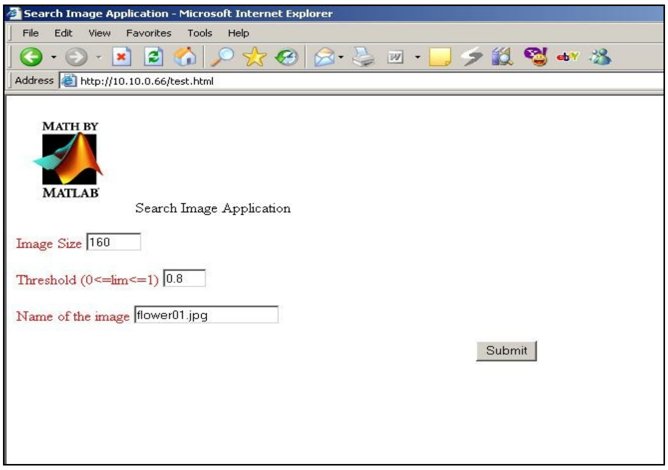


Fig. 5

IV. Conclusion

In this paper we presented our work that examines the pixel-based searching of images stored in a database. We exploited the technique of progressive wavelet correlation using Fourier methods, which led us to the following conclusions.

This technique is not yet suitable for general practical commercial usage. The reason for that is the big number of operations per picture. In the following years, with increasing processor speed, there should be a possibility for detail analysis at the rate of 1000 pictures per second. With this processing speed, it would be easy to construct a system, which is a combination of searching by description (descriptor-based search) and searching by pixels (pixel-based search). The descriptor can be used for isolating a certain part out of a big collection, which should be an object of a detailed pixel-based search in the later phase.

References

[1] M. Flickner et al., “Query by image and video content: The QBIC system,” IEEE Comp., vol. 28, pp. 23-32, Sept.1995.

[2] H. S. Stone, “Progressive Wavelet Correlation Using Fourier Methods,” IEEE Trans. Signal Processing, vol. 47, pp. 97-107, Jan. 1999.

[3] I. Stojanovic, M. Bogdanov, “Location of Objects in a JPEG Image with Progressive Wavelet Correlation using Fourier methods”, 9th Telecommunications Forum, pp. 561-564, Belgrade, Yugoslavia, Nov. 20-22, 2001.

[4] I. Stojanovic, D. Taskovski, I. Kraljevski,”Normalized Correlation Coefficients for Searching JPEG Images”, 7th Information Technologies 2002, pp. 104-107, Zabljak, Yugoslavia, 24 Feb. - 2 Mar., 2002

[5] G. K. Wallace, “The JPEG still-picture compression standard”, Commun. ACM, vol. 34, no.4, pp. 30-44, Apr. 1991.

Sets of Representative Points and Modified Hausdorff Distance for Image Registration

Yulka P. Petkova[1], Mariana Ts. Stoeva[2]

Abstract – **The present work is aroused from the striving after decreasing the computational complexity of the similarity measures, which are used in the tasks for searching and localizing the before known template into the larger image, keeping the accuracy of the measure. We propose criteria for forming the sets of representative points and a modification of the classical Hausdorff distance, which uses the specific features of these sets. We theoretically prove, that the modified distance has smaller computational complexity than classical.**

Keywords – **Hausdorff distance, Template matching, Representative points, Image registration.**

I. Introduction

Systems for pattern recognition and image processing have great applications in many fields as for example: character recognition, scene analysis, analysis of medical signals and images, person identification, face recognition and also in the robot technique in different areas of industry. One of the most frequently solved tasks in these systems is the task of searching, registration and localizing a properly chosen template into a larger image. This task is being solved using different methods for coinciding properly chosen sets of representative points and using different distances to evaluate matching between them. The main goal of the improvements of these methods is to increase or to keep their reliability and to decrease their computational complexity as well.

In the literature there are many classifications of registration methods. One of them is for example that, proposed in [1]. According to this classification, methods for image registration can be divided into the following classes:

- methods, which directly use the intensity value in each pixel, i.e. correlation methods;
- methods, which use the transformation in the frequency domain, i.e. methods, based on the Fast Fourier Transformation;
- methods, which use the representation of the template with so called low level features, for example – edges, corners, and contours, these are so called "feature based" methods;
- methods, which use high level characteristics, like identified object (or parts of objects) or relevancies between their features; these are so called graph-theoretic methods.

[1] Yulka P. Petkova is with the Department of Computer Science and Technologies, Technical University of Varna, 1, Studentska Str., Varna 9010, Bulgaria, Email: jppet@mbox.digsys.bg

[2] Mariana Ts. Stoeva is with the Department of Computer Science and Technologies, Technical University of Varna, 1, Studentska Str., Varna 9010, Bulgaria

Each of these methods can be realized in different way, using different similarity (or matching) measures. The most successful matching measures are based on the distance transformations, because of their high stability to missing or partly occluded data [2], [3], [4]. In these methods binary templates represent objects. Most frequently these template contain information about the coordinates of the representative points. Hausdorff distance is one of the measures, which determine the distance between two sets of points. It measures the extent to which each point of a template set lies near some point of an image set and vice versa.

Most frequently sets of representative points are formed from the edge points, from the corner points or from the contour points, because these points bring the biggest information about the image.

Our goal in this paper is directed to the forming of the sets of representative points and also to a modification of the Hausdorff distance in order to decrease its computational complexity, using the features of the point sets and keeping the worth of the measure.

II. Forming the Set of Representative Points

A. Edge Points

In [7] we propose a new edge definition, based on the interrupted first derivative of intensity function. In practice at the points of interruption second derivative has local extreme values. The rule for determining if a point belongs to an edge (for one-dimensional case) is:

$$\text{If} \quad \left| \underset{i}{extr}\left(\frac{d^2 I(x)}{dx^2} \right) \right| > \theta, \quad \text{then} \quad i \in E, \tag{1}$$

where $I(x)$ is the intensity function; E is a set of points i, which are edge points; θ is a before settled threshold.

Here we propose to divide such formed set E into two subsets, depending on the sign of the local extremum. We propose the following **definition:**

Definition 1: The set E consists of two subsets, which are formed according the following criteria:

$$\text{If} \quad \left(\underset{i}{extr}\left(\frac{d^2 I(x)}{dx^2} \right) \right) > 0, \quad \text{then} \quad i \in E_L \tag{2.1}$$

and

$$\text{If} \quad \left(\underset{i}{extr}\left(\frac{d^2 I(x)}{dx^2} \right) \right) < 0, \quad \text{then} \quad i \in E_H \tag{2.2}$$

The set of all edge points E is a sum of both subsets

$$E = E_L \cup E_H \,. \tag{3}$$

B. Informative points, selected with the method of equipotential planes

In [6] we propose extracting the representative points to be conformed to the criterion of D-optimality. According to this criterion the most informative points lie on the protruded peripheral wrapping of the object [5]. Thus, we propose the choice to be made by equipotential planes, which are parallel to the plane *xOy* and which cut the three-dimensional image profile (relief) on proper intensity levels. Thus, the extracted points outline the horizontal contours of the local "hollows" and "hills" from the three-dimensional image profile.

Here we propose to divide such formed sets into two subsets, depending on their belonging to local "hollow" or "hill", according to the following **definition:**

Definition 2: Set of points EP consists of two subsets, which are formed according to the following criteria:

$$\text{If}\begin{pmatrix} [(PI(i,j)\le Pt_{max})\cap(PI(i+1,j)>Pt_{max})]\cup \\ \cup[(PI(i,j)\ge Pt_{max})\cap(PI(i+1,j)<Pt_{max})] \end{pmatrix}, \text{then}\,(i,j)\in EP_H \tag{4.1}$$

$$\text{If}\begin{pmatrix} [(PI(i,j)\le Pt_{min})\cap(PI(i+1,j)>Pt_{min})]\cup \\ \cup[(PI(i,j)\ge Pt_{min})\cap(PI(i+1,j)<Pt_{min})] \end{pmatrix}, \text{then}\,(i,j)\in EP_L \tag{4.2}$$

where: $PI(i,j)$ is the intensity at the point (i,j); Pt_{min}, Pt_{max} are intensity values, which determine the distance between the *xOy* plane and the planes, cutting low and high parts of the relief ("hollows" and "hills"); EP_L, EP_H are sets of points, outlining "hollows" and "hills".

The set of all informative points EP is an union of both subsets:

$$EP = EP_L \cup EP_H \tag{5}$$

Both subsets can be additionally enriched in addition with the points, which describe "ridges" and "tablelands" in the separated "hills" and "valleys" and "lowlands" in the separated "hollows", as we propose in [8].

As it is seen both sets of informative points (E and EP) are composed of two subsets, which are formed according to the criteria (2.1) and (2.2) or (4.1) and (4.2). The idea, which arises here, is to use this property of the sets when we use the Hausdorff distance for determining the distance between them.

III. Modified Hausdorff Distance

The classic Hausdorff distance is defined [4] as follows:

$$H(A,B) = max(h(A,B),\, h(B,A)) \tag{6}$$

where $A = \{a_1, a_2, \dots a_p\}$ and $B = \{b_1, b_2, \dots b_q\}$ are finite, not empty sets of points, i.e. $p, q \neq 0$, and $h(A,B)$ and $h(B,A)$ are the directed distances.

Let the compared sets of points consist of two subsets, which are obtained using ones and the same criteria, particularly these, proposed in the previous section:

$$A = A_1 + A_2\,;\; A_1 = \{a_{11}, a_{12}, \dots, a_{1p_1}\};\; A_2 = \{a_{21}, a_{22}, \dots, a_{2p_2}\}$$
$$\text{where } p_1 + p_2 = p \text{ and } p_1, p_2 \neq 0 \tag{7.1}$$

$$B = B_1 + B_2\,;\; B_1 = \{b_{11}, b_{12}, \dots, b_{1q_1}\};\; B_2 = \{b_{21}, b_{22}, \dots, b_{2q_2}\}$$
$$\text{where } q_1 + q_2 = q \text{ and } q_1, q_2 \neq 0 \tag{7.2}$$

Then the equation (6) can be written as:

$$\begin{aligned} H(A,B) &= H(A_1 + A_2, B_1 + B_2) = max(h(A,B),\, h(B,A)) = \\ &= max(h(A_1 + A_2), h(B_1 + B_2)) \end{aligned} \tag{8}$$

We formulate the following **lemma:**

Lemma 1: Each of the directed distances between two sets is determined as the greater of the distances between the subsets "of the same names", i.e.

$$h(A,B) = max(h(A_1,B_1),\, h(A_2,B_2)) \tag{9.1}$$

$$h(B,A) = max(h(B_1,A_1),\, h(B_2,A_2)) \tag{9.2}$$

Proof: According to the definition [4] $h(A,B) = \max_{a\in A}\min_{b\in B}\|a-b\|$, but $A = A_1 + A_2$ and $B = B_1 + B_2$, which allows to write:

$$h(A,B) = \max_{a\in A} min\left(\min_{b_1\in B_1}\|a-b_1\|\,,\ \min_{b_2\in B_2}\|a-b_2\|\right) =$$

$$= max\begin{pmatrix} \max\limits_{a_1\in A_1}\left(min\left(\min\limits_{b_1\in B_1}\|a_1-b_1\|\,,\ \min\limits_{b_2\in B_2}\|a_2-b_2\|\right)\right), \\ \max\limits_{a_2\in A_2}\left(min\left(\min\limits_{b_1\in B_1}\|a_2-b_1\|\,,\ \min\limits_{b_2\in B_2}\|a_2-b_2\|\right)\right) \end{pmatrix},$$

but $\min\limits_{b_1\in B_1}\|a_1-b_1\| < \min\limits_{b_2\in B_2}\|a_1-b_2\|$ and $\min\limits_{b_2\in B_2}\|a_2-b_2\| < \min\limits_{b_1\in B_1}\|a_2-b_1\|$, according to the way of these subsets forming - in our case, according to the criteria (2.1) and (2.2) or (4.1) and (4.2).

Hence,

$$min\left(\min_{b_1\in B_1}\|a_1-b_1\|\,,\ \min_{b_2\in B_2}\|a_1-b_2\|\right) = \min_{b_1\in B_1}\|a_1-b_1\|\,.$$

Analogously:

$$min\left(\min_{b_1 \in B_1} \|a_2 - b_1\| \ , \ \min_{b_2 \in B_2} \|a_2 - b_2\| \right) = \min_{b_2 \in B_2} \|a_2 - b_2\| \ .$$

Thus, we can write:

$$h(A,B) = max\left(\max_{a_1 \in A_1} \min_{b_1 \in B_1} \|a_1 - b_1\| \ , \ \max_{a_2 \in A_2} \min_{b_2 \in B_2} \|a_2 - b_2\| \right),$$

but $\max_{a_1 \in A_1} \min_{b_1 \in B_1} \|a_1 - b_1\|$ is nothing else than $h(A_1, B_1)$ and $\max_{a_2 \in A_2} \min_{b_2 \in B_2} \|a_2 - b_2\|$ is nothing else than $h(A_2, B_2)$. Therefore: $h(A,B) = max(h(A_1,B_1), h(A_2,B_2))$, which we want to prove.

Analogously we can prove that

$$h(B,A) = max(\ h(B_1,A_1),\ h(B_2,A_2)).$$

The just proved claims give us a right to write the equation (8) as follows:

$$\begin{aligned} H(A,B) &= H(A_1 + A_2, B_1 + B_2) = \\ &= max\begin{pmatrix} max(\ h(A_1,B_1),\ h(A_2,B_2)), \\ max(h(B_1,A_1),\ h(B_2,A_2)) \end{pmatrix} \end{aligned} \qquad (10)$$

The last equation (10) is the base on which we formulate the following **definition:**

Definition 3: Modified Hausdorff distance is defined as:

$$H_M(A,B) = max\begin{pmatrix} max(\ h(A_1,B_1),\ h(A_2,B_2)), \\ max(h(B_1,A_1),\ h(B_2,A_2)) \end{pmatrix} \qquad (11)$$

where: $A = A_1 + A_2$; $A_1 = \{a_{11}, a_{12}, ..., a_{1p_1}\}$; $A_2 = \{a_{21}, a_{22}, ..., a_{2p_2}\}$; $p_1 + p_2 = p$ and $p_1, p_2 \neq 0$; $B = B_1 + B_2$; $B_1 = \{b_{11}, b_{12}, ..., b_{1q_1}\}$; $B_2 = \{b_{21}, b_{22}, ..., b_{2q_2}\}$; $q_1 + q_2 = q$; and $q_1, q_2 \neq 0$.

A_1, B_1 and A_2, B_2 satisfy ones and the same criteria.

Properties, which are valid for the classical Hausdorff distance, are valid for the modified distance (11) as well.

1) It is *not negative,* i.e.

$$H_M(A,B) = H_M(A_1 + A_2,\ B_1 + B_2) \geq 0;$$

2) It is *identical*, which means that the distance between two identical sets is zero, i.e.

$$H_M(A,A) = H_M(\ A_1 + A_2, A_1 + A_2\ ) = 0;$$

3) It is *symmetrical*, which means that:

$$\begin{aligned} &H_M(A,B) = H_M(\ B,A\ ) = H_M(A_1 + A_2,\ B_1 + B_2\ ) = \\ &= H_M(B_1 + B_2,\ A_1 + A_2\ ) \end{aligned}$$

Proofs of these properties are trivial and we don't consider them. We pay more attention to the following property, which is used when three sets of points are compared and the distances between them are evaluated. This property is:

4) The *triangle inequality*:

$$H_M(\ A,B\ ) + H_M(\ B,C\ ) \geq H_M(\ A,C\ )$$

or

$$\begin{aligned} &H_M(\ A_1 + A_2, B_1 + B_2\ ) + H_M(\ B_1 + B_2, C_1 + C_2\ ) \geq \\ &\geq H_M(\ A_1 + A_2, C_1 + C_2\ ) \end{aligned} \qquad (12)$$

The sense of this inequality is the following: if the distance between the compared sets *A* and *B* is small and if the distance between the sets *B* and *C* is small as well, than it can be claimed that the distance between *A* and *C* is also small, i.e. images, represented by *A* and *C* are similar.

In order to estimate if the triangle inequality (12) is satisfied when the modified Hausdorff distance is used, let answer the question: if the both of the compared subsets are the same, will the similarity measure work properly. With the other words, is there any possibility to make a wrong conclusion, that both sets are similar, i.e. the searched template to be recognized into the image.

In this connection we formulate the following **lemma:**

Lemma 2: If $A = A_1 + A_2$, $B = B_1 + B_2$ and $A_1 \equiv B_1$, but $A_2 \neq B_2$, then

$$H_M(A,B) = max(\ h(A_2,B_2),\ h(B_2,A_2))$$

or if $A_1 \neq B_1$, but $A_2 \equiv B_2$, then

$$H_M(A,B) = max(\ h(A_1,B_1),\ h(B_1,A_1)),$$

where A_1, A_2, B_1, B_2 are not empty sets and they are formed by ones and the same criteria.

Proof: According to Eq. (9.1)

$$h(A,B) = h(A_1 + A_2,\ B_1 + B_2\ ) = max(h(A_1,B_1), h(A_2,B_2)).$$

If $A_1 \equiv B_1$, then $h(A_1,B_1) = 0$, because of identity. If $A_2 \neq B_2$, then $h(A_2,B_2) > 0$, because of no negativity. Hence, $h(A,B) = h(A_2,B_2)$. According to identity and symmetry $h(B_1,A_1) = 0$, and $h(B_2,A_2) > 0$. It follows that we can express Eq. (9.2) as:

$$h(B,A) = h(B_1 + B_2,\ A_1 + A_2\ ) = h(B_2,A_2).$$

Finally, in this case the Hausdorff distance is:

$$H_M(A,B) = H_M(A_1 + A_2, B_1 + B_2) = max\left(h(A_2,B_2),\ h(B_2,A_2)\right).$$

Thus, the first part of the Lemma 2 is proved. The second part can be proved analogously.

The claim from the Lemma 2 can be expressed in the following way as well: the distance between two sets consisting of two subsets, which are obtained by ones and the same criteria, is estimated with the distance between the different subsets "of the same names", if the other subsets "of the same names" coincide.

From this formulation it obviously follows that the triangle inequality is satisfied, since the decision about the closeness between two sets makes only if their different subsets are close enough.

The formulated in this section definitions and lemmas are in the base of our decision for using the Hausdorff distance in a new way, in order to improve its computational complexity.

IV. Computational Complexity of the Modified Hausdorff Distance

According to [4] the computational complexity of the classical Hausdorff distance is $O_H(pq)$, where p and q are the numbers of elements of the compared sets.

Considering Eqs. (7.1) and (7.2) we estimate computational complexity of Hausdorff distance as follows:

$$\begin{aligned} O_H(pq) &= O_H((p_1+p_2).(q_1+q_2)) = \\ &= O_H(p_1q_1+p_1q_2+p_2q_1+p_2q_2) = \\ &= O_H((p_1q_1+p_2q_2)+(p_2q_1+p_1q_2)) = \\ &= O_H(p_1q_1+p_2q_2)+O_H(p_2q_1+p_1q_2) \end{aligned} \tag{13}$$

If the sets $A = A_1 + A_2$ and $B = B_1 + B_2$ satisfy criteria (2.1), (2.2) or (4.1), (4.2), then, according to Lemma 1 it is not necessary to compare the subsets with "different names" (A_1 with B_2, and A_2 with B_1). Thus, we can write:

$$O_{H_M}(p_1q_2+p_2q_1) = 0$$

and

$$O_{H_M}(pq) = O_{H_M}(p_1q_1+p_2q_2) \tag{14}$$

Comparing (13) and (14), it obtains:

$$O_H(pq) = O_{H_M}(p_1q_1+p_2q_2)+O_H(p_2q_1+p_1q_2),$$

from where it follows that:

$$O_{H_M}(pq) = O_H(pq)-O_H(p_2q_1+p_1q_2),$$

i.e.

$$O_{H_M}(pq) < O_H(pq) \tag{15}$$

Inequality (15) shows that the computational complexity of the modified Hausdorff distance is smaller than that of the classical distance, which is our goal.

V. Conclusion

In the present paper we theoretically propose a modification of Hausdorff distance, used for image comparison, considering the specific features of the compared sets of representtative points. We also propose two way for forming these sets, using the before proposed by us methods for representative points extraction [6], [7], [8]. We prove that the computational complexity of the modified Hausdorff distance is smaller than this of the classical distance.

Our next efforts will be directed to developing and investigating different algorithms for evaluation the proposed modified Hausdorff distance and its application in the tasks for searching and localizing the template into a larger image.

References

[1] B. S. Reddy, B. N. Chatterji, An FFT-Based Technique for Translation, Rotation and Scale-Invarian Image Registration, *IEEE Transactions on Image Processing,* Vol. 5, No8, pp. 1266-1271, August 1996

[2] D. M. Gavrila, Multi-feature Hierarhical Template Matching Using Distance Transforms, *in Proc. of IEEE International Conference on Pattern Recognition,* Brisbane, Australia, 1998.

[3] D. M. Gavrila, V. Philomin, Real-Time Object Detection for "Smart" Vehicles, *Proc. of IEEE International Conference on Computer Vision,* pp. 87-93, Kerkyra, 1999.

[4] D. P. Huttenlocher D. P., G. A. Klanderman, W. J. Rucklidge, Comparing Images Using the Hausdorff Distance, *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol.15, No9, Sep.1993

[5] D. S. Tyanev, Algorithms for pattern recognition and their application in technical diagnostics, PhD. Thesis, Sofia, 1991

[6] D. S. Tyanev, Y. Petkova, A New Method for Important Points Extraction, *International Scientific Conference CompSysTech 2004,* Rousse, Bulgaria, 17-18 June 2004, pp. III.11-1 – 11-8.

[7] D. S. Tyanev, Y. Petkova. About the Possibilities of a New Edge Definition in Black and White Images, *ISSE 2005, 28th International Spring Seminar on Electronics Technology,* Wiener Neustadt, Austria, May 19-22, 2005.

[8] Y. P. Petkova, D. S. Tyanev, *"An addition to the equipotential method for representative point extraction"*, Journal "Computer Science and Technologies", pp. 44-49, Varna, 2006.

Analysis of Complex Hadamard Transform Properties

Rumen P. Mironov[1], Roumen K. Kountchev[2]

***Abstract* – An analysis of Complex Hadamard Transform properties for 1D and 2D signals is presented. The "basis" images for 2D CHT and energy spectrums are obtained and evaluation of coefficients distribution in complex spectrum space for test images is made.**

***Keywords* – digital signal processing, orthogonal transforms, Hadamard transform, image analysis.**

I. INTRODUCTION

Discrete orthogonal transforms [1], [2] have been used extensively in the area of N dimensional signal processing, spectral analysis, pattern recognition and etc. The Walsh Hadamard Transform is a fairly simple transform and has found applications in data compression and watermarking involving image transmission, storage and security. The idea of using complex, rather than integer transforms matrices for spectral processing and analysis has been shown in [3], [4], [5] and [6]. From the Complex Hadamard Transform (CHT), several complex decisions diagrams are derived.

In this paper analysis of more general Complex Hadamard Transform properties for 1D and 2D signals are investigated, which are important for their applications in digital signal processing area. The 'basis' images for 2D CHT are obtained by using a developed MATLAB program simulation. The similar properties with well-known Hadamard Transform are received, what show that CHT can be used by the same way in more complicated analysis and processing [5], [6], [7].

II. MATHEMATICAL DESCRIPTION

One dimensional Complex Hadamard Transform

The forward and inverse one dimensional Complex Hadamard Transform (1D CHT) can be represented by the following equations [6]:

$$\left| \begin{array}{l} y(u) = \sum_{v=0}^{N-1} c(u,v).x(v) = y_{\mathrm{Re}}(u) + j y_{\mathrm{Im}}(u) \\ x(v) = \frac{1}{N} \sum_{u=0}^{N-1} c^{*}(u,v).y(u) \quad \text{for} \quad u,v = \overline{0, N-1} \end{array} \right. , \tag{1}$$

where: $N = 2^n$; $y(u)$, $x(v)$ are input and output N dimensional discrete signals; $j = \sqrt{-1}$; $y_{\mathrm{Re}}(u), y_{\mathrm{Im}}(u)$ are the real and the imaginary parts of $y(u)$; the coefficients of CHT are:

$$\left| \begin{array}{l} c(u,v) = j^{uv} s(u,v) \\ c^{*}(u,v) = (-j)^{uv} s(u,v) \end{array} \right. ; \tag{2}$$

$$s(u,v) = \begin{cases} 1 & \text{for } n = 2 \\ (-1)^{\sum_{r=3}^{n} \lfloor u/2^{r-1} \rfloor \lfloor v/2^{r-1} \rfloor} & \text{for } n = 3,4,5..... \end{cases} \tag{3}$$

is the sign function. Here $\lfloor . \rfloor$ is an operator, which represents the integer part of the result, obtained after the division.

From the equations (1),(2) and (3) follow that for $u=2p$

$$\begin{aligned} y(2p) &= \sum_{v=0}^{N-1} j^{2pv} s(2p,v) x(v) = \\ &= \sum_{v=0}^{N-1} (-1)^{pv} s(2p,v) x(v) = y_{\mathrm{Re}}(2p) \end{aligned} \tag{4}$$

and in the result $M(2p)=|y(2p)|$ and $\varphi(2p) = 0$, where $M(.)$ and $\varphi(.)$ are amplitude and phase frequency spectrum coefficients respectively.

When $u=4p+3$ for the odd coefficients of the transform the following are fulfilled:

$$\begin{aligned} y(4p+3) &= \sum_{v=0}^{N-1} j^{4pv} j^{3v} s(4p+3,v) x(v) = \\ &= \sum_{v=0}^{N-1} (-j)^{v} j^{4pv} s(4p+3,v) x(v) = y^{*}(4p+1) \end{aligned} \tag{5}$$

Equations (4) and (5) are valid for rows u as well for columns v and therefore the coefficients of CHT are symmetric about rows and columns.

In the results can be summarized that the even coefficients of CHT are real and the odd one are complex conjugated.

From the equations (2) the CHT basis matrix of order 2^n can be calculated for n=2:

$$[CH_4] = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & j & -1 & -j \\ 1 & -1 & 1 & -1 \\ 1 & -j & -1 & j \end{bmatrix} \quad [CH_4]^{*} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -j & -1 & j \\ 1 & -1 & 1 & -1 \\ 1 & j & -1 & -j \end{bmatrix} \tag{6}$$

The basis matrices of order 2^n (n>2) can be received as the Kroneker product of a number of identical "core" matrices of order 2^{n-1} in the following way:

$$[CH_{2^n}] = \begin{bmatrix} [CH_{2^{n-1}}] & [CH_{2^{n-1}}] \\ [CH_{2^{n-1}}] & -[CH_{2^{n-1}}] \end{bmatrix}. \tag{7}$$

[1]Rumen P. Mironov is with the Faculty of Communication Technics and Technology, Technical University of Sofia, Kl. Ohridsky 8, 1000 Sofia, Bulgaria, E-mail: rpm@tu-sofia.bg

[2]Roumen K. Kountchev is with the Faculty of Telecommunications, Technical University of Sofia, Boul. Kl. Ohridsky, 8, Sofia 1000, Bulgaria. E-mail: rkountch@tu-sofia.bg

As a sample, the basis Complex Hadamard Transform matrix of order 8, calculated by the equation (7) is :

$$[CH_8]=\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & j & -1 & -j & 1 & j & -1 & -j \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -j & -1 & j & 1 & -j & -1 & j \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & j & -1 & -j & -1 & -j & 1 & j \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -j & -1 & j & -1 & j & 1 & -j \end{bmatrix}.$$

In conclusion, using (1) forward and inverse CHT can be generalized in matrix form as:

$$\left| \begin{aligned} \vec{Y} &= [CH_N]\,\vec{X} \\ \vec{X} &= \frac{1}{N}[CH_N]\,\vec{Y} \end{aligned} \right. \quad \text{for:} \quad \left| \begin{aligned} \vec{Y} &= \{y(u)\,/\,u=\overline{0,N-1}\} \\ \vec{X} &= \{x(v)\,/\,v=\overline{0,N-1}\} \end{aligned} \right. . \tag{8}$$

From the above equations the following mathematical properties can be established:

$$\left|\det[CH_N]\right|^2 = N^N \tag{9}$$

$$[CH_N].[CH_N]^* = N[I] \tag{10}$$

$$[CH_N]^{-1} = \frac{1}{N}[CH_N]^* \tag{11}$$

$$[CH_N].[CH_N]^t = [CH_N]^t.[CH_N] = N[I] \tag{12}$$

Two dimensional Complex Hadamard Transform

The common results, obtained from the one dimensional Complex Hadamard Transform can be generalized for two-dimensional Complex Hadamard Transform. In this case the 2D signals (images) can be represented by the input matrix $[X]$ with the size NxN. The result is a spatial spectrum matrix $[Y]$ with the same size. The corresponding equations for the forward and the inverse 2D CHT are:

$$\left| \begin{aligned} [Y] &= [CH_N]\,[X]\,[CH_N] \\ [X] &= \frac{1}{N^2}[CH_N]\,[Y]\,[CH_N] \end{aligned} \right. \tag{13}$$

The symmetry of CHT coefficients allows 2D CHT to be accomplished in two steps. The first one is 1D CHT for every row the image and the second one is 1D CHT for the columns. This difference of transformation makes easier the calculations and the symmetry guarantees that the correlations between image elements in horizontal and vertical direction will influence in the same way the determination of transformed elements. The same considerations can be made for two steps calculation of the inverse 2D CHT.

III. Experimental Results

The 2D CHT can be expressed in different way by the equation:

$$[X] = \sum_{k=0}^{N-1}\sum_{l=0}^{N-1} y_{kl}[T_{kl}] \tag{14}$$

where $[T_{kl}]$ is the matrix of "basis" image with consecutive number (k,l). This expression can be presented as image decomposition $[X]$ in order on N^2 "basis" images with weighted coefficients y_{kl}. Using equations (6) and (14) these images of order 4 are simulated on MATLAB and are shown in Fig. 1a.

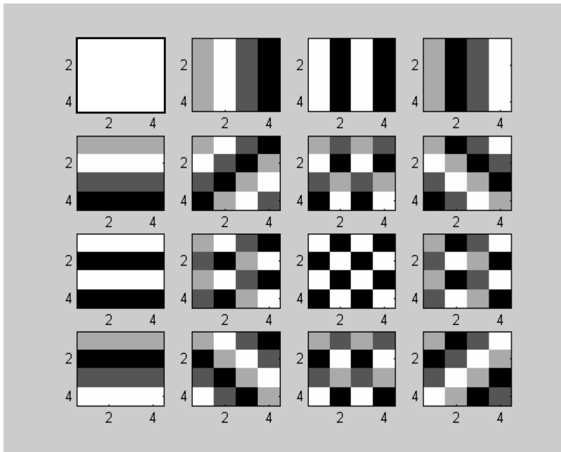


Fig. 1a. "Basis" images of order N=4 for 2D CHT.

In this figure the values (+1) are colored with white level, the values (-1) – with black level, the values (+j) – with white gray level and the values (–j) – with dark gray level. It is show that the received "basis" images look like the non-ordered "basis" images of real Hadamard Transform (HT).

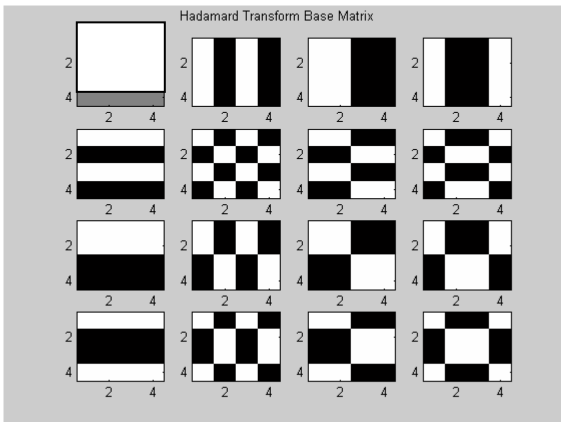


Fig.1b. "Basis" images of order N=4 for 2D HT.

On Fig. 1b are shown the same "basis" images simulated on MATLAB by real Hadamard Transform, arranged in naturally order.

On Fig. 2 are shown the "basis" images for 2D CHT of order 8, received in the same way.

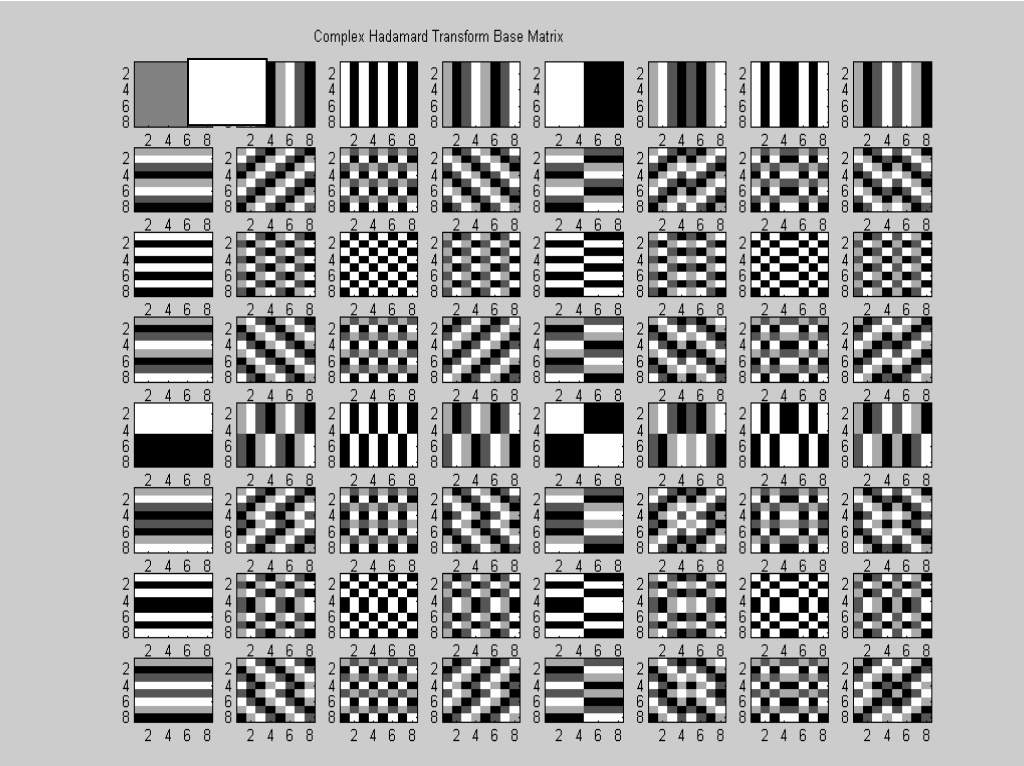


Fig.2. Basis images of order 8 for 2D CHT.

For the analyses of spectral distribution between the coefficients of 2D CHT a test image “LENNA”, shown in Fig.3, with size 512x512 and 256 gray levels is used. This image is transformed by the 2D CHT with kernel 16x16. By this way the input image is divided on 1024 sub-images with size 16x16 and is calculated by MATLAB 6.5 program. In Fig.4a and Fig. 4b the averaged amplitude frequency spectrums of all sub-blocks for two-dimensional Complex Hadamard Transform and two dimensional real Hadamard Transform respectively, are shown. On Fig. 4c the averaged phase frequency spectrum calculated for all sub-blocks, for two-dimensional Complex Hadamard Transform, is shown.

IV. Conclusion

A class of Complex Hadamard Transformation is presented. The general principles of complex matrices construction of high order for 1D and 2D transforms are given. The basic properties of CHT are discussed. The obtained amplitude spectrums for CHT and HT are practically identical and show that both can be used in similar applications.

The presented Complex Hadamard Transform can be used in digital signal processing for spectral analysis, pattern recognition, digital watermarking, coding and transmission of one-dimensional and two-dimensional signals.

Acknowledgement

The authors thank the National Fund for Scientific Research of the Bulgarian Ministry of Education and Science for the financial support by the contract I-1304/2003.

References

[1] N. Ahmed, K. R. Rao. Orthogonal Transforms for Digital Signal Processing, Springer-Verlag Berlin, Heidelberg, 1975.

[2] W. K. Pratt. Digital Image Processing, John Wiley&Sons, N.Y., 2001.

[3] B. Falkowski, S. Rahardja. "Complex Spectral Decision Diagrams", *Proc. of the 26th Int. Symposium on Multiple Valued Logic*, Vol. ISMVL’96, 1996.

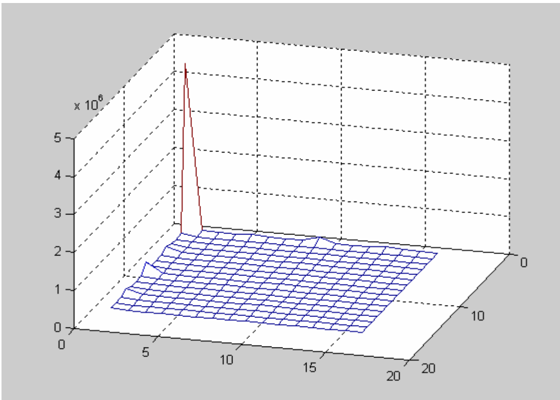
[4] S. Rahardja, B. Falkowski. “Complex Composite Spectra of Unified Complex Hadamard Transform for Logic Functions”, *IEEE Trans. on Circuits and Systems-II: Analog and Digital Signal Processing*, Vol. 47, No. 11, November 2000.

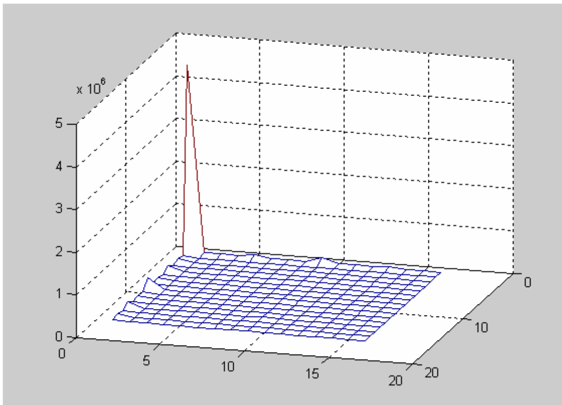
[5] B. Falkowski, S. Rahardja. “Complex Hadamard Transforms: Properties, Relations and Architecture”, *IEICE Trans. Fundamentals*. Vol. E87-A, No.8, August 2004.

[6] R. Kountchev, R. Mironov. “Audio Watermarking in the Phase-Frequency Domain”, XL Intern. Scientific Conference on Information, Communication and Energy Systems and Technologies, ICEST’2005, Nis, Serbia and Montenegro, 2005.

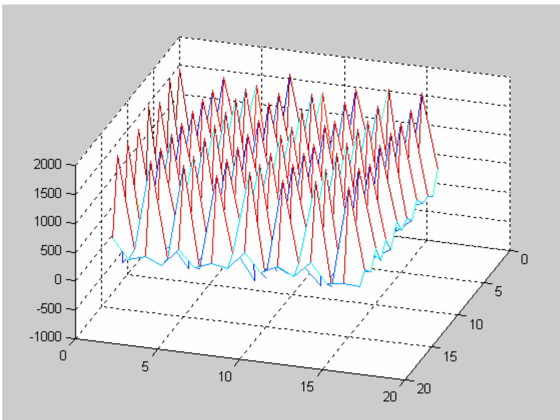
[7] A. D. Poularikas. The Transforms and Applications Handbook, Second Ed., CRC Press, 2000.

Fig. 3. Test image “LENNA” with size 512x512 pixels and 256 gray levels.





a) 2D CHT amplitude spectrum; b) 2D HT amplitude spectrum;



c) 2D CHT phase;

Fig. 4. Averaged amplitude and phase spectrums of image “LENNA”

Rotation Angle Estimation of Scanned Handwritten Cursive Text Documents

Ivo R. Draganov[1], Antoaneta A. Popova[2]

Abstract – **In this paper we propose an algorithm that estimates the rotation angle presenting in scanned images of handwritten cursive text. It combines a technique of tracking the separate words connectivity, narrowing the range of possible rotation angles to be checked and the extra obtained results are found useful in the following stages of a complete handwritten text recognition system.**

Keywords – **handwritten text recognition, angle estimation**

I. INTRODUCTION

It is a common task to check if the input image in most of the optical character recognition (OCR) systems for both printed and handwritten text has non-zero rotation angle. If this is so it is indication for either the sheet of paper has not been correctly aligned in the scanner's bed and or it is an effect due to the native characteristics of the writer handwriting or different styles for the text lines used. The rotation angle for the whole image (page) in all these cases should be correctly estimated and a new image with the right alignment used in the further processing steps.

There are numerous approaches trying to solve the problem mentioned above. We narrow it concerning only images containing handwritten cursive text. In general – the angle between every text line and the base of the image could be different, as well as the angle between any pair of the lines. Furthermore any line may contain certain linear parts that form different angles with the image baseline or speaking in other words – there may be curved lines. This makes the problem even harder to solve.

One of the very first attempts to solve this problem concerning printed text documents was done by Glauberman [1]. In the early days of the OCR systems he suggested the use of horizontal projections line-by-line counting the number of pixels forming the text in the image and thus forming the vertical histogram. Getting all the histograms for different angle projections (in a certain range with a given step) and finding the one corresponding to the minimal entropy directly gives us the rotation angle we are looking for. This basic principle lies down in great number of other approaches and usually leads to very good results.

[1]Ivo R. Draganov is with the Faculty of Communications and Communications Technologies, Technical University, Kliment Ohridski 8, 1000 Sofia, Bulgaria, E-mail: idraganov@abv.bg

[2]Antoaneta A. Popova is with the Faculty of Communications and Communications Technologies, Technical University, Kliment Ohridski 8, 1000 Sofia, Bulgaria, E-mail: antoaneta.p@komero.net

Pal et al. [2] propose a technique for multi-oriented text lines detection and their skew estimations. They use grouping boxes around single characters which help them to determine their belonging to certain words. Afterwards they find key points from the detected words which form reference lines. These lines then are used to extract whole lines of text recognizing the accurate angle to which each line is rotated. Artistic effects are not considered as a serious barrier in front of this approach.

In [3] Shivakumara et al. use static and dynamic thresholds in the process of separation every text line from the image including the usage of the projection profile. They also use linear regression analysis to estimate the skew angle for each text line after the separation. This approach is considered to be a step ahead including the Glauberman principle and optimizing the whole process of segmentation the lines from the image.

Clark and Mirmehdi in [4] also use projection profiles but here they are introduced to locate the horizontal vanishing point of the text plane. Thus they segment the lines of text and then reveal the style of justification of the paragraphs. Vertical vanishing point is obtained from analyzing the change in line spacing. This method is another serious step in estimating rotation (of the whole document) and skew (of the separate lines) angles and originally applicable for images of documents taken by a camera (not only a scanner) at different angles and without the knowledge of the focal length.

A final example for the expansion of the projection profiles in finding the rotation and skew angles and the lines separation as well, is given in [5]. Messelodi and Modena propose there the usage of a great number of heuristic filtering rules, as well as the coordinates of the centers and vertexes of the bounding rectangles around every candidate line, word and character. They implement three-level alignment segmentation for every line candidate which leads to good results even in images taken from complex real scenes containing some sort of text (printed, handwritten, with artistic effects etc.).

Our goal is to extend the basic algorithm using projection profiles [1] for finding the rotation angle of the whole page in such way that it can performs faster and in the same time can process images containing handwritten text which lines are not straight. The last property is absent in the most algorithms proposed so far for solving this problem.

In the next second section of this paper we describe such an algorithm and in the third one we give some experimental results. Afterwards we make our conclusions in the fourth part.

II. PROPOSED ALGORITHM

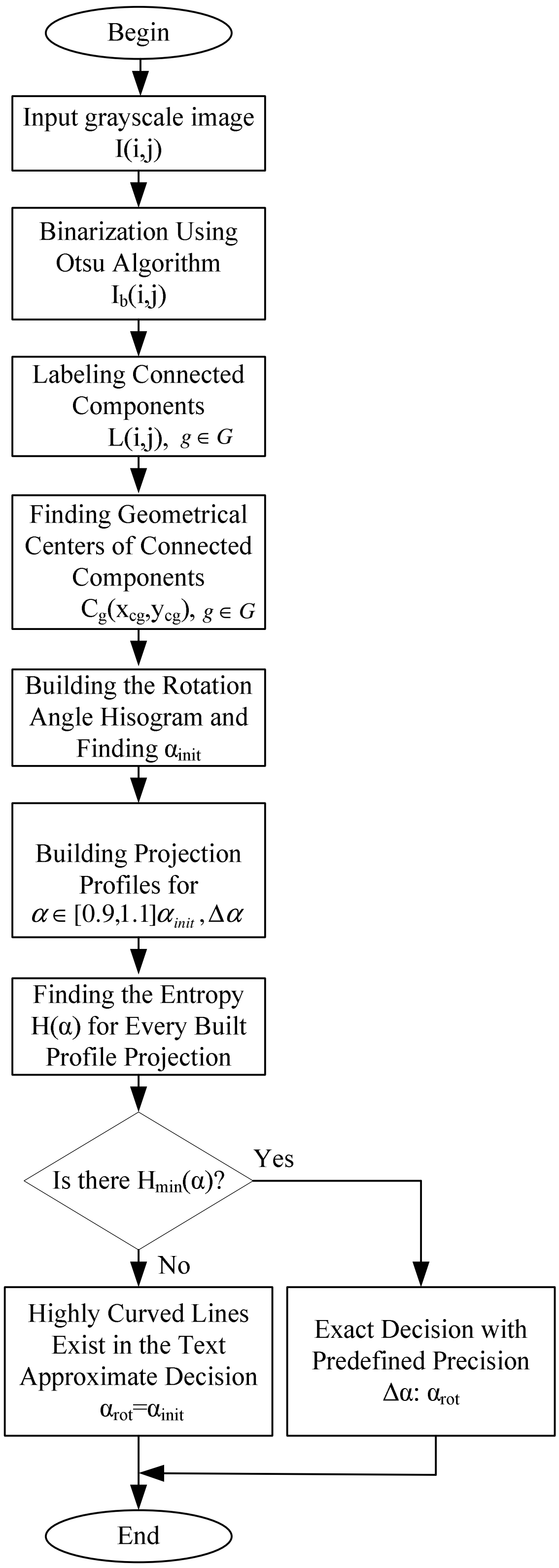


Fig. 1. Flowchart of the proposed algorithm

Narrowing the possible input only to handwritten text and knowing the advantages of the Messelodi-Modena method, described in [5], we propose a faster algorithm (Fig.1) due to the limited input mentioned. One heuristic rule is used concerning the values of one threshold, as well as we do not need to build a boundary rectangles forming line, word and character blocks and perform three-level segmentation. As for the word/character segmentation we suggest using of an indexing (labeling) algorithm which has proven its efficiency. Thus suggested algorithm can be described in the following steps:

1) Input a grayscale image, described by the intensity function $I(i,j) \in [0,255], i \in [0, M-1], j \in [0, N-1]$ containing the handwritten text scanned. We assume that the sheet of paper containing the original text has no background elements (e.g. stripes) and no strikes over the text as well.

2) Applying intensity segmentation with one threshold (binarization) using Otsu algorithm:

$$I_b(i,j) = \begin{cases} 1, I(i,j) \geq \theta_{opt} \\ 0, I(i,j) < \theta_{opt} \end{cases}, \qquad (1)$$

for $i \in [0, M-1], j \in [0, N-1]$. θ_{opt} - the optimal intensity threshold estimated by the Otsu algorithm.

3) Labeling (indexing) all elements (pixels) in the image *(i,j)* with a number that differs for elements belonging to different words or separate characters. We use the mask shown on Fig. 2a).

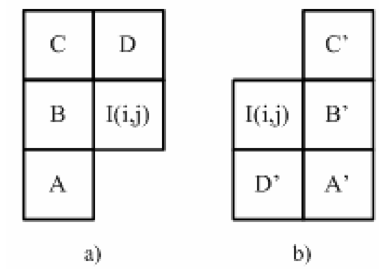


Fig. 2. Working masks used for labeling

Using the four neighbour elements *A, B, C* and *D* we label the current element *(i,j)* if its value is 1 using:

$$L(i,j) = \min\{L_1(A), L_2(B), L_3(C), L_4(D)\}, \qquad (2)$$

where *L(i,j)* is the index associated with *(i,j)*. Then we make a table of connectivity as shown in Table I. It shows the label of every element and with which other elements is connected. The first scan should be made vertically from top to bottom and from left to right.

TABLE I
TABLE OF CONNECTIVITY

g-th Label	Connected with m(g) marks
2	3
3	2
4	5, 6
…	…

G is the maximal number of objects in the image. *m(g)* is the label for the *g-th* object from all the *G* objects.

A second scan should be made using the mask from Fig. 2b) and it has to be done in bottom-top and right-left directions again vertically. This second labeling is actually a checking function and in case that after is done some objects from the left column *g-th* Label in Table I have greater value than the labels of their connected components from the right column *m(g)* it is necessary these labels to be exchanged. Thus it is possible some numbers already associated with some objects to become redundant.

4) At this step the geometrical centers are found for every labeled object, using:

$$x_{cg} = \frac{\sum_{p=0}^{P-1} x_{pg}}{P}, \; y_{cg} = \frac{\sum_{p=0}^{P-1} y_{pg}}{P}, \tag{3}$$

where g is the label of the respective object, P is the number of pixels that it consists of and c stands for center. x_{cg} and y_{cg} are the coordinates of the geometrical center $C_g(x_{cg}\,,\, y_{cg})$ for the *g-th* object.

5) At this step we use a distance threshold $l_{opt.}$ It represents the average Euclidean distance $d_E(C_g, C_{g+1})$ between the centers of connected components which are neigbours from a single text line and its value is estimated experimentally over a large number of test images. Having this threshold we estimate the angle α formed by the line connecting each two centers from step 4), the distance between which is smaller than l_{opt}, and the base of the image. The process is illustrated on Fig.3. Eq. (4) represents it more clearly:

$$\alpha = \begin{cases} arctg\left(\dfrac{y_{c(g+1)} - y_{cg}}{x_{c(g+1)} - x_{cg}}\right), d_E(C_g, C_{g+1}) < l_{opt} \\ \varnothing, otherwise \end{cases}. \tag{4}$$

All estimated values for α are saved for the next step.

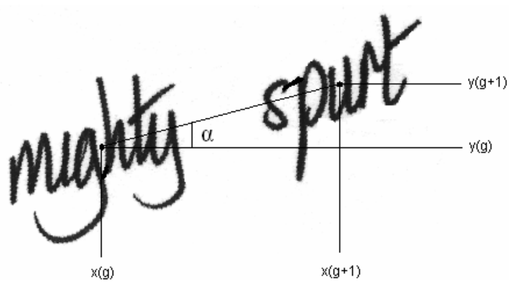


Fig. 3. Estimation of a local value for α defined by connected components centers

6) With all the values for α we can make the histogram of its distribution (Fig.4.b)). Then the global maximum is found and estimated the initial value for the angle to start testing hypothesizes for the accurate rotation angle of the whole image:

$$\alpha_{init} = f.\alpha_{\max}, \tag{5}$$

where *f* is a factor, typically *f=0.9*.

7) At this point we use $\Delta\alpha$ - the step that defines the error in rotation angle estimation α_{rot}. Typically $\Delta\alpha = 0.5°$ but it can be 0.1° as well in some cases where high accuracy is needed. Given this angle step we find the vertical projection profiles (Fig. 4c)) around the local maximum for α ($\alpha \in [0.9, 1.1]\alpha_{init}$) and the profile with a minimal entropy is the one

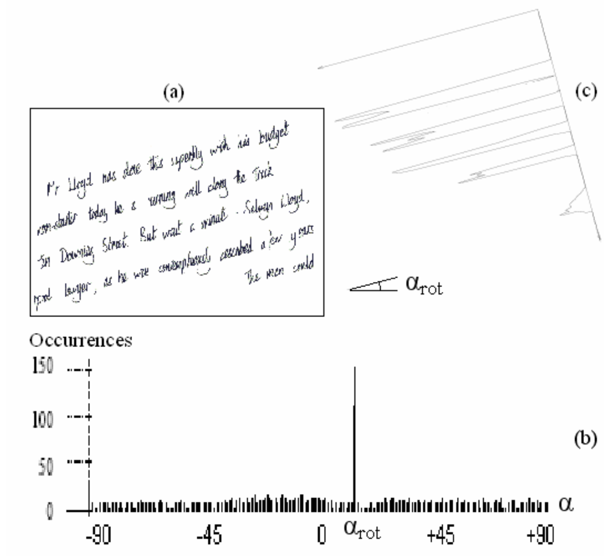


Fig. 4. Rotated text (a), estimation of the α-distribution (b), finding the histogram with minimal entropy corresponding to the angle of initial rotation (c)

corresponding to α_{rot} which is the result we are looking for so far. This is done using Eqs. (6) and (7):

$$h_n^{\alpha}(k) = \frac{h^{\alpha}(k)}{\sum_{k=0}^{M-1} h^{\alpha}(k)}, \tag{6}$$

which is the projection profile presented by a normalized histogram (the local value for a single line) for an angle α, k is the current line in the image of total M number of lines;

$$H(\alpha) = \sum_{k=0}^{M-1} - h_n^{\alpha}(k) \log_2 h_n^{\alpha}(k), \tag{7}$$

and $H(\alpha)$ is the entropy estimated from that projection profile. The minimum is reached when $\alpha = \alpha_{rot}$.

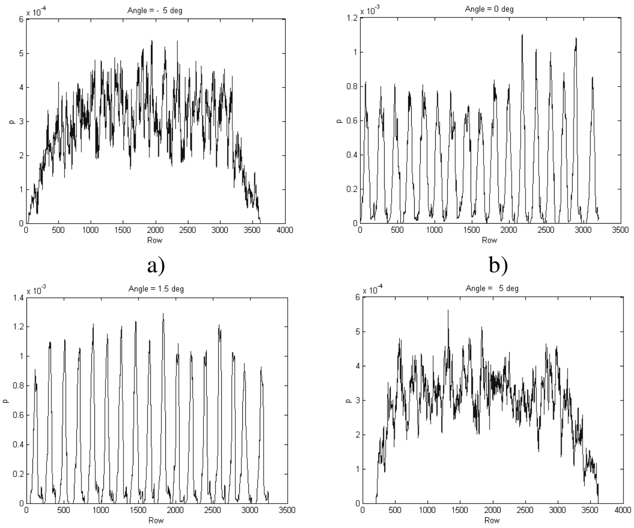
Finishing the last seventh step the goal is accomplished. It is worth noting that:

- With the estimation of α_{init} we strongly reduce the necessary number of projection profiles. Given $\Delta\alpha = 0.5°$ their number may be less than 5. Otherwise if we check all the angles from -90° to +90° with a step of 0.5° we need to estimate 361 projection profiles!
- If there are highly curved lines or such with alternatively changing slopes in different directions in the text it is obvious that no minimum for $H(\alpha)$ will be reached as for the case of the basic algorithm described in [1]. But α_{init} is a good approximate value for initial rotation angle correction of the whole page before any additional analysis being made for the exact structure of the document if needed.
- Segmenting the connected components and labeling them at this early stage is in great use in the later stages where determining the whole line positions, features extraction and the real recognition process take place.

Of course this labeling comes at a price of using some computing time but compared to the time period needed for estimating 361 projection profiles it is completely excusable.

III. Experimental results

For out experimentation we used a single page (A4 format white paper) containing handwritten with a black pen text of a single individual with 17 lines present. Then the sheet was scanned for 25 different positions of the rotating angle – from -6 to +6 degrees with a step of 0.5 degrees at 300 dpi resolution with a typical office flatbed scanner in a grayscale (256 gray levels used) mode. The initial estimating of α for non-rotated sheet (0°) showed +1.5° (due to the slope for the separate text lines originally introduced by the writer), and all the other 24 results showed correct estimation of *α* containing this initial line slope (the results were from -4.5° to +7.5° with a step of 0.5°), so 100% accuracy was reached, but it is obvious that it would be lower if some text with highly curved lines is processed.



a) b) c) d)

Fig. 5. Different profile projections: a) α=-5°, b) α=0°, c) $\alpha = \alpha_{rot}$= +1.5°, d) α=+5°

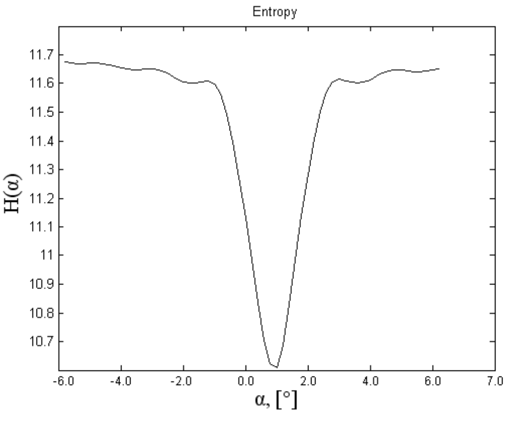
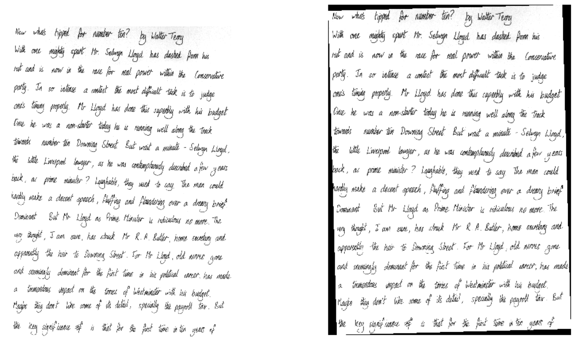


Fig. 6. The entropy as a function of α, reaching its minimum for α_{rot} = +1.5°

In Fig. 5 are presented 4 different profile projections. Fig. 5 c) corresponds to the minimum of the entropy.

The entropy *H*(α) itself is presented in Fig.6 as a function of the angle *α*.

And finally, Fig. 7a) represents the original input image of the analyzed handwritten text and Fig. 7b) represents an image with the same text but with corrected rotation angle. We used nearest neighbour approximation for maximal fast performance at the stage of rotating the image.



a) b)

Fig. 7. The original input image (a) and the corrected one after rotation the original to $\alpha = \alpha_{rot}$ = +1.5° (b)

IV. Conclusion

The proposed algorithm in seven steps produced very good results estimating extremely precisely the rotation angle of the whole text document image. It radically reduces the computational time for completing this task in comparison to direct projection profiles method. Moreover it extracts important additional information about the isolated words and characters in the text which will be used in the later stages of a complete text recognition algorithm. Thus the realization of this stage of preprocessing the input image can also be described as a module which can be easily embedded into large and complex systems aiming handwritten text recognition.

References

[1] M. Glauberman, Character Recognition for Business Machines, *Electronics*, No. 29, pp. 132–136, 1956.

[2] U. Pal, S. Sinha, B. Chaudhuri, "Multi-Oriented Text Lines Detection and Their Skew Estimation", *Third Indian Conference on Computer Vision, Graphics and Image Processing, Conference Proceedings*, Ahmedabad, India, 2002.

[3] P. Shivakumara, G. Kumar, D. Guru, P. Nagabhushan, "Skew Estimation of Binary Document Images Using Static and Dynamic Thresholds Useful for Document Image Mosaicing", *National Workshop on IT Services and Applications, Conference Proceedings,* New Delhi, India, 2003.

[4] P. Clark, M. Mirmehdi, "On the Recovery of Oriented Documents from Single Images", *Advanced Concepts for Intelligent Vision Systems, Conference proceedings,* pp. 190-197, Ghent, Belgium, 2002.

[5] S. Messelodi, C. Modena, "Automatic Identification and Skew Estimation of Text Lines in Real Scene Images", *Pattern Recognition,* Vol. 32, No. 5, pp. 789-808, 1999.

Framework for Video Editor with Support for Many Virtual Machines

Angel R. Kanchev[1], Antoaneta Popova[2]

Abstract – **Big software like video editor needs strong and stable framework. The attempt to make that framework at most open led to creation of plug-in system that can load both JavaVM [1] and .NET [2] components. Considering the multiplatform developing approach and the ability for distributed callings, the framework has become attractive for any big application.**

Keywords – **Java, .NET, plugins, plug-ins, multiplatform, framework.**

I. Introduction

The framework that is described here is made for Audio/ Video Editor but is universal enough – it can be used in any big application.

By "big" application, we will understand an application that is built by many modules and is big memory and CPU consumer. This is the reason to put on first place the requirement for fastness and size of the framework.

All requirements for such framework could be summarized this way:

1. To be fast, small and stable.
2. To be easily supported and easy for use.
3. To be maximum open and extensible.
4. To be easily ported for different platforms.

The described framework could be written for some Virtual Machine (VM) in order to be stable and multiplatform.

Currently there are two widespread virtual machines: Java VM [1] and .NET CLR (Common Language Runtime) [2].

Both virtual machines have disadvantages when they are used for big applications:

1. Interpretation of VM code (only for Just-In-Time and Install-Time compilations – managed native code is not a problem according this point).
2. Validation and verification of the VM code (correctness and security check)
3. No control over the memory usage (there is a "magical" tool – Garbage Collector).
4. Reflection – too universal type descriptions and calling conventions, which makes it too heavy.

The argument that writing native (unmanaged) code is dangerous is not strong enough. Actually there are two reasons for dangerous code – the programmer is not good (the team is not well formed) or the schedule is too tight. Neither the technology nor the language make the code more secure.

The good things in VM approach are the fast source compilation (to VM code), small size of the executables, platform independence, nice exception system and multi-language interoperation. So writing managed code is faster and easier – it is good for rapid application development (r.a.d.). For big, heavy applications, where full control over the hardware is needed, native code is still the best (however, if we put commercial arguments here the situation will be a little bit different).

The hybrid approach – like .NET with unsafe code, has the disadvantages of both systems (managed and unmanaged)… The best approach is to use one technology for the whole framework.

As a result, it is created a simple unmanaged framework that can load different virtual machines (see Table I).

TABLE I
REASONS FOR IMPLEMENTATION APPROACH

Decision	Reason
Simple	Small, stable, easy to support and use
Unmanaged	Fast and full control over the hardware
Loading of VM	It can support plug-ins for different Virtual Machines

The goal of the framework is to ease writing of open Video Editor. The key word here is open – the application must be dynamic (change of active modules at runtime) and extensib-le. Change of active modules means change of different modules that can do similar work.

Multiplatform approach is used in the framework's imple-mentation (multiplatform libraries are used and platform-de-pendant code is isolated).

There is plug-in system with cross-call capabilities (it will be described later).

II. Overall description

The framework is like mini virtual machine – it has:

- Application loaders
- GUI (Graphical User Interface) – the multiplatform library wxWidgets is used [3].
- Reflection-like system for class creation (there are identifications only for modules and class names)
- Plug-in system – the plug-in interfaces are statically hard-coded: no reflection, metadata, Interface Defini-tion Language (IDL) or any of these universal systems.

[1] Angel R. Kanchev, is with the Faculty of Communications and Communications Technologies, Technical University, Kliment Ohridski 8, 1000 Sofia, Bulgaria, E-mail: angel_kanchev@mail.bg

[2] Antoaneta A. Popova, is with the Faculty of Communications and Communications Technologies, Technical University, Kliment Ohridski 8, 1000 Sofia, Bulgaria, E-mail: antoaneta.p@komero.net

Note that any reflection-like system is heavy – even the current one should be used for major interfaces and for plug-ins only. On Fig.1, you can see the major interfaces (and the data transfer) for Video Editing system.

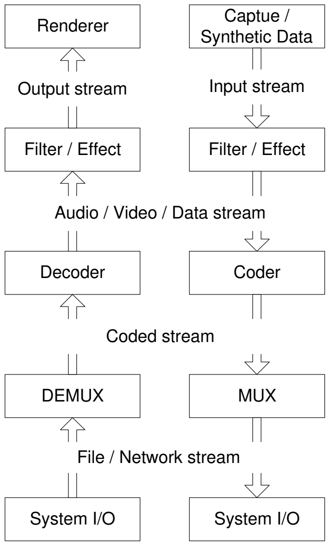


Fig. 1. Main parts of Audio/Video editor

The framework is abstract enough not to limit with Video Editor's needs, so we will talk for "Application" instead of Video Editor from now on.

From distribution point of view the application consists of three parts: loader, core and plug-ins. The application core has the major functionality of the application. The module that has the binding system between core modules, that dispatches calls to all core modules and that has the startup code for the application core is called Core. We will distinguish application core and module Core by the first capital letter.

On Fig. 2, you can see the major parts of the framework and their time relation (Core and PluginManager are modules in the application core). Arrow directions show function calls (i.e. at loading time the callings are unidirectional).

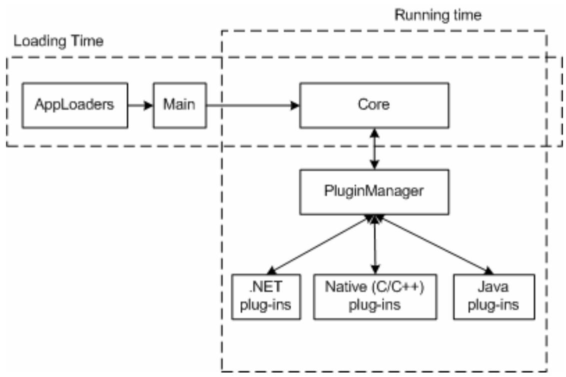


Fig. 2. Parts of the framework and their interaction in time

Note that plug-ins and PluginManager have bidirectional calls. That is – when the core needs a plug-in it uses PluginManager to initiate a call. If the plug-in needs something from the core, it can call back PluginManager, which in turn will call the necessary function in the core. This is the cross-call capability of PluginManager. For example, it allows calling .NET plug-in from inside Java plug-in (it will go through module Core – see Fig. 3). If it is necessary, PluginManager can load a VM – for now are supported JavaVM [1] and .NET CLR [4].

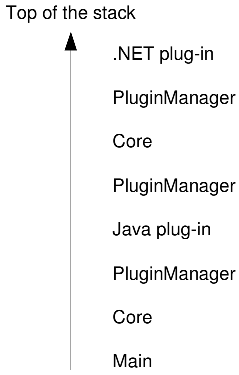


Fig. 3 Cross call stack (between Java and .NET)

On Fig.4, you can see the application loading in more details. There are many AppLoaders which goal is to find and load one dynamic library – the "Main" module. Main module has exactly one exported function – ExecuteApp. This function returns when the application must exit so the loader should exit when the function returns. There is parameter to ExecuteApp, which is "Environment Data". The application loader must build and pass correct environment data depending on what is the loader's type.

Before loading the core, Main have to check that all necessary components are available and with correct versions. After that, Main loads Core and gives it the environment data. On its turn, Core passes the environment data to PluginManager.

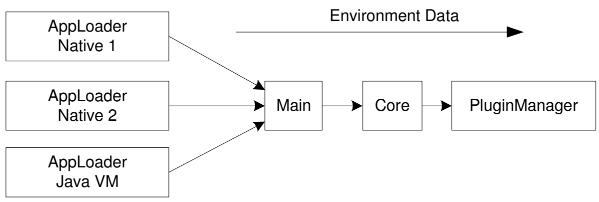


Fig. 4. Application loading

If the loading environment is different from the one necessary for a plug-in, PluginManager creates the necessary environment (i.e. loads the necessary virtual machine).

III. Framework organization

The framework organization from developer's point of view can be seen on Fig. 5. The arrow means "use" or "links to".

There are three major (or distribution) groups: application loaders (AppLoaders), application core and application plug-ins. The application core consists of many dynamic libraries (modules). On Fig. 5 are shown three of them: CoreUtilities, Core and PluginManager. These three modules are the modules from the core that participate in the framework. The application core contains additional modules that use the framework to add functionality to the core.

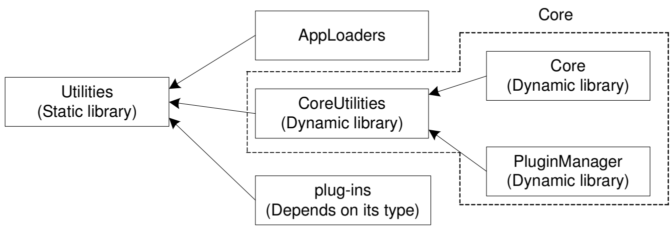


Fig. 5. Framework organization

The module CoreUtilities is used to export the classes in Utilities as dynamic library so the size of the core modules will not increase. Utility modules are modules like: .ini file parser, configuration loader, dynamic libraries manipulator and version support.

IV. Features of the Application Core

A. Component / Proxy architecture

The module Core exports the binding system for the application core. It is implementation of design pattern called "Component / Proxy" [5]. That pattern says: "component is a class that is hidden behind another class – proxy". Components can be accessed only through their proxies.

As you can see on Fig.6, one component can have many proxies but one proxy can point to one component.

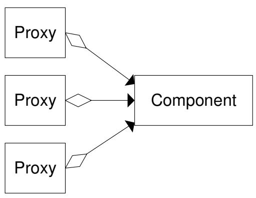


Fig. 6. Component – Proxy relations

It is very important to note – proxies are requested for creation and destruction while creation and destruction of components is automatic. This way it is possible that there are components without proxies as well as proxies without components (dead proxies). It is normal for component to be without proxies but dead proxy is "bad" thing. Such proxy has to simulate some work when its functions are called…

According creation, there are two types of components: singleton (it can live without proxies and there can exist only one instance of these components) and non-singleton (for each requested proxy, one component is created). Component can be destroyed after all its proxies are destroyed (but even though, the component could be left alive). There is manager that controls the lifetime of components and takes requests for proxies – ComponentManager.

The component and proxy creation is done via string identification that contains: Name of the Module (a dynamic library) and Name of the Component. Creation of classes using string identification is similar to reflection systems in Java and .NET.

In the current implementation in addition to the reference counting system, proxy and component lists are used. The benefit is that a component can understand when any of its proxies is destroyed and a proxy can understand if its component is destroyed.

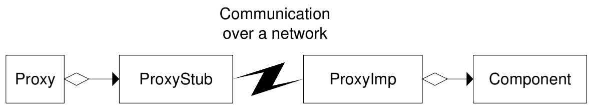


Fig. 7. Remote call between Proxy and Component

"Component / Proxy" architecture has one more benefit – it can be used for remote calls as shown in Fig. 7. For class Proxy, ProxyStub looks like a component. The same way, for class Component, ProxyImp seems to be ordinary proxy. This way Proxy and Component never understand that they are on different computers and are communicating via network. The job of ProxyStub and ProxyImp is to convert function calls to protocol requests / responses.

Static class hierarchy for the described architecture is shown on Fig. 8 (an arrow means inheritance). Class BaseComponentProxy has one pointer to IBaseComponent. This way Proxy and ProxyImp can point to either Component or ProxyStub. Class BaseNetworkProxy has basic functionality for data exchange through network.

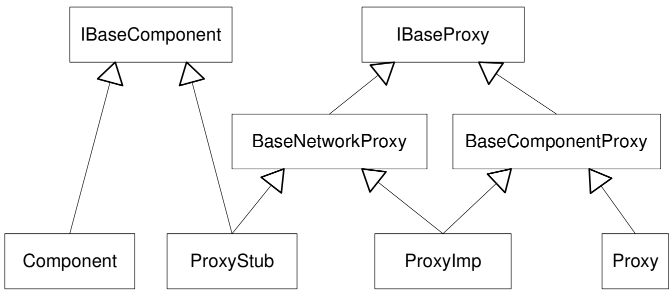


Fig. 8. Class hierarchy of C / P architecture

B. Configuration file

If you look again at Fig. 2 you will see three independent modules that are involved in the application loading: AppLoader, Main and Core. The interface between them is very tight – one function (ExecuteApp, exported by both Main and Core). In order to parameterize the loading and to create storage with common data for these modules there is a little configuration file. It could be parsed up to three times (each of the involved modules may need to parse it). Once the module Core has the information in the configuration file, the whole application core will have it.

In the main configuration file there is reference to another configuration file – for the logging system.

C. Log system

According the destination of the log strings there are 3 output types:

1. File – the log strings are written in a file
2. GUI – the log strings are put in a place where the user can see them
3. Debug – in the debugger output (for debug builds only).

The string formatting for each output type can be different. This way the developer and the user are eased at most.

There are four logging levels: Info, Warning, Recoverable Error and Fatal Error. The application exits on fatal errors…

The log configuration supports different combinations between levels and output types. In addition, there is filter on per-module basis (which modules to include / exclude from logging).

In order to be fast, the log system defines for each possible configuration different (optimized) function. There are 4 function pointers – for each log level. During initialization, these pointers are set according the configuration. The logging is done via call to the necessary function pointer…

D. Plug-in system

The plug-in system was discussed on different places in this article and was well described ideologically. It consists of core module (PluginManager) and export declarations. PluginManager exports strictly defined interfaces for use by different Virtual Machines; or by native plug-ins, written in C or compatible language (C++, Borland's Pascal). The plug-in system is tested with native (C, C++) and VM (Java and .NET) plug-ins.

If the core is loaded by loader for some VM and a plug-in for the same VM is called, PluginManager use the environment of the loader. Otherwise, it loads the VM first and then proceeds with loading and calling the plug-in. The VM loading code is called only if corresponding plug-in is requested. Once VM is loaded, it is used for all plug-ins of its type. The code is protected against missing VM so the user is not obliged to have any VM.

V. EXPERIMENTS AND CONCLUSIONS

The framework that was described so far is implemented in native C++ and is used in Video Editing software (see Fig. 1). The experiments with that software have shown that creating new module for the application core is time-consuming task. In order to ease the creation of new core modules, there is created custom wizard for Microsoft Visual Studio called "CorePackage". This wizard generates project that can be directly compiled to a dynamic library that covers the requirements for core module.

The experiments have shown the following advantages of the framework:

1. It is very open and supports native (C compatible), JavaVM and .NET CLR plug-ins (all plug-in types are tested)

2. It gives optimized, configurable and easy-to-use log system.

3. It has reflection-like system that allows module loading and class creation at runtime determined by string. When used with predefined interfaces (like the interfaces in Fig. 1) the system is powerful enough without being as heavy as Java or .NET reflection.

Allowing each block on Fig. 1 to be plug-in makes the software very dynamic and extensible. In addition, concentrating interfaces and data flow in one place (the application core) gives full control over the data.

Future work – after finishing the framework for the Video Editor, features implementation can be started (stream editor with lazy algorithm and after that – editor for each level on Fig. 1).

REFERENCE

[1] Liang, Shen: "The Java™ Native Interface", Addison Wesley Longman Inc., 1999.

[2] Microsoft Corporation: "Technical Overview of the Common Language Runtime", 2006.

[3] Smart, Julia; Robert Roebling; Vadim Zeitlin; Robin Dunn: "wxWidgets 2.5.5: A portable C++ and Python GUI toolkit", April 2005.

[4] Chakraborty, Ranjeet: "Creating a Host to the .NET Common Language Runtime", article in "The Code Project", October 2001

[5] Gamma, Erich; Richard Helm; Ralph Johnson; John Vlissides: "Design Patterns: Elements of Reusable Object-Oriented Software", Addison Wesley Longman Inc., 1998.

An Influence of the Wavelet Packet Decomposition on the Noise Reduction in Ultrasound Images

Veska M. Georgieva [1], Roumen K. Kountchev [2]

Abstract – **The wavelet packet method is a generalization of wavelet decomposition that offers a richer image analyses. It's made a trial in the paper to analyze the influence of different wavelet packet decompositions over the reduction of specific noise in medical ultrasound images. The optimal decomposition is found with the respect to the entropy-based criteria. Some results of the experiments are presented, which are made by computer simulation in MATLAB environment.**

Keywords – **Wavelet packet analyses, optimal decomposition, entropy-based criteria, noise reduction and medical ultrasound images.**

I. Introduction

The conventional methods of filtration basic on the processing in spatial domain [1], can not give a high quality for the ultrasound images, which needed for precise diagnostic. In noise reduction with wavelet transforms is used the method of the limitation the level of circumstantially coefficients. Giving the determinate threshold of their level and breaking on level the circumstantially coefficients, can be reduced the level of the noise, too. But the most interesting aspect of this problem is that the level of limitation can be determinate separately for every one coefficient. This permits to make adaptive changes of the image.

The wavelet packet methods for denoising give more complete analyses,that provides increased flexibility. They have the following important properties:

- Give a richer presentation of the image, basic on functions with wavelet forms, which consist of 3 parameters: position, scale and frequency of the fluctuations around a given position.
- Propose numerous decompositions of the image, that allows estimate the noise reduction of different levels of its decomposition.
- Allow adaptive denoising on each level of the decomposition by choice of optimal decomposition tree and optimal thresholds parameters.

In the paper is presented an analysis of the influence of the decomposition, based of entropy-criteria, on the specific noise reduction in ultrasound images. This noise is a product of specific phenomena in the human tissues impact of the ultrasound and physiological state of the patient, which can be presented as additive Gaussian white noise [2]. It exists in the diagnostic signals as "grit" in the light as in the dark regions.

The optimal decomposition can be chosen on the base of the best level and minimum of the entropy criteria.

[1]Veska M. Georgieva is with the Faculty of Communication, TU-Sofia, Kl.Ohridsky str.8, Sofia, Bulgaria, E-mail: vesg@tu-sofia.bg

[2]Roumen K. Kountchev is with the Faculty of Communication, TU-Sofia, Kl.Ohridsky str.8, Sofia, Bulgaria, E-mail: rkountch@tu-sofia.bg

Some results of the experiments are presented, which are made by computer simulation in MATLAB environment.

II. Problem Formulation

The wavelet packet analysis is a generalization of wavelet decomposition that offers a richer image analysis.

The wavelet packet atoms are waveforms [3]. Their mathematical description is given in Eq.1.

$$W_{j,n,k}(x) = 2^{-j/2} W_n(2^{-j}x - k), \quad (1)$$

where

$n \in N$ and $(j,k) \in Z^2$, j is a scale parameter, k – time-localization, n – frequency.

The basic idea of the wavelet packets is that for fixed values of j and k, $W_{j,n,k}$ analyzes the fluctuations of the signal roughly around the position $2^j.k$, at the scale 2^j and at various frequencies for the different admissible values of the parameter n.

The set of functions $W_{j,n} = W_{j,n,k}(x), k \in Z$ is (j,n) packet. For the positive values of integers j and n, wavelet packet are organized in trees. For each scale j, the possible values of parameter n are: $0,1,.....2^j - 1$, where j definite the level of decomposition and n - the position on the tree.

For the given orthogonal wavelet functions exists library of bases, called wavelet packet bases. Each of these bases offers a particular way of coding images, preserving global energy, and reconstructing exact features.

Based on the organization of the wavelet packet library, it is determinated the decomposition issued from a given orthogonal wavelets.

A signal of length $N = 2^L$ can be expand in α different ways, where α is the number of binary subtrees of a complete binary tree of a depth L.The result is $\alpha \geq 2^{N/2}$ [4]. As this number may be very large, it is interesting to find an optimal decomposition with respect to a conventional criterion. The classical entropy-based criterion is a common concept. It is looking for a minimum of the criterion. In case of denoising the 2D joint entropy of the wavelet co-occurrence matrix is used as the cost function to determine the optimal threshold. In this case 2D Discrete Wavelet Transform (DWT) is used to compose the noisy image into wavelet coefficients [5].

In the paper is presented on other approach. It is looking for the optimal from three different entropy criteria [6]. They are the followings:

- The norm entropy with $1 \le p$, where P is parameter, given in Eq.2.

$$E_1(S) = \sum_i \sum_j \left| s_{ij} \right|^p , \qquad (2)$$

where S is the image, s_{ij} are the coefficients of S in an orthonormal basis.

- The (none normalized) Shanon entropy. It is given by the Eq. 3.

$$E_2(S) = -\sum_i \sum_j s_{ij}^2 \log(s_{ij}^2) \qquad (3)$$

- The logarithm of the "energy" entropy, given in Eq. 4.

$$E_3(S) = \sum_i \sum_j \log(s_{ij}^2) \qquad (4)$$

By looking for optimal decomposition to noise reduction, on this basis we can formulate the following problems:

- To choose the level for the best tree decomposition.
- To select the optimal entropy, by the minimum of criteria.

To solve these problems, two important conditions must be realized together. The conditions are given in Eq. 5 and Eq. 6:

$$E_K(S) = \min, \; for K = 1,2,3,\ldots n , \qquad (5)$$

where E_K is the entropy in the level K for the best tree decomposition

$$s_{ij} \le t , \qquad (6)$$

where t is the threshold of the coefficients.

By determination of the threshold it is used the strategy of Birge-Massart [7]. This strategy is flexibility and allows to determinate the threshold in three directions: horizontal, vertical and diagonally. In addition the threshold can be hard or soft [4].

III. EXPERIMENTAL PART

The formulated problems are solved by computer simulation in MATLAB, version 6.5 environment with using the WAVELET TOOLBOX.

In analysis are used 20 ultrasound images from cardiology with different sizes. The original images are in different file formats: tiff, jpeg, bmp, but all of them are converted in bmp. All images are 24 bpp, in RGB system.

The simulation is made with additive Gaussian white noise with normal distribution and covariance 0.03.

The experiments are associated with assignment of some basic elements:

- The choice the level of the best decomposition

From the implemented experiments with image of size 640x480, the optimal results are obtained by the second level of the decomposition. In the lower levels of the decomposition the noise reduction is greater but the images are visual with badly quality.

- The assignment of the optimal decomposition, by using the entropy criteria.

The potentiality of the noise reduction it is tested by using of three different entropy criteria, given in Eq. 2,Eq. 3 and Eq. 4. The corresponding optimal decompositions in node (2, 0) are shown in Fig. 1 (Optimal decomposition by Shanon entropy criteria), Fig. 2 (Optimal decomposition by norm entropy criteria) and Fig. 3 (Optimal decomposition by log energy criteria).

The best results are obtained by the (none normalized) Shanon entropy.

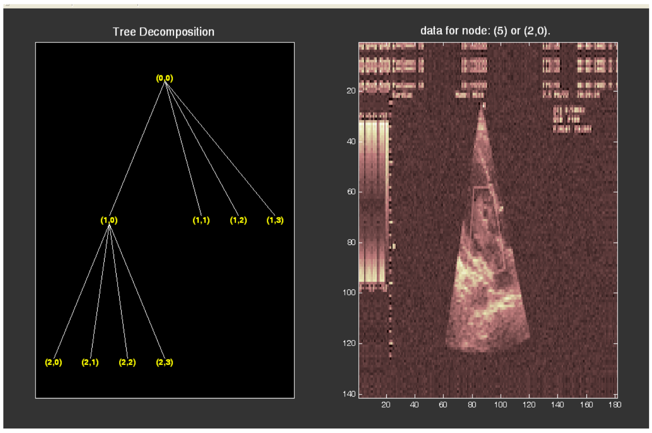


Fig. 1. Optimal decomposition by Shanon entropy criteria

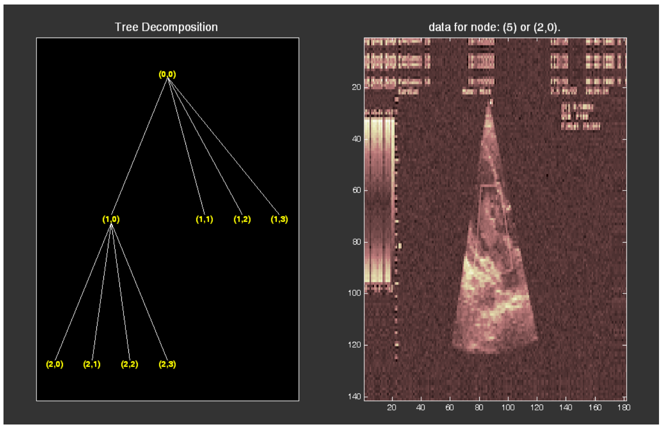


Fig. 2. Optimal decomposition by norm entropy criteria

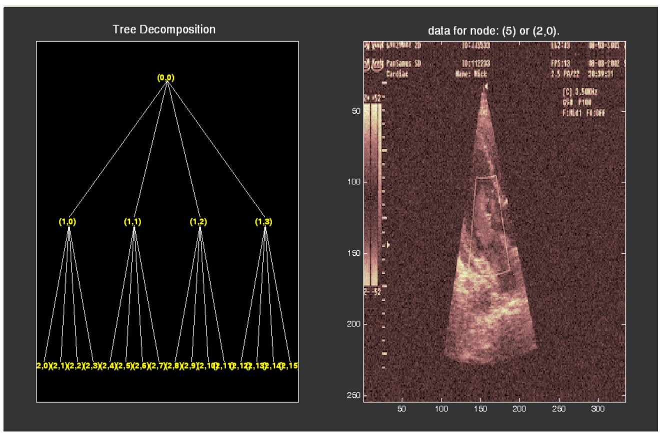


Fig. 3. Optimal decomposition by log energy criteria

The following quantitative estimations are used:

- Coefficient of noise reduction (CNR), given in Eq. 5:

$$K = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\left[x_F(i,j) - x(i,j)\right]^2}{\sum_{i=1}^{N}\sum_{j=1}^{M}\left[y(i,j) - x(i,j)\right]^2}\,, \tag{5}$$

where $x(i,j)$ is a pixel from the original image; $x_F(i,j)$ is a pixel from the filtrated image; $y(i,j)$ is a pixel from the noised image.

- Signal to noise ratio SNR_Y [dB], given in Eq. 6:

$$SNR_y = 10\lg\frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\left[x(i,j)\right]^2}{\sum_{i=1}^{N}\sum_{j=1}^{M}\left[y(i,j) - x(i,j)\right]^2} \tag{6}$$

- Signal to noise ratio SNR_F [dB], given in Eq. 7:

$$SNR_F = 10\lg\frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\left[x_F(i,j)\right]^2}{\sum_{i=1}^{N}\sum_{j=1}^{M}\left[y(i,j) - x(i,j)\right]^2} \tag{7}$$

- Effectiveness of filtration E_{FF} [dB], given in Eq. 8:

$$E_{FF} = SNR_y - SNR_F \tag{8}$$

The results for the quantitative estimations of the noise reduction are obtained by N=640 and M=480, for each component of the system YUV. The results from the simulation of noise reduction are presented in Table 1.

TABLE I
RESULT FROM SIMULATION

Entropy criteria	**YUV system**	**Threshold**	**CNR**	**SNR_Y [dB]**	**SNR_F [dB]**	**E_{FF}**
1	**2**	**4**	**5**	**6**	**7**	**8**
Norm	Y	0.2294	0.3967	5.5595	5.7739	0.2144
	U	0.1844	0.3931	5.6078	5.7893	0.1815
	V	0.2729	0.3899	5.8553	6.0174	0.1621
Shanon	Y	0.2270	0.6351	7.7001	7.9847	0.2846
	U	0.1839	0.0328	-4.5401	-4.7753	0.2352
	V	0.2729	0.0362	-4.4538	-4.6234	0.1694
Log energy	Y	0.2297	0.3999	5.5595	5.5308	0.0287
	U	0.1834	0.4023	5.6078	5.5529	0.0549
	V	0.2734	0.4010	5.8553	5.7981	0.0572

The best results from the simulation of noise reduction are obtained by using the Shanon entropy criteria. By using of the log energy and norm criteria the effectiveness of the filtration is smaller.

In Fig. 4 is presented the noised image and in Fig. 5 we can see the denoised image by using the optimal decomposition on the second level, by using (none normalized) Shanon entropy criteria. The noise reduction is realized by hard threshold.

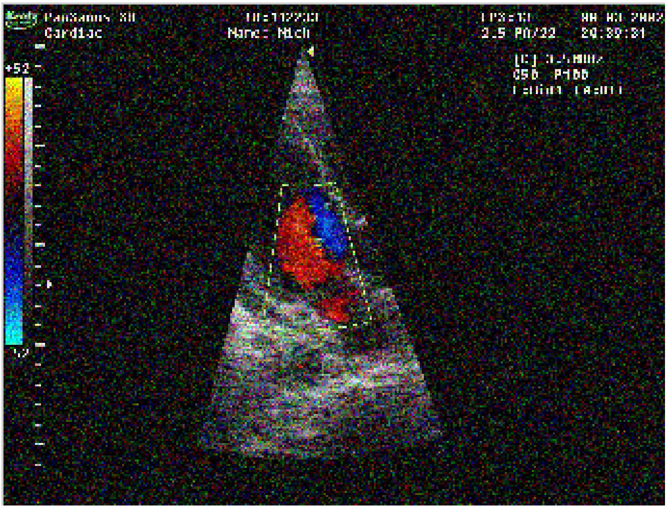


Fig. 4. Noised image

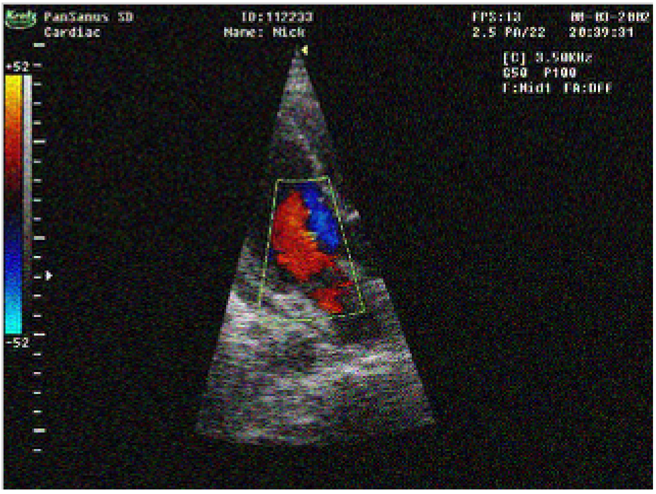


Fig. 5. Denoised image

IV. Conclusion

The problem of the optimal decomposition is important in noise reduction of ultrasound images. In the paper is presented this influence, by changing the level of decomposition and the entropy criteria. On this base is improved the advanced approach for specific additive noise reduction, using wavelet packet transformations [8]. The thoroughly advantages are:

- It is possible to reduce the specific noise in medical ultrasound images by preservation the high quality of the restored images.
- It allows a selective approach by noise reduction on each level of decomposition by choice of optimal parameters of the using wavelet packet transformations.
- It can be used by medical image processing, archiving and making of data basis.

REFERENCES

[1]. W. Pratt, Digital image processing, New York: John Wiley and Sons, 2001

[2]. Junji Maclin, Benhard Siegel, Ultrasound for Surgeons, IGAKU, N.Y., 1997

[3]. Rao Raghuveer M., Bopardicar Ajit S., Wavelet Transforms: Introduction to Theory and Applications, 1998, http://www.awl.com

[4]. D. Donoho, I. Johnston, Adapting to unknown smoothness via wavelet shrinkage, Jornal Am. Stat. Assos., 1995, 90:1200-1224

[5]. Zeyong Shan, Selin Aviyente, Image denoising based on the wavelet co-occurance matrix, IEEE Trans. on Image Processing, Vol.9, no.9, pp.1522-1531, 2000

[6]. [6] R. Coifmann, M. Wickerhauser, Entropy-based Algorithms for best basis selection, IEEE Transaction on information theory, Vol.38, №27 pp.713-718, March 1992

[7]. [7] MATLAB 6.5, User's Guide, www.mathwork.com

[8]. [8] R. Kountchev, V. Georgieva, An approach for noise reduction in medical ultrasound images, Proceedings of the TU-Sofia, Vol.55, 2005, pp. 159-166

Fingerprint Enhancement by Adaptive Filtering in Frequency Domain

Brankica M. Popović[1]

Abstract – **Fingerprint image enhancement is an essential preprocessing step in fingerprint recognition applications. A method described in this paper is based on the adaptive filtering in frequency domain using Gabor filters. Algorithm based on varianse calculation is presented and compared with one based on orientation field estimation .**

Keywords – **fingerprints, enhancement, image processing, image filtering, Gabor filter**

I. Introduction

Reliable person identification is an important problem in diverse businesses such as finance, law enforcement, access control, health care etc. Often, it is necessary to do it remotely and automatically. Biometrics, as automatic identification of people based on their physiological or behavioral characteristics, are becoming dominant over traditional means of authentication such as knowledge-based (password) and token-based (key) authentication. Among several human characteristics that can be used in biometric systems (face, retina, iris, voice, hand etc.), fingerprints are, due to their characteristics, one of the most researched, used and mature method of authentication. [1]. They have been extensively used by forensic experts in criminal investigations for decades. [2]. Popularity of fingerprint identification is based on the following properties of fingerprints: the fingerprint of a person is unique and its features remain invariant with age [3, 4].

There were different realizations of automatic fingerprint identification systems (AFIS) in the last 50 years. We can either digitalize fingerprint image taken by ink, or use inkless optical scanners to provide input image for AFIS. In either way spatial resolution of 500 dpi and amplitude resolution of 256 gray values (8 bits per pixel) are recommended. A number of operations are applied in order to extract features later used in matching process.

The goal of feature extraction in pattern recognition system (in general) is to extract information from the input data that is useful for determining its category. In the case of fingerprints a natural choice are features based directly on the fingerprint ridges and ridge-valley structure. However, the effectiveness of a feature extraction depends greatly on the quality of the images. Consequently, fingerprint image enhancement is usually the first step in most AFIS.

The rest of this paper is organized as follows. In Section II we briefly describe fingerprint structure and some methods for fingerprint image enhancement. In Section III 2-D Gabor filter is introduced, and adaptive filtering is described in Section IV. Some results and conclusions are shown in Section V.

II. Fingerprint structure

A fingerprint represents the image of the surface of the skin of the fingertip. A typical structure of a fingerprint consists of ridges (black lines) separated by valleys.

The ridge pattern in a fingerprint can be described as an oriented texture pattern with fixed dominant spatial frequency and orientation in a local neighborhood. The frequency is depending on inter-ridge spacing, and orientation on flow pattern exhibited by the ridges. The global pattern of fingerprint is used to determine the class [5]. Classification process was performed manually in police departments from the end of the 19th century [6], and is, with small modifycations, still in use today. Region of a fingerprint where the ridge pattern makes it visually prominent are called singularities [7]. There are two types of fingerprint singularities: core and delta, and they are very useful for determining fingerprint's class.

A closer analysis of the fingerprint reveals some anomalies of the ridges, such as ridge endings, bifurcations, crossovers, short ridges, etc. These local features of fingerprints, called minutiae, can be used for manual or automatic fingerprint identification. These basic features of fingerprints (singularities and minutiae) are shown in Fig. 1.

Fig. 1. Basic features of fingerprints: minutia- ridge ending (in square), ridge bifurcation (in circle), singularities core (X) and delta (in triangle)

Enhancement may be viewed as a process of improving the clarity of the ridge structure in the fingerprint image [8, 9, 10]. Result is expected to be more suitable than the original, for visual examination and automatic feature extraction. Although noise content is reduced, enhancement process can also introduce false ridges, resulting in false or missing minutiae.

[1] Brankica M. Popovic is with the Police Academy, Humska 22, 11000 Belgrade, Serbia and Montenegro, E-mail: cica@ptt.yu

There have existed a variety of algorithms for image enhancement that can be used in case if fingerprint images. One group of algorithms, such as change of contrast and histogram equalization [11], are dealing with each pixel independly. Other algorithms are considering local neighbourhood of pixel. Typicall algorithms from this class are linear filtering in spatial and frequency domain, nonlinear filtering, adaptive filtering etc.

Since in local area, the ridges and vallyes have well-defined frequency and orientation, it is natural to use oriented filters. There were some research in directional filtering in spatial domain [3, 12]. In frequency domain Fourier transforms [9, 13, 14], and Gabor filters [10, 15] are analyzed. If filters in use are adapted to ridge orientation, filtering process is called adaptive.

III. 2 - D GABOR FILTER

The Gabor filters are recognized as a very useful tool in computer vision and image processing applications such as texture analysis, image compresion etc. Gabor filters ere very useful both in frequency and spatial domain, due to their frequency-selective and orientation-selective properties. Impulse response of these filters, which are by the way band-pass filters, are very similar to impulse response of receptive fields in the brain's visual cortex [16]. By simple adjustment of mutualy independent parameters, Gabor filters can be configured for different shapes, orientations, different width of band pass and different central frequences. Properly tuned, Gabor filter can filter an image, maintaining only regions of a given frequency and orientation, and this has profound implications for research in fingerprint image analyse and enhancement using this filter [10, 15].

An even Symmetric Gabor filter general form in the spatial domain is [10]:

$$h(x, y, \phi, \omega) = e^{-0.5\{[x/\delta_x]^2+[y/\delta_y]^2\}} \cos[\omega(x\cos\phi + y\sin\phi)] \quad (1)$$

where ϕ is the orientation of the Gabor filter, ω is the frequency of the sinuusoidal plane wave along the x –axis, δ_x and δ_y are the standard deviations of the Gaussian envelope along the x and y axes, respectively.

Fig. 2 shows an example of Gabor filter and its response in spatial and frequency domain.

Parameters ω, δ_x i δ_y , for optimal Gabor filter depends on inter-ridge distance in fingerprint image. Since the variations of inter ridge distance, in the fingerprint database avaliable to the author, is small (distance is about 5 pixels,f=1/5=0.2) , we set parameter ω to be $\omega=2\pi\times0.2$, and $\delta_x = \delta_y = 4.0$. Trade-off in selection of δ_x and δ_y is done based on empirical data [10], so that the filter is robust to noise, but still can capture ridge information at fine level. 16 different orientations are examined, creating Gabor filter bank with orientations $\phi = i\pi/16$.

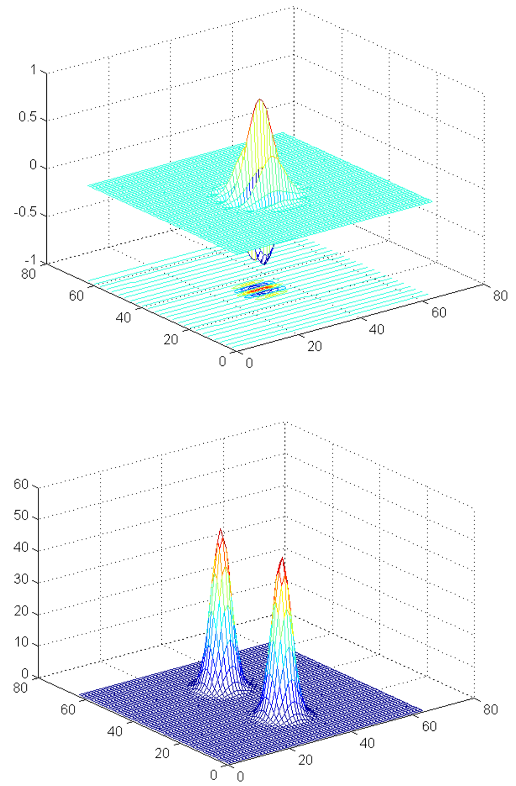


Fig. 2. The Gabor filter and its response in spatial and frequency domain

IV. ADAPTIVE FILTERING

Bank of oriented Gabor filters are applied to input fingerprint image. Filtering is performed in frequency domain resulting in set of 16 filtered images, where each of them emphasizes one ridge orientation (namely $\phi = i\pi/16$). Those filtered image are combined in order to get enhanced image. We applied and examined two different methods.

First we applied method described in [10]. Additional information about dominant ridge orientation is used. We create block-direction image by determination dominant orientation in blocks dimension of $w\times w$, by the following formula [17]:

$$\theta_d = \frac{1}{2}\tan^{-1}\left(\frac{\sum_{i=1}^{w}\sum_{j=1}^{w} 2G_x(i,j)G_y(i,j)}{\sum_{i=1}^{w}\sum_{j=1}^{w}\left(G_x(i,j)^2 - G_y(i,j)^2\right)}\right), G_x, G_y \neq 0 \quad (2)$$

where, G_x and G_y are the gradient magnitudes obtained using 3×3 sobel masks. Smoothing process presented in [5] is then applied.

In enhancement process, pixels in one block of enhanced image take the value of pixels on the same position from the filtered image which emphasizes determined orientation for corresponding block.

Second realization is based on following idea:

- Each Gabor filter clearly emphasize ridge in certain direction, while the rest has no specific texture.
- The variance of the pixel intensities in block is high if there are ridges and valleys in those block, otherwise variance is low.

We propose the following enhancement process:

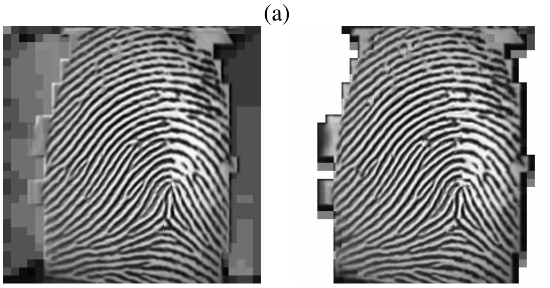
- For each filtered image we calculate the variance of intensities in block, in order to get rank of variance values;
- Pixels in one block of enhanced image take the value of pixels on the same position from the filtered image where maximal variance for the block is calculated;
- To avoid chess-field effect, we combined blocks (with appropriate ponder factor), of three filtered images: one with highest calculated variance and two other that are next in the rank.

V. RESULTS AND CONCLUSION

Original image and enhanced images obtained by described methods are shown in Fig. 3.

In order to compare effectiveness of presented methods, we applied algorithm for minutiae extraction described in [18]. Then we compared number of extracted minutiaes with the one determined by the expert. Thinned binarized images with extracted minutiae are shown in Fig. 4.

(a)



(b) (c)

Fig. 3. (a) Original image, (b) enhanced image obtained by first method [10] (using information of block orientation), (c) enhanced image obtained by second method (max variance in block)

In enhanced image obtained by applying first method, total of 54 minutiae is found, opposite to 47 found by the expert. Among them, 36 are matched, 11 are missing, and 18 are false minutiae.

(a) (b)

Fig. 4. Thinned binarized enhanced image
(a) obtained by first method
(b) obtained by second method

In enhanced image obtained by applying second method, total of 57 minutiae is found, 39 are matched, 8 are missing, and 18 are false minutiae.

At the same time if we apply algorithm for minutiae detection directly on original image, a total of 128 minutiae is found, 36 are matched, 11 missing and 92 are false.

Relatively high number of false minutiae both in enhanced images, and in original one, is mostly result of binarization and thinning algorithms applied in minutiae extraction algorithm.

We can see what there a clear benefit in using presented enhancement techniques is. Although the number of false minutiae is still too high, it is reduced multiple times.

Similar results are obtained for other images in our database. So what method should be applied?

It seems logical that if we already spent some time in orientation image estimation, we should use that information without wasting more time in variance calculation.

However, second method is competitive in results with the first one. Since most algorithms for segmentation background from foreground of fingerprint image are based on variance calculation in block, it is optimal to use that information to perform enhancement as well. Since in that case there is no need for orientation image estimation, some time is obviously saved.

Also, some modification of second method can be used in artifacts removal (letters, interruptions etc.), and that will be object of our further research. We are also considering including the information of maximal dispersion of the pixel intensities in blocks, in enhancement process.

REFERENCES

[1] B. Popović, “Use of biometric systems in fighting organized crime”, *Organized Crime: Facts and measures of protection Meeting Proceedings*, pp. 810-831, Belgrade, Serbia and Montenegro, 2005

[2] Federal Bureau of Investigation. *The Science of Fingerprints: Classification and Uses*, Washington, D.C., U.S. Government Printing Office, 1984

[3] B.M. Mehtre, “Fingerprint image analysis for automatic identification”, *Machine Vision and Applications*, vol. 6, pp. 124-139, 1993

[4] B. Miller, “Vital signs of identity”, *IEEE Spectrum*, vol. 31, no. 2, pp. 22-30, 1994

[5] M. Popović, B. Popović, “Computer classification of fingerprints”, *NBP*, vol. 2, no. 2, pp. 107-120, Beograd, 1997

[6] E. Henry. *Classification and uses of finger prints*. Routledge, London, 1900

[7] N. Yager, A. Amin, “Fingerprint verification based on minutiae features: A review”, *Pattern Anal Applic.* Vol. 7, pp. 94-113, 2004

[8] T.F. Krile, J.F. Walkup, “Enhancement of fingerprints using digital and optical techniques”, *in Image Analysis Applications*, Eds. R. Kasturi, M.M. Trivedi, M. Dekker, New York, 1990

[9] B.G. Sherlock, D.M. Monro, K. Millard, “Fingerprint enhancement by directional Fourier filtering”, *IEEE Proc. Vision Image Signal Process.*, vol. 141, no. 2, pp. 87-94, 1994.

[10] L. Hong, Y. Wan, A.K. Jain, “Fingerprint image enhancement: Algorithm and performance evaluation”, *IEEE Trans. Pattern Anal. Machine Intell.* Vol. 20(8), pp. 777-789, 1998

[11] K.R. Castelman. D*igital Image Processing*. Prentice Hall, Englewood Cliffs, 1996

[12] L. O'Gorman, J.V. Nickerson, “An approach to fingerprint filter design”, *Pattern Recognition*, vol. 22, no.1, pp.29-38, 1989

[13] T. Kamei, M. Mizoguchi M, “Image Filter Design For Fingerprint Enhancement”, *Int. Symp. Computer Vision – ISCV Proceedings*, pp. 109-114, Coral Gables, FL, USA, 1995

[14] A. Willis, L. Myers, “A cost-effective fingerprint recognition system for use with low-quality prints and damaged fingertips”, *Pattern Recognition,* vol. 34, pp. 255-270, 2001

[15] J. Yang, L. Lin, T. Jiang, Y. Fan, “A modified Gabor filter design method for fingerprint image enhancement”, *Pattern Recognition Letters*, vol. 24, pp. 1805-1817, 2003

[16] J.G. Daugman, “Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters”, *J. Optical Soc. Amer.*, vol 2 (7), pp. 1160-1169, 1985

[17] N.K. Ratha, S. Chen, and A.K. Jain, “Adaptive flow orientation-based feature extraction in fingerprint images”, *Pattern Recognition*, vol. 28, no. 11, pp. 1657-1672, Nov. 1995.

[18] B. Popović, M. Popović, “Automatic fingerprint feature extraction for person identification", *3rd DOGS, Conference Proceedings*, Novi Sad, Yugoslavia, 2000

Improved Spatial-Temporal Moving Areas Detection Method Resistant to Noise

Vesna Zeljkovic[1] and Dragoljub Pokrajac[2]

Abstract — **We discuss the resilience of the spatial temporal moving objects detection algorithm on various types of additive and multiplicative noise. Video is decomposed into the spatio-temporal blocks and dimensionality reduction is used to suppress the influence of noise. The moving object detection algorithm based on spatial temporal analysis is subsequently applied.**

Keywords — **Video analysis, Motion Detection, Principal Component Analysis, Noise Reduction.**

I. Introduction

We evaluate the performance of the improved spatial-temporal moving object detection method, resistant to noise. Our main goal is to demonstrate that this novel technique is resistant to influence of various types of noise and to augment the reasons for such desirable behavior.

A common feature of the existing approaches for moving objects detection is the fact that they are pixel based [1-5]. General drawback of these methods is their sensitivity on various types of noise that may exist in video frames due to influence of insufficient illumination, video amplifier, lens system, electromagnetic interference, etc. Recently, a moving object detection algorithm based on principal components analysis is proposed [6]. It is demonstrated [7, 8] that the application of principal component analysis can contribute to significant reduction of the number of false positives (background objects falsely labeled as moving).

In this paper, we propose to apply principal component projection as preprocessing step of our pixel-based technique for detection of moving objects [9]. With the proposed preprocessing step, we combine the pixel and region texture information. More precisely, we decompose a given video into overlapping spatiotemporal blocks, e.g., 5x5x3 blocks, and then apply a dimensionality reduction technique to obtain a compact scalar representation of color or gray level values of each block. Subsequently, the illumination robust moving object detection algorithm that detects and tracks moving objects is applied.

Observe that we go away from the standard input of pixel values that are known to be noisy and the main cause of instability of video analysis algorithms. In contrast, the application of principal components instead of original vectors is expected to retain useful information while suppressing successfully the destructive effects of noise [10].Hence, we have anticipated that the proposed technique will provide motion detection robust to various types of noise that may be present in video sequence. In our earlier paper [11] we demonstrated that the proposed technique is resilient to various levels of Gaussian noise. This paper demonstrates the robustness of the proposed technique to other noise types that may exist in video frames on a test video sequence from PETS repository', (available at ftp://pets.rdg.ac.uk/)

II. Methodology

The technique for moving object detection we use consists of the following major phases: extraction of the 3D filter coefficients with the PCA analysis; dimensionality reduction by spatiotemporal blocks; image filtering of a current frame with the noise removal filter; and detection of moving objects applying the pixel based method for moving object detection resistant to illumination changes.

We treat a given video as three-dimensional (3D) array of gray pixels $p_{i,j,t}$, i=1,…,X; j=1,…,Y; t=1,…,Z with two spatial dimensions X, Y and one temporal dimension Z. We use spatiotemporal (3D) blocks represented by N-dimensional vectors $\mathbf{b}_{I,J,t}$, where a block I, J spans ($2T$+1) frames and contains N_{BLOCK} pixels in each spatial direction per frame ($N=(2T+1)\times N_{BLOCK}\times N_{BLOCK}$). To represent the block vector $\mathbf{b}_{I,J,t}$ by a scalar while preserving information to the maximal possible extent, we use principal component analysis [10] (see Fig. 1 for illustration).

In principal component analysis, we estimate sample mean and covariance matrix of representative sample of block vectors corresponding to the considered types of movies and use the first eigenvector of the covariance matrix **S** (corresponding to the largest eigenvalue). This eigenvector represents the coefficients of the 3D filter used for dimensionality reduction (that suppresses the noise). In practical realizations, the 3D filter can be emulated by three 2D filters applied on three subsequent frames.

After the dimensionality reduction, we apply the following pixel based algorithm for moving object detection and tracking [9]. Consider image sequence S consisting of N video frames. The sliding mask A_t is applied on every frame t.

We calculate the pixel variance in order to estimate the potential movement in the observed area, as follows:

$$\sigma_t^2(I)=\frac{1}{card\{A_t\}}\sum_{m\in A_t}\left(\frac{B_m}{C_m}K_t-median\{A_t\}\right)^2, \qquad (1)$$

[1]Vesna Zeljkovic is with Applied Mathematics Research Center, Delaware State University, 1200 N DuPont Hwy, Dover DE 19901, USA, E-mail: vzeljkovic@desu.edu

[2]Dragoljub Pokrajac is with Computer and Information Science Department and with Graduate Department of Applied Mathematics and Theoretical Physics, Delaware State University, 1200 N DuPont Hwy, Dover DE 19901, USA, E-mail: dpokraja@desu.edu

where pixel intensities within mask A_t are denoted with B_m for a reference—background frame that does not contain changing regions and with C_m for a current frame (where we are identifying moving objects). The estimated mean of the pixel intensity ratio within A_t is denoted with μ_{A_t} .

The illumination compensation coefficient is defined as

$$K_t = \frac{\sum_{m \in A_t} C_m}{\sum_{m \in A_t} C_{1m}} = \frac{\mu_t}{\mu_1}, \qquad (2)$$

where C_{1m} is pixel intensity for the first frame in the sequence.

The algorithm performs the analysis in time and space domains simultaneously, contributing to its resistance to the illumination changes and reducing the false detection, i.e. artifacts. We average estimated pixel variances for three successive pairs of frames and threshold this average to determine the presence of moving objects. This represents temporal aspect of analysis. The ratio of pixel intensities in A_t between two frames is used to estimate the pixel variance $\sigma_t^2(I)$ for the following three pairs of successive frames: frames t-3 and t-2, t-2 and t-1 and t-1 and t, where t is the current frame. Thus, we obtain three pixel variances for three successive and corresponding variance pair $\sigma_{t-2}^2(I)$, $\sigma_{t-1}^2(I)$ and $\sigma_t^2(I)$. Subsequently, we compute the average value of these three variances as follows:

$$\bar{\sigma}^2(I) = (\sigma_{t-2}^2(I) + \sigma_{t-1}^2(I) + \sigma_t^2(I)) / 3, \qquad (3)$$

After that, the mean value is subtracted from the pixel variance of the current and previous frame:

$$\sigma_t^2 * (I) = \sigma_t^2(I) - \bar{\sigma}^2(I) \qquad (4)$$

If $\sigma_t^2 * (I) \geq \varepsilon$ (a suitable threshold), the center of A_t is marked as changing region, i.e. as a moving area.

The proposed algorithm performs analysis in three dimensions, two spatial and one temporal. The time analysis, additionally to space analysis, helps with correct moving object detection and augments the precision of the algorithm. This algorithm, it should be pointed out, exploits local intensity and does not require color information. Hence, the method can easily be used for gray movies, i.e. BW movies or infra-red monochromatic movies.

III. Results

We have demonstrated the performance of the proposed approach on sequences from the Performance Evaluation of Tracking and Surveillance (PETS) repository[1]. Processed video-sequences are available on our web site: http://ist.temple.edu/~pokie/data/ACIVS2005. Here, we present results on a video sequence from PETS2001[1] (here referred to as the *Outdoor video* sequence). Video consisted of 2688 frames PAL standard (25 frames per second, 576×768 pixels per frame). In our experiments we use $N_{BLOCK} = 5$, thus the length of a block vector $\mathbf{b}_{I,J,t}$ is $N = 75 = 5\times5\times3$.

We experimented with additive Gaussian, Salt&Pepper, multiplicative ("speckle") and Poisson noise [12]. The additive Gaussian noise was zero mean, with variance ranging from 0.1 to 0.5. The Salt&Pepper noise densities varied from 0.05 to 0.2. The variance of the speckle noise ranged from 0.1 to 0.5. In Fig. 2, we demonstrate the effects of selected levels of each noise type on frame 2500 of the *Outdoor video* sequence.

The result of the proposed algorithm on Frame 2500 of the *Outdoor video* sequence is illustrated in Fig. 3, for the same types and intensities of noise.

As we can see, the proposed technique is able to successfully and precisely detect moving objects even in case of relatively strong noise influence. The moving car in foreground and slow-moving white van are identified as well as the pedestrian on the left part of the scene. There are no false moving objects identification and no artifacts are introduced.

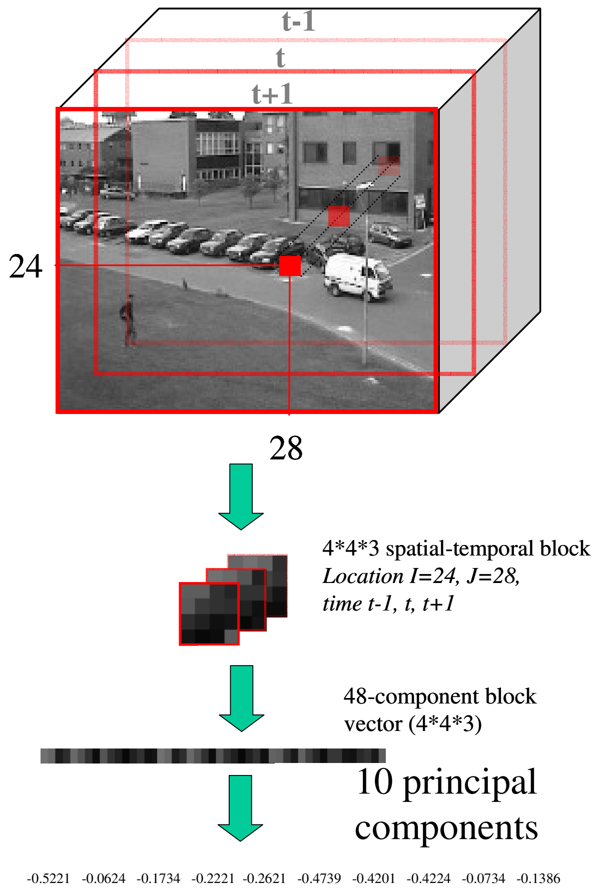


Fig. 1. Dimensionality reduction using spatial-temporal blocks.

[1]ftp://pets.rdg.ac.uk/PETS2001/DATASET1/TESTING/CAMERA1_JPEGS/

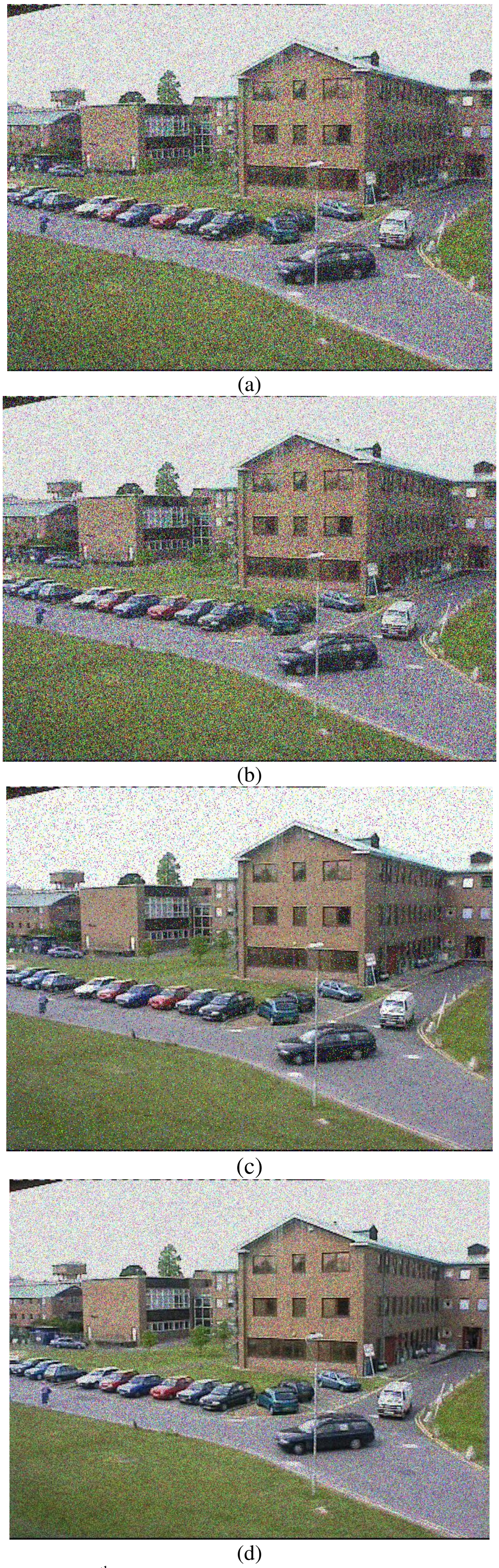


Fig. 2. The 2500th frame of *Outdoor video* under a) Gaussian zero-mean noise with variance 0.1; b) Salt and paper noise with density 0.1; c) Poisson noise; d) Speckle noise with variance 0.1.

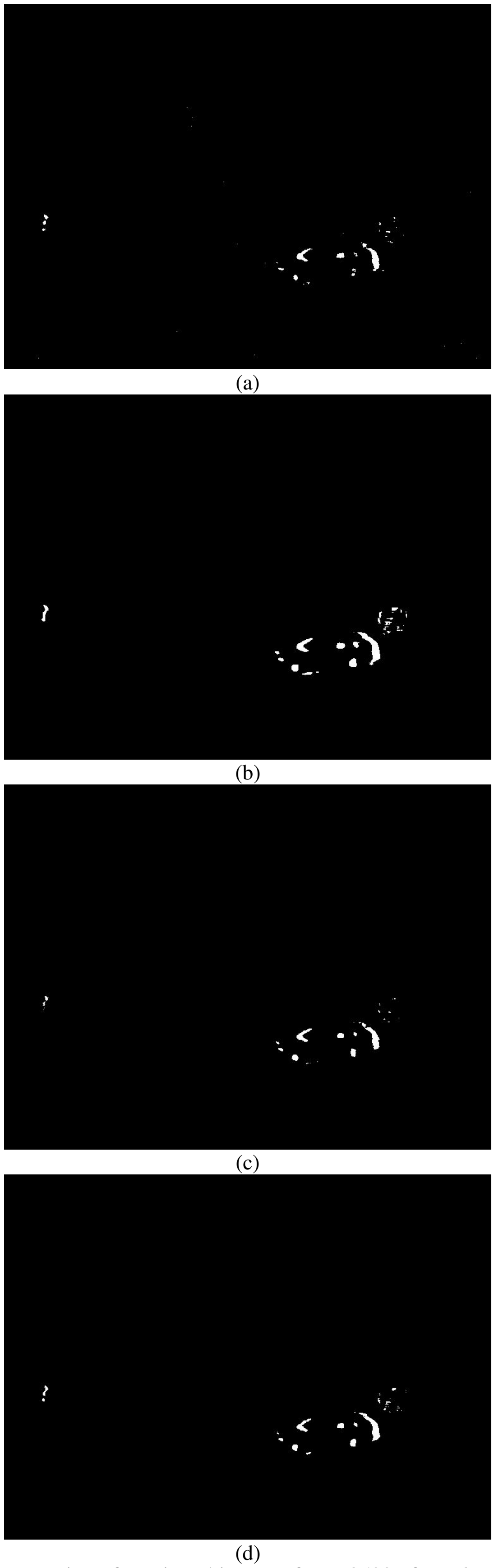


Fig. 3. Detection of moving objects on frame 2500 of *Outdoor video under* a) Gaussian zero-mean noise with variance 0.1; b) Salt and paper noise with density 0.1; c) Poisson noise; d) Speckle noise with variance 0.1

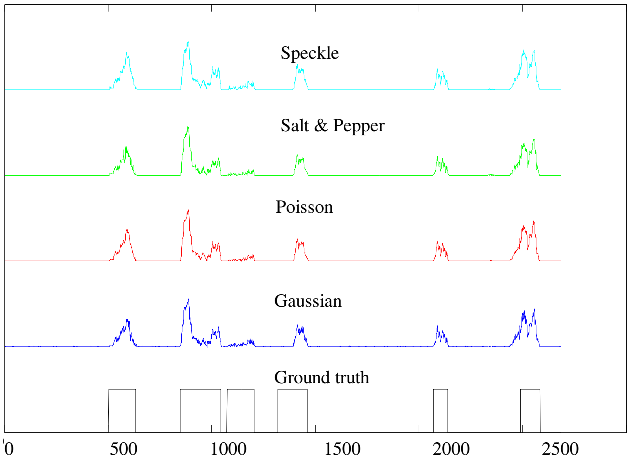


Fig. 4. Percentage of identified moving objects at spatial window (350,510; 500,600) calculated for *Outdoor* video with Gaussian zero-mean noise with variance 0.1, Poisson noise, Salt and paper noise with density 0.1, Speckle noise with variance 0.1, compared with hand-labeled ground truth (presence of moving object in the video as observed by a human)

To demonstrate the influence of varying noise levels on the performance of our algorithm, we computed spatial-windows based evaluation statistics. We counted the number of identified moving block within a pre-specified spatial window and normalized it with the number of spatial blocks in the same window. We hand-labeled the observed spatial window by denoting time intervals when a moving object is present in the window in order to compare the result of automatic detection of moving objects with "ground truth".

In Fig. 4, we show the computed statistics for sequence with Gaussian zero-mean noise with variance 0.1, Poisson noise, Salt and paper noise with density 0.1, and Speckle noise with variance 0.1, as well as ground truth moving objects detection in rectangular region (350, 510; 500, 600).

It can easily be observed that, in spite of various types of noise, it is still possible to detect a moving object in a window by properly thresholding the observed statistics. Observe also that such identification agrees with the ground truth.

IV. Conclusions

In this paper we have demonstrated that our moving object detection algorithm based on spatiotemporal blocks and linear variance-preserving dimensionality reduction is resistant on the influence of various types of noise.

We evaluated performance of the applied algorithm on benchmark videos from Performance Evaluation of Tracking and Surveillance (PETS) repository. As a performance measure we, in addition to a visual evaluation, used spatial-windows based evaluation statistics compared to human-identified ground truth moving objects detection.

The results indicate that a proper detection is still possible in spite of significant levels of additive or multiplicative noise. As we experimentally shown, this could be explained by inherent capability of employed dimension reduction techniques to efficiently suppress noisy component. Our work in progress is oriented towards theoretical justification of such behavior. Also, we would like to point out that we have not experimented with the influence of moving object velocity. That is also the part of our work in progress.

Acknowledgement

V. Zeljkovic has been partially supported by DoD HBCU/MI Infrastructure Support Program (45395-MA-ISP Department of Army). D. Pokrajac has been partially supported by NIH-funded Delaware IDeA Network of Biomedical Research Excellence (INBRE) Grant, DoD HBCU/MI Infrastructure Support Program (45395-MA-ISP Department of Army), National Science Foundation (NSF) Infrastructure Grant (award # 0320991) and NSF grant "Seeds of Success: A Comprehensive Program for the Retention, Quality Training, and Advancement of STEM Student" (award #HRD-0310163).

References

[1] Jain, R., Militzer, D., and Nagel, H. "Separating Nonstationary from Stationary Scene Components in a Sequence of Real World TV Images", In Proc. International Joint Conference on Artificial Intelligence IJCAI 77 (Cambridge, MA, 1977), 612–618.

[2] I. Haritaoglu, D. Harwood, and L. Davis, "W4: Real-Time Surveillance of People and Their Activities", IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI) 22(8) (2000), pp. 809–830.

[3] N. M. Oliver, B. Rosario, and A. P. Pentland, "A Bayesian Computer Vision System for Modeling Human Interactions", IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI) 22(8) (2000), pp. 831–843.

[4] Remagnino, P., G. A. Jones, N. Paragios, and C. S. Regazzoni, eds., Video-Based Surveillance Systems, Kluwer Academic Publishers, 2002.

[5] C. Wren, A. Azarbayejani, T. Darrell, and A.P. Pentland, "Pfinder: Real-time Tracking of the Human Body", IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI) 19(7) (1997), pp. 780–785

[6] Pokrajac, D., Latecki, L. J. "Spatiotemporal Blocks-Based Moving Objects Identification and Tracking", In Proc. IEEE Int. Workshop Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS), Nice, France, 2003.

[7] V. Zeljković, D. Pokrajac, L.J. Latecki, "Noise Robust Spatial-Temporal Algorithm for Moving Objects Detection", proc. 49th ETRAN Conf., 2005, in press.

[8] D. Pokrajac, V. Zeljković, L.J. Latecki, "Noise-Resilient Detection of Moving Objects Based on Spatial-Temporal Blocks", proc. 47th ELMAR Conf., 2005, in press.

[9] V. Zeljkovic, D. Pokrajac, A. Dorado, and E. Izquierdo, "Application of the Improved Illumination Independent Moving Object Detection Algorithm on the Real Video Sequence", In Proc. 6th International Workshop on Image Analysis for Multimedia Interactive Services, WIAMIS 2005.

[10] Jolliffe, I. T, Principal Component Analysis, 2nd edn., Springer Verlag, 2002.

[11] D. Pokrajac, V. Zeljković, "Influence of the Gaussian Noise Spatial-Temporal Method for Moving Objects Detection", submitted to 11th International Conference on Computer Analysis of Images and Patterns (CAIP) 2005.

[12] Gonzalez, R., Woods, R., Digital Image Processing, Prentice-Hall, 2002.

Designing of scalar quantizer based on the hybrid model for the Laplacian source

Zoran H. Peric[1], Jelena R. Nikolić[2]

Abstract – **In this paper we suggest a hybrid model for scalar quantizer construction in order to achieve the quantization performances arbitrarily close to those of the optimal scalar quantization. This model is based on the combination of two quantization techniques. One of them is the companding technique and the other one is the Lloyd-Max's procedure for designing scalar quantizers. The suggested mixed technique has low implementtation complexity. Furthermore, it has a little bit greater complexity than companding technique, but enables better quantizer's performance, very close to those of optimal Lloyd-Max's scalar quantizers. The recommended hybrid model presents the general model of the scalar quantizer considered in [1].**

Keywords – **Hybrid model, Laplacian source, scalar quantizer**

I. Introduction

One of the important issues from the engineer's point of view is the design and implementation of quantizers to meet the performance objectives. Lloyd [2] and Max [3] respectively proposed an algorithm to compute optimum quantizers using mean-square error distortion measure. Namely, they gave the nonlinear quantization procedure in order to minimize the quantization noise. The primary goal when designing an optimal Lloyd-Max's quantizer is to select the representation levels and the decision thresholds so as to provide the minimum possible average distortion for a fixed number of quantization levels N. Particularly, Lloyd-Max's algorithm is an iterively algorithm, which performs in each iteration calculation of all representation levels and decision thresholds of the N levels scalar quantizer. The size of necessary calculation is the deficiency of this algorithm, especially when designing scalar quantizers with a large number of quantization levels. However, a nonuniform quantization can also be achieved by compressing the input signal, than quantizing it with a uniform quantizer and expanding the quantized version of the compressed signal using a nonuniform transfer characteristic inverse to that of the compressor. The described quantization technique is called the companding technique. Bad approximation of the input signal in the region of high amplitudes is the deficiency of the companding technique which is the consequence of the extensive outermost cells. In order to reduce the size of the necessary calculation in comparision to that for Lloyd-Max's algorithm, as well as to improve the deficiency introduced by using the companding technique, we suggest one model, denoted here as the *hybrid model*. This model is based on the combination of the companding technique and the Lloyd-Max's procedure of designing scalar quantizers. Namely, applying the companding technique to N-2L inner cells and Lloyd-Max's procedure to 2L outer cells it is possible to design N-levels scalar quantizer. The suggested model is very simple for analysis and it provides an almost optimal design of scalar quantizers. Therefore, in this paper we continue the research in the field of finding as simple as possible method for designing optimal scalar quantizers. We perform an exact and complete analysis of the hybrid model considering the Laplacian input signals. Furthermore, we derive the expression for determining the decision thresholds and the representation levels of the considered scalar quantizer. Thus, knowing the decision threshold t_{2K-1} it is possible to determine the support region of the observed scalar quantizer, ranging $(-t_{2K-1}, t_{2K-1})$. Optimal determining of the support region has been considered by a lot of researchers [1], [4], [5]. The problem of determining the support region was considered by Sangsin Na and David Neuhoff [1]. They applied the companding technique when they calculated the values of the representation levels and the decision thresholds of N-2 inner cells and Lloyd-Max's procedure when calculated the values of the first and the last representation levels. Namely, they considered the special case for L=1, of the hybrid model suggested in this paper. At the end of this paper we are considering the performances (relative distortion error) of the quantizers designed by using the hybrid model. We show that these performances are arbitrarily close to those of optimal scalar quantizers.

II. Hybrid model

Let us consider an N-level nonuniform scalar quantizer Q for the Laplacian input signals. Scalar quantizer Q is defined with Q: $R \rightarrow C$, as a functional mapping of the set of real numbers R onto the set of the output representation. The set of the output representation constitutes the code book:

$$C \equiv \{y_1, y_2, y_3, \cdots y_N\} \subset R \qquad (1)$$

that has the size $|C|=N$. The output values, y_j, are called the representation levels. The nonuniform scalar quantizer Q is defined with the set of the output values and with the partition of the input range of values onto N cells i.e. intervals α_j, j=1,2,...,N. Cells α_j are defined with the decision thresholds $\{t_0, t_1, \ldots, t_N\}$, such that $\alpha_j=(t_{j-1}, t_j]$, j=1,2,...,N. A quantized signal has value y_j when the original signal belongs to the quantization cell α_j. Hence, N-level scalar quantizer is defined as a functional mapping of an input value x onto an output representation, such as:

$$Q(x) = y_j, \quad x \in \alpha_j . \qquad (2)$$

[1]Zoran H.Perić is with the Faculty of Electronic Engineering, University of Nis, Aleksandra Medvedeva 14, 18000 Nis, Serbia
E-mail: peric@elfak.ni.ac.yu
[2]Jelena R. Nikolić is with the Faculty of Electronic Engineering, University of Nis, Aleksandra Medvedeva 14, 18000 Nis, Serbia
E-mail: njelena@elfak.ni.ac.yu

The negative thresholds and the representation levels are symmetric to their nonnegative counterparts. Hence, the considered quantizer can be depicted by using positive values of the decision thresholds $0=t_{N/2}<t_{N/2+1}<\ldots<t_{N-1}<t_N=\infty$ and representation levels $y_{N/2+1}<y_{N/2+2}<\ldots<y_N$. Let us denote the distances from the representative levels to the nether decision thresholds, i.e. reconstruction offsets δ_j $j=1,\ldots,N$, that are necessary when calculating the parameters of the scalar quantizers.

Nonuniform quantization can be achieved by the following procedure: compress the signal x using a nonlinear compressor characteristic $c(\cdot)$, quantize the compressed signal $c(x)$ with a uniform quantizer, expand the quantized version of the compressed signal using a nonlinear transfer characteristic $c^{-1}(\cdot)$ inverse to that of the compressor. The corresponding structure of a nonuniform quantizer consisting of a compressor, a uniform quantizer, and expandor in cascade is called the compandor. Let us define the compressor function $c(x)$ similarly as in [6]:

$$c(t_j) = -1 + 2\frac{\int_{-\infty}^{t_j} p^{1/3}(x)dx}{\int_{-\infty}^{+\infty} p^{1/3}(x)dx}. \tag{3}$$

Also, the following equation is valid for the compressor function $c(x)$ [7]:

$$c(t_j) = -1 + \frac{2j}{N}. \tag{4}$$

When the values of the input signal x are within the $(-\infty,\infty)$ range, the values of $c(t_j)$ are copied into the $[-1,1]$ range by using thus defined compressor function. Decision thresholds t_j, $j=1,2,\ldots,N-1$ can be determined by equating the last two equations:

$$t_j = \frac{3}{\sqrt{2}}\ln\left(\frac{N}{2(N-j)}\right) \qquad j=1,\ldots,N-1. \tag{5}$$

In order to simplify the equations that are necessary for the scalar quantizer's parameters calculation we can introduce the relation $K=N/2$. Thresholds t_{2K-L}, $L \ll K$, can be determined by using the following expression:

$$t_{2K-L} = \frac{3}{\sqrt{2}}\ln\left(\frac{K}{L}\right). \tag{6}$$

In this paper we are considering the hybrid model based on the combination of two quantization techniques. One of them is the companding technique and the other one is the Lloyd-Max's procedure for designing scalar quantizers. Namely, applying the companding technique to the range $(-t_{2K-L}, t_{2K-L})$ (inner region), i.e. to N-$2L$ inner cells $\alpha_{L+1},\ldots,\alpha_{2K-L}$, and Lloyd-Max's procedure to union of ranges $(t_0, -t_{2K-L})$ and (t_{2K-L}, t_{2K}) (outer region), i.e. $2L$ outer cells $\alpha_1,\ldots,\alpha_L$ and $\alpha_{2K-L+1},\ldots,\alpha_{2K}$ it is possible to design the N-levels scalar quantizer. The widths of the outer cells $\alpha_1,\ldots,\alpha_L$ and $\alpha_{2K-L+1},\ldots,\alpha_{2K}$ are constant and independent of the number of quantization levels N. Considering that fact the performances of the designed quantizer will be better than those of the quantizer realized by using the companding technique. Therefore, by using the compressor function $c(x)$, defined with Eqs. (3) and (4), we can use Eq. (6) to calculate the edge of the inner region t_{2K-L}. In order to calculate the decision thresholds and the representation levels of the outer region we use the well known values of the reconstruction offsets $\delta_{2K-L+1},\ldots,\delta_{2K}$ that are calculated in case of optimal Lloyd-Max's scalar quantizers [7]:

$$y_{2K-i+1} = t_{2K-i} + \delta_{2K-i+1}, \qquad i=1,\ldots,L \tag{7}$$

$$t_{2K-i+1} = y_{2K-i+1} + \delta_{2K-i+2}, \qquad i=2,\ldots,L \tag{8}$$

The values of the decision thresholds and representation levels of scalar quantizers realized by using the companding technique are not optimal. The goal of the suggested hybrid model of quantizers is to make, as much as possible, the decision thresholds and the representation levels to be optimal. Also, when designing N-level Lloyd-Max's scalar quantizer it is necessary to know all the values of the decision thresholds and the representation levels. Hence, in such a case $4K$ values should be memorized. However, when designing N-level scalar quantizer, based on the hybrid model, it is required to know the edge of the inner region t_{2K-L} and the set of L values of the reconstruction offsets $\delta_{2K-L+1},\ldots,\delta_{2K}$, i.e. the set of $L+1$ values. Thus, sparing the memory space simpler solution of hardware can be achieved. This is particularly of interest when designing scalar quantizers with large number of quantization levels $N=2K$. Hence, the compromise between the design complexity and the distances from the optimal solution of the scalar quantizer designing problem should be obtained. Furthermore, this model is the generalized model which for $L=K$ presents the model of the Lloyd-Max's quantizer while in case of $L=0$ presents the model of the quantizer realized by using the companding technique. Choosing the values of L it is possible to arbitrarily approach the optimal solution of the scalar quantizer construction problem.

III. THE QUANTIZER PERFORMANCES

The performance of a quantizer is often specified in terms of $SNRQ$ (signal to quantization noise ratio), given by [8]:

$$SNRQ = 10\log_{10}\left(\frac{\sigma^2}{D}\right) \tag{9}$$

measured in decibels, with σ^2 denoting the variance of x. Here we assume the unit variance input signal, therefore $SNRQ$ can be given by:

$$SNRQ = 10\log_{10}\left(\frac{1}{D}\right). \tag{10}$$

Let us define the relative distortion error δ such as:

$$\delta = \frac{D - D^{opt}}{D^{opt}} \tag{11}$$

where D^{opt} is the optimal distortion value. Also, let us denote the optimal value of $SNRQ$ with $SNRQ^{opt}$. Introducing the relation:

$$\Delta SNRQ = SNRQ - SNRQ^{opt} \tag{12}$$

the Eq. (11) becomes:

$$\delta = 10^{-\frac{\Delta SNRQ}{10}} - 1 . \tag{13}$$

In analyzing the behavior of the quantizer, it is preferable to use relative quantities, like signal to quantization noise ratio and relative distortion error instead of absolute quantities, such as distortion. Relative parameters portray the behavior of the quantizer in a way that is independent of the signal level and hence is more general. Good distortion approximation of quantizers based on the companding technique can be achieved by using Bennett's integral [9], [10] ranging $[-t_{2K-L},+t_{2K-L}]$:

$$D = \frac{1}{12(2K-2L)^2}\left(\int_{-t_{2K-L}}^{t_{2K-L}} p^{1/3}(x)dx\right)^3 + 2\sum_{j=2K-L}^{2K-1}\int_{t_j}^{t_{j+1}}(x-y_{j+1})^2 p(x)dx. \tag{14}$$

III. Numerical results

Table I provides numerical values of the the relative distortion error δ, calculated for L=0,1,2,4, when the number of quantization levels varies (N=32, 64, 128). Namely, one confirmation of the hybrid model validity is given by Table I. It is apparent that for L=8, for large enough number of quantization levels N, numerical values of the relative distortion error δ are below 0.005 [11], whereby the one of stopping criterion that allows interruption of the Lloyd-Max's algorithm is satisfied. Hence, it is obvious that for L=8, recommended the hybrid model enables optimal scalar quantizer designing. Summary of the numerical values for *SNRQ*, calculated for L=0,1,2,4,8, when the number of quantization levels are N=32, 64, 128, is given by Table II. Assimilating the appropriate values from Table II and Table III, one can notice that when the value of L grows it is approximately possible to approach to optimal values of the *SNRQ*.

Table I
Numerical values of the relative distortion error δ, calculated for L=0,1,2,4, 8, when the number of quantization levels varies (N=32, 64, 128).

δ	N=32	N=64	N=128
L=0	0.0713	0.0348	0.0181
L=1	0.0341	0.0168	0.0093
L=2	0.0213	0.0106	0.0062
L=4	0.0110	0.0057	0.0038
L=8	0.0041	0.0023	0.0021

Table II
Numerical values of the *SNRQ*, calculated for L=0,1,2,4, 8, when the number of quantization levels varies (N=32, 64, 128).

SNRQ	N=32	N=64	N=128
L=0	23.5709	29.5915	35.6121
L=1	23.7244	29.6675	35.6499
L=2	23.7785	29.6941	35.6631
L=4	23.8226	29.7155	35.6736
L=8	23.8522	29.7299	35.6807

Table III
Optimal reference of the *SNRQ*, ($SNRQ^{OPT}$), when the number of quantization levels varies (N=32, 64, 128).

	N=32	N=64	N=128
$SNRQ^{opt}$	23.87	29.74	35.69

III. Conclusion

The suggested hybrid model for scalar quantizer construction enables sophisticated relation, i.e. compromise between design complexity and the distances from the optimal solution of the scalar quantizer designing problem. The depicted model is based on the quantization technique that is very simple and convenient for use. It is very important to point out that for fixed L, by using the suggested hybrid model, when the number of quantization levels N varies, the size of the necessary calculation of scalar quantizers' parameters values is constant. Furthermore, the required memory space remains constant. However, the size of the necessary calculation and the capacity of memory space for Lloyd-Max's quantizers grow with the number of quantization levels N. In such a way, comparing Lloyd-Max's quantizers considerable contribution can be achieved by using the proposed hybrid model. Also, by using the suggested hybrid model for designing scalar quantizers with large enough number of quantization levels N, it is possible to achieve nearly optimal values of *SNRQ*, i.e. the order of difference is 10^{-2}. This is yet another proof of the hybrid model validity. Accordingly, the analyses shown here is of a practical importance because it can be of great help to engineers.

References

[1] Sangsin Na and David L. Neuhoff, "On the Support of MSE-Optimal, Fixed-Rate, Scalar Quantizers", *IEEE Transactions on Information Theory*, Vol. 47, No. 7, November 2001., pp. 2972-2982.

[2] S. P. Lloyd, "Least squares quantization in PCM", unpublished memo., Bell Lab.,1957; *IEEE Transactions on Information Theory*, Vol. IT-28, Mar. 1982, pp. 129-137.

[3] J. Max, "Quantizing for minimum distortion", *IRE, Transactions on Information Theory*, Vol. IT-6, Mar. 1960., pp. 7-12.

[4] Sangsin Na, "On the Support of Fixed-Rate Minimum Mean-Squared Error Scalar Quantizers for a Laplacian Source", *IEEE Transactions on Information Theory*, Vol. 50, No. 5, May 2004., pp. 937-944.

[5] Z. Perić, J. Nikolić, D. Pokrajac, "Analysis of support region for Laplacian source's scalar quantizers", *Proceedings of 7th International Conference on Telecommunications in Modern Satelite, Cable and Broadcasting Services TELSIKS 2005*, Vol. 2, pp. 491-494, Niš, Serbia and Montenegro, 2005.

[6] Neil Judell and Louis Scharf, "A Simple Derivation of Lloyd's Classical Result for the Optimum Scalar Quantizer", *IEEE Transactions on Information Theory,* Vol. 32, No. 2, March 1986., pp. 326-328.

[7] N.S. Jayant, Peter Noll, *Digital coding of waveforms*, Prentice-Hall, New Jersey, 1984, Chapter 4, pp. 129-139.

[8] Wai C. Chu, *Speech coding algorithms*, John Wiley & Sons, New Jersey, 2003, Chapter 6, pp. 161-165.

[9] Sangsin Na and David L. Neuhoff, "Bennett's Integral for Vector Quantizers", *IEEE Transaction on Information Theory*, Vol.41, July 1995., pp. 886-900.

[10] J. A. Bucklew and G. L. Wise, "Multidimensional asymptotic quantization theory", *IEEE Transactions on Information Theory*, Vol. IT-28, March 1982., pp. 239-247.

[11] Robert Gray, Quantization and data compression, Lecture notes, Stanford University, 2004.

Optimal Product Pyramid Vector Quantization of Memoryless Laplacian Source

Zoran H. Perić[1], Aleksandra Ž. Jovanović[2] and Olivera D. Milanović[3]

Abstract – **In this paper the optimal product pyramid Z^{n-1} lattice vector quantization of memoryless Laplacian source will be considered. The asymptotic analysis based on minimum distortion criterion will be performed. We will show that obtained optimal distortion is function of scalar case Bennet's integral and on this base we will find optimal multidimensional radial companding characteristic. In that way, applying the quantization technique that has small implementation complexity, we will achieve the signal to quantization noise ratio which in some cases differs from that of optimal vector quantization for about 0.5 dB.**

Keywords **– Optimal product vector quantization, Z^{n-1} lattice based quantization, optimal multidimensional radial companying-ding characteristic**

I. INTRODUCTION

In order to achieve the conversion of analog signals into efficient digital representation and the compression of digital information into the fewest possible bits, the vector quantization was developed. Vector quantization is an extension of the simple scalar quantization to multidimensional spaces and it can yield smaller average mean squared error per dimension than scalar quantization for the case of fine quantization [1], [2]. Also, the rapid advance in digital signal processor chips made possible low cost implementation of complex vector coding techniques.

The quantizer presented in this paper is designed for a memoryless Laplacian source. One reason for studying the memoryless Laplacian source is that it naturally arises in numerous applications. For example, the first approximation to the long-time-averaged probability density function (pdf) of speech amplitudes is provided by Laplacian model [1]. Also, in a number of papers the vector quantization of memoryless Laplacian source was analyzed since the probability density function of the difference signal for an image waveform follows the Laplacian function [3].

In this paper we consider the optimal product vector quantization which in comparisson with the unrestricted optimal vector quantization has a little worse performances but the lower implementation complexity. One quantization technique that can be applied for quantizer design is suggested in [4]. In that paper the importance of source geometry and lattice quantization was noted. Author considered the weighted pyramid vector quantizers for Laplacian sources and used one established approximate heuristic distortion formula for optimal weighted pyramid vector quantization. Here we derive the exact equation for the optimal total distortion.

Multidimensional companding on optimal high rate quantization was introduced by Gersho [5]. He pointed at the difficulties of doing optimal quantization with companding vector quantization. Bucklew also showed that asymptotically optimal unrestricted vector quantizers for vector dimensions 3 and greater can not be implemented using a companding structure, except for a very restricted class of source densities [6], [7]. A significant contribution to the understanding of companding was made in [8]. In [8] and [9] one analytic solution for radial companding characteristic was suggested. The proposed solution is independent of source type. However, author of [10] showed that in the case of Gaussian source quantization, the signal to quantization noise ratio obtained applying the companding characteristic obtained in [8] and [9] is for about 2.5 dB smaller than that of optimal vector quantization and he gave the better solution. In [10] the proposed heuristic solution gives performance for about 1 dB worse than that of optimal vector quantization in the case of Gaussian source quantization.

In order to respect source geometry we observe the pyramid vector quantization. Also, since we observe the equal points number per hyperpyramid equally distributed, all these points would map into the same points on hyperpyramid that has radius 1. Because of that it is convenient to apply the Z^{n-1} lattice on unit pyramid surface and after that radial map the obtained n-1 dimensional cells into n dimensional cells. Then we perform the asymptotic analysis based on minimum distortion criterion. After some mathematical manipulation we express the normalized moment of n dimensional cell in function of the normalized moment of n-1 dimensional cell projection. Than, by optimizing the total distortion, we find the optimal hyperpyramid number and distortion. We show that obtained distortion is function of scalar case Bennet's integral and on this base we find the optimal multidimensional radial companding characteristic. In that way, applying the quantization technique that has smaller implementation complexity, we achieve the signal to quantization noise ratio which in some cases differs from that of optimal vector quantization for about 0.5 dB.

[1]Zoran H. Perić is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Nis, Serbia and Montenegro, peric@ elfak.ni.ac.yu

[2]Aleksandra Ž. Jovanović is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Nis, Serbia and Montenegro, sanda@elfak.ni.ac.yu

[3]Olivera D. Milanović is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Nis, Serbia and Montenegro, oliveram@elfak.ni.ac.yu

II. Asymptotic Analysis of Optimal Product Pyramid Vector Quantization

For n-dimensional vector $\mathbf{x} = [x_1 \quad x_2 \quad \cdots \quad x_n]^T$ consisting of independent and identically distributed (i.i.d.) Laplacian variables x_i with zero mean and unit variance, the joint probability density function of $\boldsymbol{x}$ is

$$f_{\mathbf{x}}(\mathbf{x}) = \prod_{i=1}^{n} f(x_i) = (2)^{-\frac{n}{2}} \cdot \exp\left(-\sqrt{2}\sum_{i=1}^{n}|x_i|\right). \tag{1}$$

The contour of constant probability density function (pdf) is given by

$$\sum_{i=1}^{n}|x_i| = -\frac{\sqrt{2}}{2}\ln\left[f_c 2^{\frac{n}{2}}\right] \equiv g_0,\ g_0 \geq 0, \tag{2}$$

where f_c is the value of pdf. This is an expression for the n-dimensional pyramid with radius g_0, where we define the radius as

$$g = \sum_{i=1}^{n}|x_i| = \|\mathbf{x}\|_n^1. \tag{3}$$

The radius g is also random variable and has a pdf given as

$$f_g(g) = \frac{2^{\frac{n}{2}}}{\Gamma(n)} g^{n-1} e^{-\sqrt{2}g},\quad g \geq 0. \tag{4}$$

Using the previous formulation for g, we can perform the coordinate transformation and express the vector $\boldsymbol{x}$ through the vector intensity, i.e. the vector amplitude g and the location vector $\mathbf{s} = [s_1 \quad s_2 \quad \cdots \quad s_n]^T$

$$\mathbf{x} = g\mathbf{s}. \tag{5}$$

The n dimensional vector $\boldsymbol{s}$ is projection of vector $\boldsymbol{x}$ on hyperpyramid that has radius 1, i.e. one endpoint of vector $\boldsymbol{s}$ is placed on unit hyperpyramid.

In order to respect source geometry, we consider that a nonlinear vector quantization of Laplacian source has L representative amplitude levels determined with equation $\hat{g}_i = const.,\quad i = 1,...,L$. In accordance with this assumption, the decision amplitude levels are g_i, $i = 0,\ldots, L$. Also, since we analyze the product vector quantization, we have the equal points number per amplitude levels M equally distributed. Therefore, we have that

$$LM = N, \tag{6}$$

where N is total points number of quantizer. If we take into consideration that number of amplitude levels L is very large, the required memory in the case of product nonuniform quantization is far less than that in the case of unrestricted nonuniform quantization when we should store the points number on each amplitude levels M_i, $i = 1,\ldots,L$.

In the case of product quantization the points from different amplitude levels map into same points on hyperpyramid that has radius 1. Because of that we can apply the Z^{n-1} lattice on unit pyramid surface and after that radial map the obtained n-1 dimensional cells into n dimensional cells with representatives equally distributed on observed hyperpyramides (Fig 1.). So, if we sign the quantization cell as $c_{i,j}$, where i determines amplitude level, i=1,…, L, j the cell on given representative amplitude level, j=1,…, M, we can write that representative of cell $c_{i,j}$ is

$$\hat{\mathbf{x}}_{i,j} = \hat{g}_i\hat{\mathbf{s}}_j,\quad j = 1,...,M,\ i = 1,...,L. \tag{7}$$

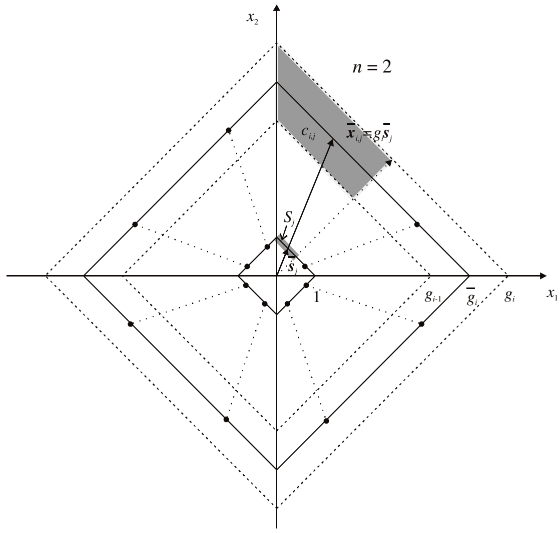


Fig. 1. Illustration of pyramid Z^{n-1} lattice vector quantization.

Now, we should define the normalized second order moment, i.e. the inertia moment of cell $c_{i,j}$

$$M^n = \frac{1}{n}\frac{1}{\mathrm{vol}(c_{i,j})^{1+\frac{2}{n}}}\int_{c_{i,j}} d(\mathbf{x}, \hat{\mathbf{x}}_{i,j})d\mathbf{x}, \tag{8}$$

where vol($c_{i,j}$) is the volume of the cell $c_{i,j}$

$$\mathrm{vol}(c_{i,j}) = \frac{2^n}{M\Gamma(n+1)}\left(g_i^n - g_{i-1}^n\right) \tag{9}$$

and $$d(\mathbf{x}, \hat{\mathbf{x}}_{i,j}) = \left(\|\mathbf{x}\|_n^1\right)^2 + \left(\|\hat{\mathbf{x}}_{i,j}\|_n^1\right)^2 - 2\|\mathbf{x}\|_n^1\|\hat{\mathbf{x}}_{i,j}\|_n^1 \cos\varphi. \tag{10}$$

If we take into consideration that corresponding unit vectors to vectors $\boldsymbol{x}$ and $\hat{\mathbf{x}}_{i,j}$ are $\boldsymbol{s}$ and $\hat{\mathbf{s}}_j$, respectively, the Eq. (10) also applies to them

$$d(\mathbf{s}, \hat{\mathbf{s}}_j) = 2 - 2\cos\varphi. \tag{11}$$

Beside this, for great number of amplitude levels L, we can assume that $g \approx \hat{g}_i$ inside one cell, i.e. we can write

$$d(\mathbf{x}, \hat{\mathbf{x}}_{i,j}) = (g - \hat{g}_i)^2 + \hat{g}_i^2 d(\mathbf{s}, \hat{\mathbf{s}}_j). \tag{12}$$

We also know that for great number of amplitude levels L, the volume of cell $c_{i,j}$ can be approximated as

$$\mathrm{vol}(c_{i,j}) \approx \frac{2^n}{M\Gamma(n)}\hat{g}_i^{n-1}\Delta g_i = S(\hat{g}_i)\Delta g_i, \tag{13}$$

where $S(\hat{g}_i)$ is effective surface of one cell from hyperpyramid that has radius $\hat{g}_i$ and $\Delta g_i = g_i - g_{i-1}$.

Similar, we can define the normalized second order moment of $c_{i,j}$ cell projection on unit hyperpyramid S_j

$$G(\Lambda)=\frac{1}{n-1}\frac{1}{\mathrm{vol}(S_j)^{1+\frac{2}{n-1}}}\int_{S_j} d(\mathbf{s},\hat{\mathbf{s}}_j)ds\,, \tag{14}$$

where n-1 is dimension of cell S_j and

$$\mathrm{vol}(S_j)=S(\hat{g}_i=1)=\frac{2^n}{M\Gamma(n)}\,. \tag{15}$$

If we substitute Eq. (12) in Eq. (8) simultaneously using the relation $d\mathbf{x}=\hat{g}_i^{n-1}dgd\mathbf{s}$, after some mathematical manipulation we express the normalized moment of n dimensional cell in function of the normalized moment of n-1 dimensional cell projection

$$M^n=\frac{1}{n}\frac{1}{\mathrm{vol}(c_{i,j})^{1+\frac{2}{n}}}\cdot \cdot\left[\frac{\Delta g_i^2}{12}+(n-1)G(\Lambda)\left[S(\hat{g}_i=1)\right]^{\frac{2}{n-1}}\hat{g}_i^2\right]. \tag{16}$$

Now, we can find distortion per dimension of the jth cell on the ith representative amplitude level ($c_{i,j}$)

$$D_{i,j}=\frac{1}{n}\int_{c_{i,j}} d(\mathbf{x},\hat{\mathbf{x}}_{i,j})f_{\mathbf{x}}(\mathbf{x})d\mathbf{x},\ j=1,2,...,M_i,\ i=1,2,...,L\,. \tag{17}$$

We perform asymptotic analysis, i.e. we assume that pdf of input vector $f_{\boldsymbol{x}}(\boldsymbol{x})$ is constant inside cell and has value $f_{\mathbf{x}}(\hat{\mathbf{x}}_{i,j})$. Under this assumption we can write that

$$D_{i,j}=M^n\mathrm{vol}(c_{i,j})^{1+\frac{2}{n}}f_{\mathbf{x}}(\hat{\mathbf{x}}_{i,j})=M^n\mathrm{vol}(c_{i,j})^{\frac{2}{n}}P_{i,j}\,, \tag{18}$$

where $P_{i,j}$ is the probability that the vector $\boldsymbol{x}$ belongs to cell $c_{i,j}$

$$P_{i,j}=\int_{c_{i,j}} f_{\mathbf{x}}(\mathbf{x})d\mathbf{x}=f(\hat{\mathbf{x}}_{i,j})\int_{c_{i,j}} d\mathbf{x}=f(\hat{\mathbf{x}}_{i,j})\mathrm{vol}(c_{i,j})\,. \tag{19}$$

After that the granular distortion of cells that have representatives on the ith amplitude representative level is

$$D_i=\sum_{j=1}^{M}D_{i,j}=M^n\mathrm{vol}(c_{i,j})^{\frac{2}{n}}P_i\,, \tag{20}$$

where P_i is a probability that vector $\boldsymbol{x}$ is located between hyperpyramides that have radii g_i and g_{i-1}

$$P_i=\sum_{j=1}^{M}P_{i,j}=\int_{g_{i-1}}^{g_i} f_g(g)dg\approx f_g(\hat{g}_i)\Delta g_i\,. \tag{21}$$

On the other hand, we also observe the nonuniform quantization through the companding vector quantization. Because of that, we define the multidimensional radial companding characteristic $h(g)$. This function maps compressor input range [0,+∞) into output range [0,1). Since we consider the great number of amplitude levels L, we use following relation in our analysis

$$\left.\frac{dh(g)}{dg}\right|_{g=\hat{g}_i}=\frac{\frac{1}{L}}{\Delta g_i}\,, \tag{22}$$

i.e.

$$\Delta g_i=\frac{1}{L\left.\frac{dh(g)}{dg}\right|_{g=\hat{g}_i}}\,. \tag{23}$$

Substituting the Eqs. (16), (23) and (21) into Eq. (20), we get that distortion per dimension for whole product nonuniform vector quantizer of signal generated by Laplacian source can be written as follows

$$D=\sum_{i=1}^{L}D_i=\frac{1}{L^2 12n}\sum_{i=1}^{L}\frac{f_g(\hat{g}_i)}{\left(\left.\frac{dh(g)}{dg}\right|_{g=\hat{g}_i}\right)^2}\Delta g_i+ \\ +\frac{(n-1)G(\Lambda)\left[S(\hat{g}_i=1)\right]^{\frac{2}{n-1}}}{n}\sum_{i=1}^{L}\hat{g}_i^2 f_g(\hat{g}_i)\Delta g_i \,. \tag{24}$$

If we apply the Riemann integral definition on the right hand of (24), we can replace sum with integral

$$D=\frac{1}{L^2 12n}\int_0^{+\infty}\frac{f_g(g)}{\left(\frac{dh(g)}{dg}\right)^2}dg+ \\ +\frac{(n-1)G(\Lambda)\left[S(\hat{g}_i=1)\right]^{\frac{2}{n-1}}}{n}\int_0^{+\infty}g^2 f_g(g)dg \,. \tag{25}$$

If we substitute Eq. (15) in Eq. (25) and take into consideration that $M=N/L$, we get the expression for total distortion per dimension of product pyramid Z^{n-1} lattice vector quantization D in function of hyperpyramides number L

$$D=\frac{l_0}{L^2}+\frac{tL^{\frac{2}{n-1}}}{N^{\frac{2}{n-1}}}\,, \tag{26}$$

$$l_0=\frac{1}{12n}\int_0^{+\infty}\frac{f_g(g)}{\left(\frac{dh(g)}{dg}\right)^2}dg\,, \tag{27}$$

$$t=\frac{2^{\frac{n+1}{n-1}}(n-1)(n+1)G(\Lambda)}{\left[\Gamma(n)\right]^{\frac{2}{n-1}}}\int_0^{+\infty}g^2 f_g(g)dg\,. \tag{28}$$

Now we can perform the asymptotic analysis based on minimum distortion criterion. By optimizing the total distortion, i.e differentiating D with respect to L and equalizing with zero, we find the optimal hyperpyramid number and distortion

$$L_{opt}=(n-1)^{\frac{n-1}{2n}}l_0^{\frac{n-1}{2n}}t^{-\frac{n-1}{2n}}N^{\frac{1}{n}}\,, \tag{29}$$

$$D_{opt}=n(n-1)^{-\frac{n-1}{n}}l_0^{\frac{1}{n}}t^{\frac{n-1}{n}}N^{-\frac{2}{n}}\,, \tag{30}$$

respectively. Since n, t and N are given parameters, D_{opt} depends only of scalar variable l_0. If we carefully observe the expression for l_0, we see that Eq. (27) corresponds to Bennet's integral form in the case of scalar quantization. Since in this way we show that obtained distortion is only function of scalar case Bennet's integral, we can find optimal multidimensional radial companding characteristic. Namely, we conclude that multidimensional radial companding characteristic has the same form as the optimal scalar companding characteristic

$$h(g)=\frac{\int_0^g x^{\frac{n-1}{3}}\exp\left(-\frac{\sqrt{2}x}{3}\right)dx}{\left(\frac{3}{\sqrt{2}}\right)^{\frac{n+2}{3}}\Gamma\left(\frac{n+2}{3}\right)}\,. \tag{31}$$

In [11] the optimal piecewise uniform vector quantization of memoryless Laplacian source was analyzed. Applying this analysis, the optimal multidimensional radial companding characteristic (31) can be also obtained, but the procedure would be more complicate.

After short mathematical manipulation, we obtain that optimal product Z^{n-1} lattice vector quantization distortion and optimal number of hyperpyramid are

$$D_{opt} = 2^{\frac{n-2}{n}} 3^{\frac{n+1}{n}} n^{\frac{n-1}{n}} (n+1)^{\frac{n-1}{n}} \left[\frac{\Gamma\left(\frac{n+2}{3}\right)}{\Gamma(n)} \right]^{\frac{3}{n}} G(\Lambda)^{\frac{n-1}{n}} N^{-\frac{2}{n}}, \quad (32)$$

$$L_{opt} = 2^{-\frac{2n-1}{n}} 3^{\frac{n^2-1}{2n}} n^{-\frac{n-1}{2n}} (n+1)^{-\frac{n-1}{2n}} \cdot \frac{\Gamma\left(\frac{n+2}{3}\right)^{\frac{3(n-1)}{2n}}}{\Gamma(n)^{\frac{n-3}{2n}}} G(\Lambda)^{-\frac{n-1}{2n}} N^{\frac{1}{n}}. \quad (33)$$

respectively.

In Table I we show the obtained results for signal to quantization noise ratio *SQNR*

$$SQNR = 10 \log \frac{1}{D} \quad (34)$$

in function of dimension n for bit rate

$$R = \frac{1}{n} \log_2 N = 8 \text{ bits/dimension}. \quad (35)$$

Simultaneous the experimental results are obtained ($SQNR^{e}$). Since the confidence of simulation is 95%, the performed simulation run of 1000 vectors shows good matching with theoretical results.

Table also contains the values of signal to quantization noise ratio in the case of optimal unrestricted nonuniform vector quantization

$$SQNR^{\mathrm{OVQ}} = 10 \log \frac{1}{\frac{1}{6}\left(\frac{n+2}{n}\right)^{n+2} N^{-\frac{2}{n}}}, \quad (36)$$

and difference $\Delta SQNR = SQNR^{\mathrm{OVQ}} - SQNR$. It is obvious from Table I that difference between signal to quantization noise ratio of optimal and optimal product vector quantization decreases with the increase of the dimension n and for $n = 32$, $\Delta SQNR = 0.55$ dB. Furthermore, for large dimension ($n = 125$), optimal product pyramid vector quantization performances converge to those of optimal vector quantization ($\Delta SQNR = 0.19$ dB). On the other hand, comparing Eq. (32) with Eq. (36), we note that $\Delta SQNR$ is independent of N, i.e. bit rate R.

TABLE I
SIGNAL TO QUANTIZATION NOISE RATIO IN FUNCTION OF DIMENSION

N	16	24	32	125
$SQNR$ [dB]	45.82	46.22	46.44	47.00
$SQNR^{e}$ [dB]	46.54	46.57	46.82	47.11
$SQNR^{\mathrm{OVQ}}$ [dB]	46.74	46.91	46.99	47.19
$\Delta SQNR$ [dB]	0.92	0.69	0.55	0.19
G [dB]	4.12	4.52	4.74	5.3

If we take into consideration that signal to noise ratio of optimal scalar vector quantization is 41.7 dB for 8 bits per sample, we see that obtained gain with our quantizer $G = SQNR - 41.7$ dB is large (the last row of Table I).

III. CONCLUSION

In this paper we perform an exact and complete asymptotic analysis of optimal product pyramid Z^{n-1} lattice vector quantization of memoryless Laplacian source. We derive the expressions for the optimal hyperpyramid number, the optimal total distortion and the optimal multidimensional companding characteristic. The presented analysis is very simple and convenient for implementation. Opposite to classical product pyramid lattice vector quantization the hyperpyramides during this quantization are not equidistant which enables us determination optimal multidimensional radial companding characteristic.

Results show that suggested quantization model gives the signal to quantization noise ratio which in some cases differs from that of optimal vector quantization for about 0.5 dB. It proves that optimal product pyramid Z^{n-1} lattice vector quantization in comparisson with the unrestricted optimal vector quantization has a little worse performances but the lower implementation complexity. In order to simplify implementation of considered vector quantization, the model should be linearized applying the piecewise uniform product vector quantizer.

REFERENCES

[1] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Boston, Dordrecht, London: Kluwer Academ.Pub, 1992.

[2] R. M. Gray, D. L. Neuhoff, "Quantization", *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2325-2384, Oct. 1998.

[3] N.S.Jayant and P.Noll, *DIGITAL CODING OF WAVEFORMS Principles and Applications to Speech and Video*, Prentice-Hall, New Jersey, 1984.

[4] T. R. Fischer, “Geometric Source Coding and Vector Quantization”, *IEEE Trans. Inform. Theory*, vol. 35, no. 1, pp. 137-144, Jan. 1989.

[5] A. Gersho, “Asymptotically Optimal Block Quantization”, *IEEE Trans. Inform. Theory*, vol. 25, no. 4, pp. 373-380, July 1979.

[6] J. Bucklew, “Companding and random quantization in several dimensions”, *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 207-211, Mar. 1981.

[7] J. Bucklew, “A note on optimal multidimensional companders”, *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 279, Mar. 1983.

[8] P. Moo, “Asymptotic analysis of lattice-based quantization”, Ph. D. dissertation, Univ. Michigan, Ann Arbor, 1998.

[9] P. Moo and D. Neuhoff, “Optimal compressor functions for multidimensional companding”, in *Proc. IEEE Int. Symp. Information Theory,* Ulm, Germany, 1997, p. 515.

[10] J. Samuelsson, ”Multidimensional Companding Quantization of the Gaussian Source”, *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1343-1351, May 2003.

[11] Z. Perić, V. Stanković, A. Jovanović, S. Bogosavljević, “Optimal Piecewise Uniform Vector Quantization of the Memoryless Laplacian source”, *Journal of Electrical Engineering*, Vol. 56, No. 7-8, pp. 200-204, 2005.

Adaptive code book with neural network for CELP speech signal coding

Sn. Pleshkova- Bekjarska[1]

Abstract - **The adaptive code book in CELP speech signal coding method is one of the most intensive calculation parts of the coding algorithm. It is very important to prove the possibility to represent this code book with a suitable neural network and make an analysis of the performance and time of calculations comparison with ordinary adaptive code book. This is the goal of this article. There are present the results of a Matlab simulation with the proposed neural network embedded in a CELP coding algorithm.**

Keywords - **CELP speech coding, adaptive code book, neural networks.**

I. Introduction

The adaptive code book in CELP coder represent the long term prediction filter (LTP) [1], which is described as:

$$d(n) = x(n) - b.d(n-T), \qquad (1)$$

where

$d(n)$ is the difference or error signal;

$x(n)$ - the input signal;

$d(n-T)$ - the signal of time delay T equal of pitch period;

b – coefficient of long term prediction filter (LTP).

In the Federal Standard FS1016 [2] it is proposed to change the open loop long term prediction filter with a closed loop method – adaptive code book. The principle of adaptive code book is shown in Fig. 1.

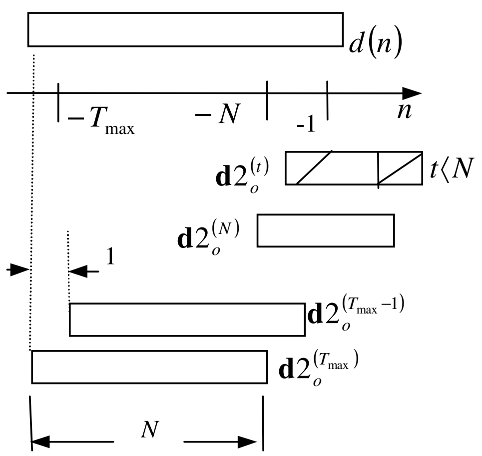


Fig. 1. Principle of adaptive code book

The content of the book is shown as a one dimensional vector $d(n)$ and parts of this vector are the current vectors for the range of pitch period $T = 20 \div 140$.

In the present article it is proposed to represent the adaptive code book with a neural network.

II. The structure of the neural network as an adaptive code book

It can be investigated, that the most useful type of a neural network as an adaptive code book is the adaptive linear neural network is shown in Fig. 2.

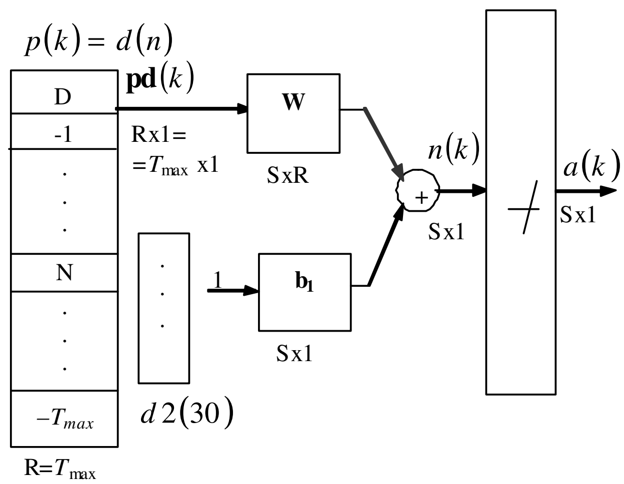


Fig. 2. The adaptive linear neural network as an adaptive code book

The equations, which describe this structure are given here:

$$p(k) = s(n), \qquad (2)$$

$$\text{pd}(k) = p(k), p(k-1), \ldots, p(k-R) = s(n), s(n-1), \ldots, s(n=N). \qquad (3)$$

$$p(k) = d(n); \qquad (4)$$

$$pd(k) = d(n), d(n-1), \ldots, d(n-N), \ldots, d(n-T_{\max}). \qquad (5)$$

They gives also the relationships between the symbols used in CELP adaptive code book and the symbols related with neural network terminology.

[1] Sn. Pleshkova is with the Faculty of Communications and Communication Technologies, Technical University, Kliment Ohridski 8, 1000 Sofia, Bulgaria.
E-mail: snegpl@tu-sofia.bg

III. Simulation program of neural network adaptive code book representation

The simulation program made for examination of the proposed neural network as an adaptive code book in a CELP coder is shown in Fig. 3.

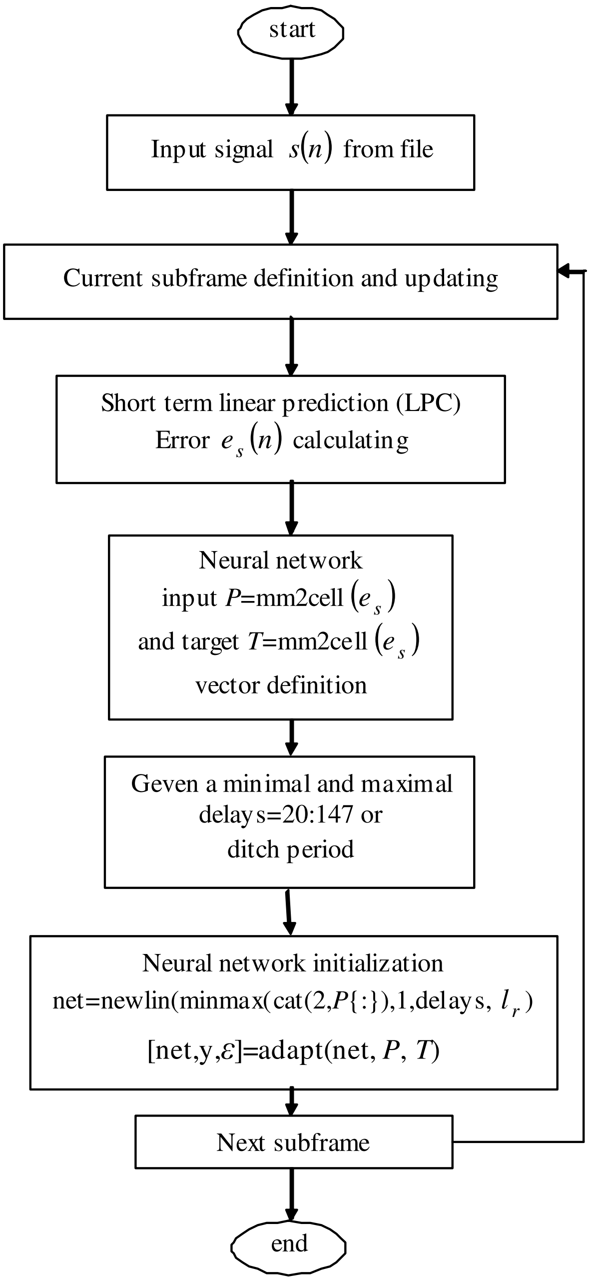


Fig. 3. Simulation program of adaptive neural network

There are involved in this Matlab simulation program all necessary actions for speech signal $s(n)$ input from file, current subframe defining and updating, short term linear predict-tion for error e_s calculating etc.

Neural network is initialized with input and target vectors:

$$P = \text{num2cell}(e_s); \quad T = \text{num2cell}(e_s). \tag{6}$$

$$\text{net} = \text{newlin}(\text{minmax}(\text{cat}(2,P\{:\})),1,\text{delays},\ l_r) \tag{7}$$

The minimal and maximal pitch period or delays are setting 20 and 147, respectively to give the neural network taped delay line initialization.

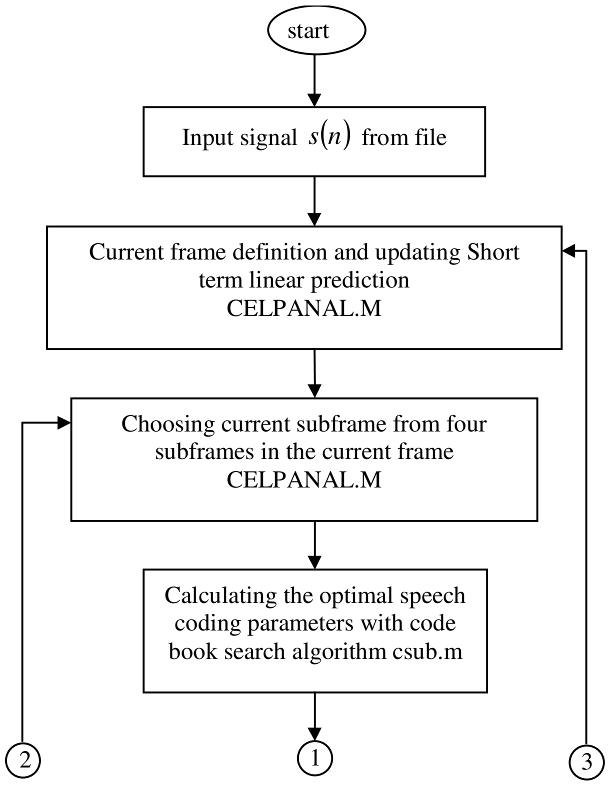
Finally it is made a neural network training:

$$[\text{net},\text{y},\varepsilon]=\text{adapt}(\text{net},\ P,\ T) \tag{8}$$

IV. Neural network embedded as an adaptive code book in standard FS1016 CELP coder

The described Matlab program is made as an independent program. It gives the results, that it is possible to represent the adaptive code book with a neural network. But it is more realistic, to embedded this neural network in the Federal Standard CELP algorithm FS1016, and to make the comparison of speech signal CELP coding using both standard adaptive code book and embedded neural network.

The possibility to make the replacement of standard adaptive code book with neural network is shown in Fig. 4.



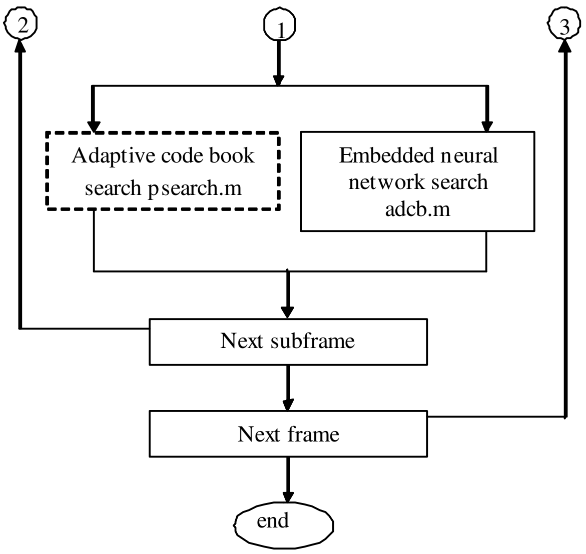


Fig. 4. The replacement of standard adaptive code book with neural network

V. RESULTS FROM THE NEURAL NETWORK SIMULATION

The results, which gives the simulation of neural network as an adaptive code book in an independent Matlab program and as an embedded function in the Federal Standard FS1016 CELP coding algorithm are shown in the next ten figures. The time diagram of input speech signal $s(n)$ is shown in Fig. 5.

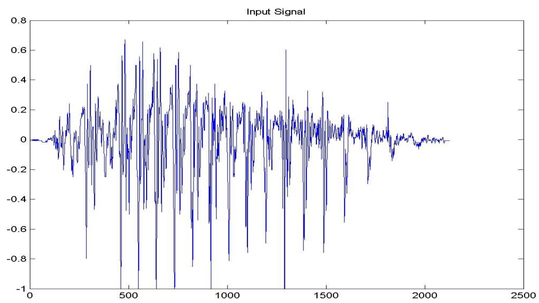


Fig. 5. Time diagram of input speech signal.

The result of short term linear prediction (LPC) error (e_s) estimation in the current frame in given in Fig. 6.

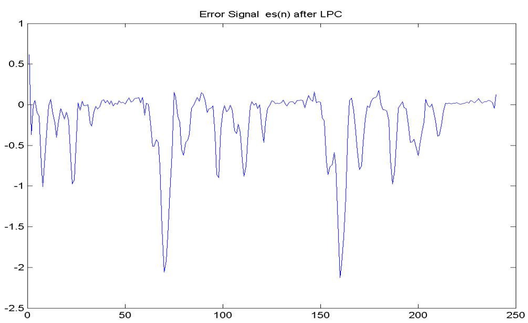


Fig. 6. The error signal of LPC.

The presences of the residual pitch period correlation in the error e_s is shown in Fig. 7.

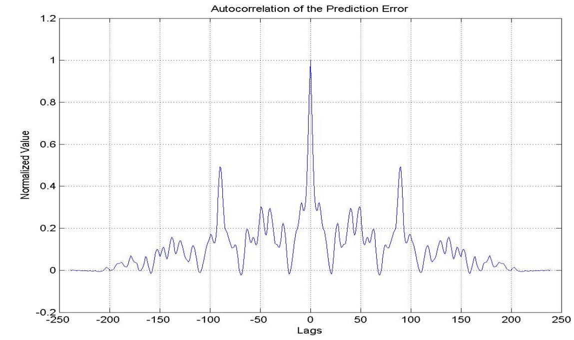


Fig. 7. The residual pitch period correlation in the error signal.

The amplitude and phase spectrum of current frame is given in Fig. 8.

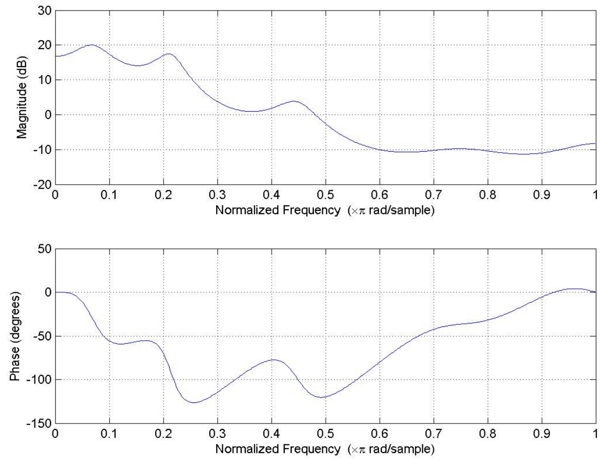


Fig. 8. The amplitude and phase spectrum of current frame

The guarantee of presences of pitch period in the error signal e_s is the autocorrelation function of this signal – Fig. 9.

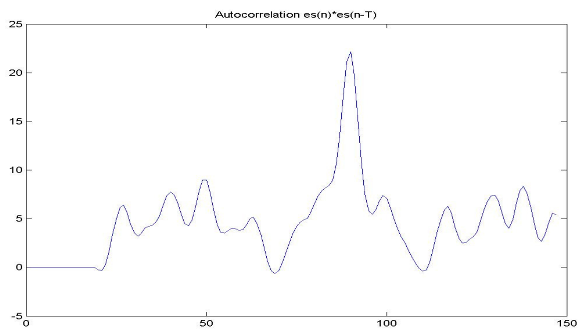


Fig. 9. The pitch period in the error signal.

The error signal after neural network simulation and training show the place of a strong minimum, which can be used to calculate pitch period T with an appropriate precision – Fig. 10.

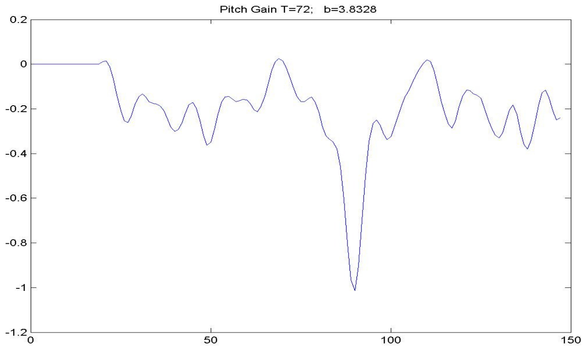


Fig. 10. The error signal after neural network simulation for long term filter pitch period *T* calculation.

The same is shown in Fig. 11 for long term filter coefficient *b* calculation.

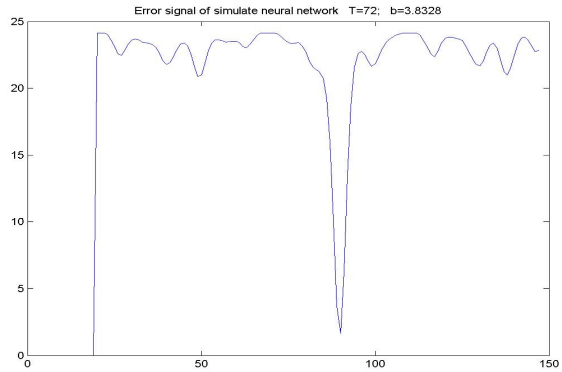


Fig. 11. The error signal after neural network simulation for long term filter coefficient *b* calculation.

An overall picture of the total algorithm of CELP coding and decoding of a speech signal using an embedded neural network an adaptive code cook in the Federal Standard FS1016 is shown in the Fig. 12 and Fig. 13. The fig. 12 represent a comparison in time domain of input and decoded speech signal. The same comparison in spectral domain is given in the Fig. 13.

VI. CONCLUSION

The theoretical analysis and the presented practical Matlab simulations of proposed neural network shows, that in is really possible to represent the adaptive code book in CELP coding method as an adaptive linear neural network. The time and spectral comparisons gives the assurance to embedding of the proposed neural network as a part of Federal Standard FS1016 CELP algorithm.

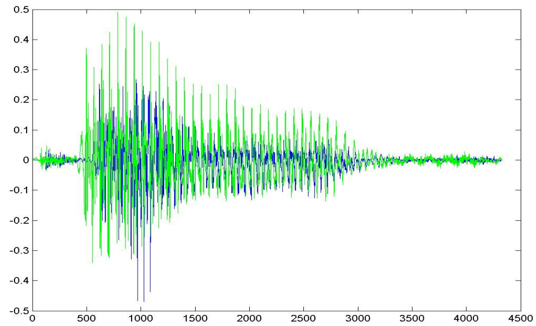


Fig. 12. Comparison in time domain of input and decoded speech signal.

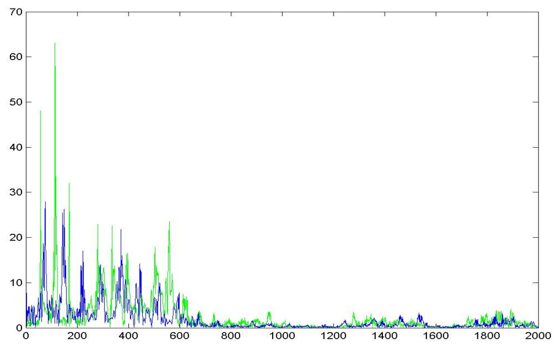


Fig. 13. Comparison in time domain of input and decoded speech signal.

REFERENCE

[1]. Atal B.S., Schroeder M.R. "Stochastic coding of Speech at Very Low Bit Rates", Proceedings of ICC, pp. 1610-1630,1984.

[2]. Federal Standard 1016, "Telecommunications: Analog to digital conversion of radio voice by 4800 bit/second code excited linear prediction (CELP)", National Communication System-Office of Technology and Standards, Feb. 1991.

Theoretical Analysis as a Function of the Total Q Factor Loudspeaker Characteristics

Ekaterinoslav S. Sirakov [1]

Abstract – **This work considers researching of the following characteristics of a directly radiating electrodynamic loudspeaker as a function of the total Q factor: frequency response of the amplitude and phase of the sound pressure created, of the group time delay, of the module and the complex impedance.**

For the purposes of the theoretical analysis of the above loudspeaker characteristics depending on the responses of the total Q factor, three regions are considered, where equations describing the output signal at input excitation - Heaviside function - are offered for each of them.

Keywords – **Loudspeaker, frequency response, step response.**

The transitional function of a loudspeaker and a closed-box loudspeaker system [1÷7] can be described by the following equation (1):

$$W(s,Q)=\frac{\dfrac{s^2}{\varpi_s^{\,2}}}{\dfrac{s^2}{\varpi_s^{\,2}}+\dfrac{s}{\omega_s}\cdot\dfrac{1}{Q}+1} \tag{1}$$

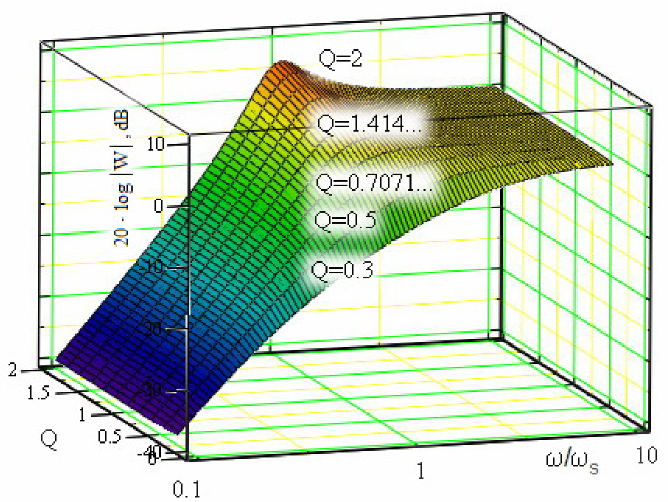


Fig. 1. Normalized Amplitude-Frequency Response of the Sound Pressure Created by the Loudspeaker (Q=0.3÷2)

This function is analogical to a second-order high-pass filter (40 dB/dec. cutoff).

Where: $s=\sigma+i.\varpi$ is the complex frequency variable,

f_s is the resonance frequency ($\varpi_s=2.\pi.f_s$) of the loudspeaker.

Fig. 1 shows the normalized characteristic of the sound pressure created by the loudspeaker $20.\log\left|W\left(\frac{i.\varpi}{\varpi_s},Q\right)\right|$ as a function of the frequency normalized to $\varpi_s=2.\pi.f_s$ - in a logarithmic scale - and the total Q factor (Q=0.3÷2).

These responses for Q greater than $1/\sqrt{2}$ i.e. 0.7071 are second-order Chebyshev equal-ripple alignments [1].

For $Q=0.7071$ (i.e. $1/\sqrt{2}$) is a second-order Butterworth maximally-flat alignment.

For $Q=0.5$ is a second-order Linkwits-Riley alignment

The responses for $Q=0.3\div0.7$ are second-order alignment for real world loudspeakers.

Table 1 shows the DYNAUDIO loudspeakers parameters [16].

Loud-speaker	fs [Hz]	Qts	1w/1m dB	W to fs dB
T-330 D	750	0.2	92	-10
D-21 AF	1300	0.41	91	-4
D-28/2	880	0.41	89	-5
D-260	1000	0.48	90	-4
M-560D	325	0.35	91	-5
D-52 AF	350	0.4	88	-6
D-54 AF	325	0.3	92	-7
D-76 AF	350	0.9	89	0
17 W-75XL	42	0.44	89	BOX
24 W-75	32	0.35	90	BOX
30 W-54	22	0.36	92	BOX

Table 1. DYNAUDIO Loudspeaker parameters [16]

Where:
fs [Hz] is the loudspeaker resonance frequency,
Qts – total Q factor,
1w/1m – sensitivity - i.e. the sound pressure created - in dB – along the axis of radiation at a distance normalized to 1m and 1W power supplied to the loudspeaker.
W to fs – loudspeaker sensitivity for the resonance frequentcy, fs, in dB.

The first four lines of Table 1 apply to high-frequency (tweeter) dome loudspeakers (T-330 D, D-21 AF, D-28/2 and D-260), and then follow the parameters of middle-frequency dome loudspeakers (M-560D, D-52 AF, D-54 AF AND D-54 AF). They are designed as closed-box systems and their Producer's brochures list the so called Thiele/Small parameters and characteristics [16].

The parameters of the low-frequency (woofer) loudspeakers (17 W-75XL, 24 W-75 and 30 W-54) determine the selection of the BOX: closed-box, variovented box, bass-reflex, trans-

[1]Ekaterinoslav S. Sirakov is with the Department of Radio engineering, Faculty of Electronics, Technical University-Varna, Studentska Street 1, Varna 9010, Bulgaria, E-mail: katio@mail.bg, katiosirakov@abv.bg

mission line, band pass, etc. The selected box and the actual loudspeaker, they both determine the parameters and the characteristics in the low-frequency region [1÷7].

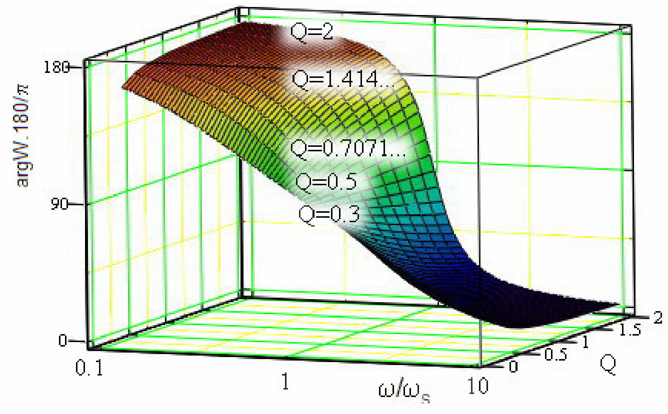


Fig. 2. Normalized frequency response of the phase of the sound pressure created by the loudspeaker where Q factor is the parameter.

The phase of the sound pressure created by the loudspeaker, as a function of the total Q factor, is plotted in Fig. 2.

$$\phi(s,Q) = \arg(W(s,Q)).\frac{180}{\pi} \tag{2}$$

The displacement of the loudspeaker voice-coil, as defined by formula (3) is analogical to a function of a low-pass second-order filter.

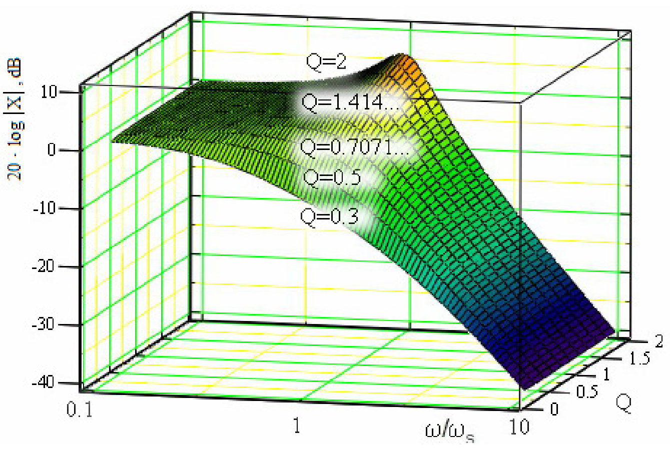


Fig. 3. Normalized voice-coil displacement of the loudspeaker as a function of normalized frequency response for parameter total Q factor.

Fig. 3 shows the normalized voice-coil displacement magnitude $20.\log\left|X\left(\frac{i.\varpi}{\varpi_s},Q\right)\right|$ with frequency normalized to ϖ_s as a function of the total Q factor.

$$X(s,Q) = \frac{1}{\frac{s^2}{\varpi_s^{\ 2}} + \frac{s}{\omega_s}.\frac{1}{Q} + 1} \tag{3}$$

The group time delay can be defined as:

$$G\left(\frac{i.\varpi}{\varpi_s},Q\right) = -\frac{d}{d\varpi}\left(\phi\left(\frac{i.\varpi}{\varpi_s},Q\right)\right) \tag{4}$$

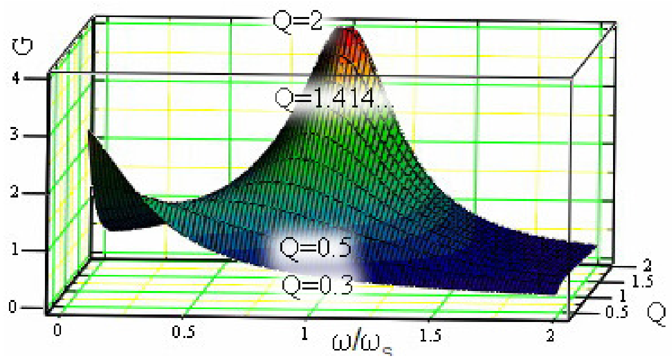


Fig. 4. Frequency response of the group time delay of a loudspeaker, where parameter is the Q factor.

The equation [1÷7], which defines the input impedance, includes the sum of the voice-coil electrical impedance $R_e + i.\omega.L_e$ and the inserted impedance of the loudspeaker $Z_{вн}$:

$$Z_{LS} = R_e + i.\omega.L_e + \frac{(B.L)^2}{r + \frac{1}{i.\omega.c} + i.\omega.m} \tag{5}$$

where: R_e is the voice-coil DC resistance,

L_e - voice-coil inductance,

$B.L$ - force factor magnet system,

r - mech. resistance,

m - moving mass,

c - suspension compliance.

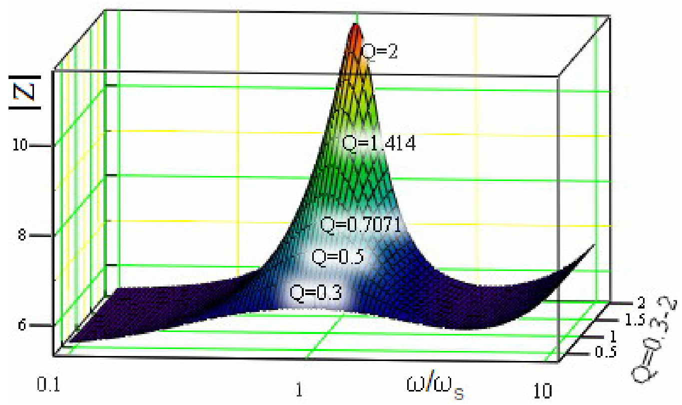


Fig. 5. Magnitude of the loudspeaker impedance depending on the normalized frequency as a function of the Q factor.

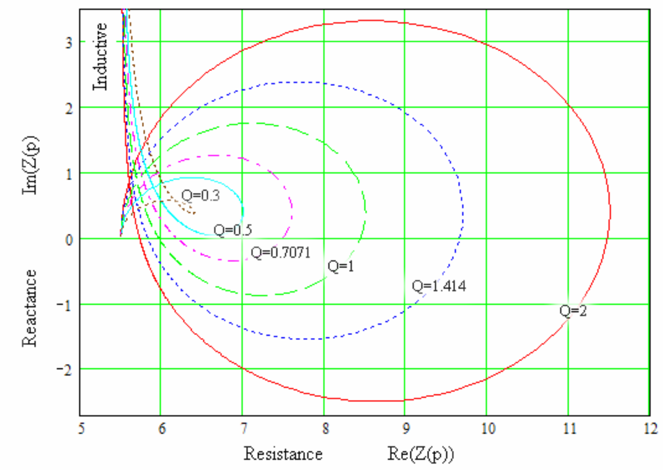


Fig. 6. Complex Impedance (Nyquist plots)

Loudspeaker Step Response

The input step function is:

$$\Phi(t) := if\,(t < 0,0,1) \tag{6}$$

The excitation is the Heaviside function for which the Laplace transform is $1/s$ [8].

$$\Phi(\mathrm{tn})\ \mathrm{laplace}, \mathrm{tn}\ \rightarrow \frac{1}{\mathrm{s}}$$

With the program MathCad symbolic transform [13,14], *invlaplace*, from the step function (1) we find an equation which describes the sound pressure created by the loudspeaker as a function of time with single input excitation and Q factor - a parameter:

$$\frac{1}{s}.W(s)invlaplace, s \rightarrow \tag{7}$$

The following three spaces can be defined depending on the values of the total Q factor:

1. Damped oscillation - at Q>0.5

$$Out1(t) = \left\{e^{-\alpha_1.t}.\left[\cos(\alpha_2.t) - \alpha_3.\sin(\alpha_2.t)\right]\right\}\Phi(t) \tag{8}$$

$$\alpha_1 = \frac{1}{2.Q},\ \alpha_2 = \frac{1}{2}.\sqrt{4 - \frac{1}{Q^2}}\ \text{и}\ \alpha_3 = \frac{1}{Q.\sqrt{4 - \frac{1}{Q^2}}}$$

2. Critical value - at Q=0.5

$$Out2(t) = \left\{e^{-t}.[1 - t]\right\}\Phi(t) \tag{9}$$

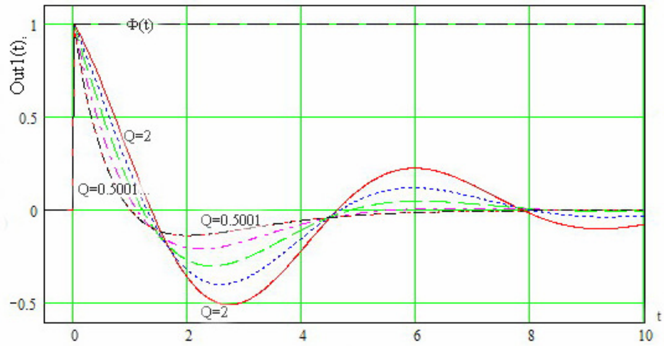


Fig. 7. Normalized step response of a loudspeaker at Q=0.501, 0.7071, 1, 1.414 and 2, according to Eq. (8).

3. Aperiodic damping at Q<0.5

$$Out3(t) = \left\{e^{-\alpha_1.t}.\left[\cosh(\alpha_4.t) - \alpha_5.\sinh(\alpha_4.t)\right]\right\}\Phi(t) \tag{10}$$

$$\alpha_1 = \frac{1}{2.Q},\ \alpha_4 = \frac{1}{2}.\sqrt{\frac{1}{Q^2} - 4}\ \text{и}\ \alpha_3 = \frac{1}{Q.\sqrt{\frac{1}{Q^2} - 4}}$$

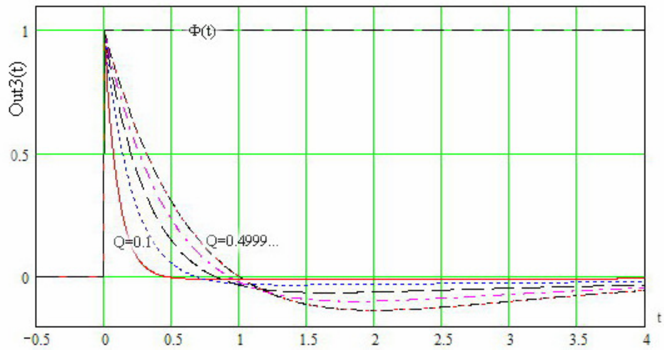


Fig. 8. Normalized step response of a loudspeaker at Q=0.1, 0.2, 0.3, 0.4 and 0.499... according to Eq. (10)

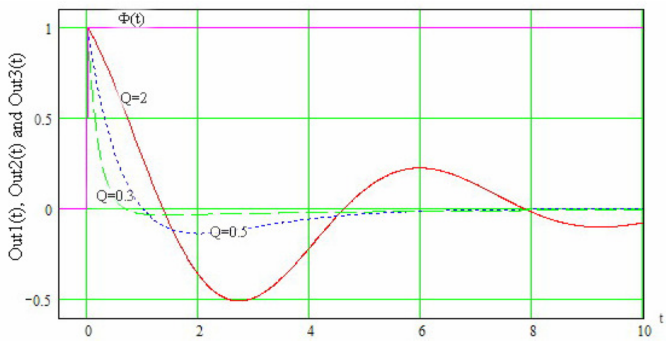


Fig. 9. Normalized step response of a loudspeaker at Q=0.3, 0.5 and 2, according to Eqs. (8), (9) and (10).

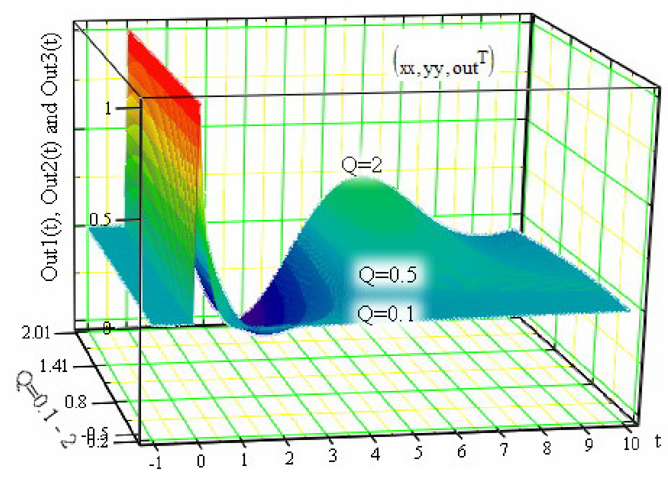


Fig. 10. 3D Normalized step response of a loudspeaker at Q=0.1÷2, according to Eqs. (8), (9) and (10).

Conclusion

The value of the total Q factor Q=0.5 appears to be critical for the analysis of the loudspeakers step response. The sound pressure created can be described with equation (9), and Fig. 9 shows the normalized step response.

The loudspeaker step response at Q>0.5 is with specific damped oscillations with the frequency of the mechanical resonance. Real world high-quality loudspeakers with a total Q factor of (0.5<Q<2) are hardly ever offered by producers.

Values of the Q factor within 0.1<Q<0.5 are typical for the real world high- and middle-frequency dome and low-frequency loudspeakers (see Table 1). Their step response can be described with Eq. (10), and the normalized step responses in Fig. 8 and Fig. 10 are marked with aperiodic damping of the signal.

The results obtained in this work can be used for theoretical analysis, design and production of loudspeakers, closed-box systems, etc.

References

[1] Small R. H., "Closed-Box Loudspeaker Systems Part I: Analysis," J. Audio Eng. Soc.,, vol. 20, Number 10, pp. 798-808, (1972 Dec.).

[2] Вълчев, Ив. Й., "Електроакустика", Държавно издателство "Техника", София, 1975г.

[3] W. Marshall Leach "Loudspeaker driver phase response: the neglected factor in crossover network design" 64th Convention November 2-5, 1979, New York City

[4] Алдошина И.А., А.Г.Войшивилло "Высококачественные акустические системы и излучатели", Москва, Радио и связь, 1985, стр. 64-102.

[5] Попянев, Д. Ф. "Електроакустични преобразуватели", Държавно издателство "Техника", София, 1988г.

[6] Евстатиев Г. "Ръководство за лаб. упр. по Основи на звукотехниката" ТУ, Варна 1990г.

[7] М. Момчеджиков и Сн. Плешкова, "Ръководство за лабораторни упражнения по основи на звукотехниката", издателство "Нови знания", София, 1999г.

[8] Poularikas Al. D, *The Handbook of Formulas and Tables for Signal Processing,* CRT Press LLC, 1999.

[9] Sirakov Ekaterinoslav S.,Atanaska A. Angelova and Georgi K. Evstatiev, Transitional characteristics of the Loudspeaker systems, ICEST 2003, 16-18 October 2003, Sofia, Bulgaria, pp 241÷242.

[10] Е.С. Сираков. RLC - вериги за съгласуване импеданса на високоговорител. "*Научно-техническо списание /ISSN 0861-4717/ "Електротехника и Електроника"*, бр. 1÷2, 2001г. стр. 49÷53.

[11] Е.С. Сираков. Макро-моделиране на електродинамичен високоговорител с директно излъчване в Spice формат. "*Proceedings of the Intern. Scient. Conf. of Energy and Information Systems and Technologies 2001 (EIST)"*, June 7-8,2001, Bitola, Rep. of Macedonia, vol. II, pp. 494-499.

[12] A.A. Angelova, E.S. Sirakov, G.K. Evstatiev. Theoretical analysis of Frequency, Pulse and Transitional characteristics of Loudspeaker (Part I). "*XXXIX Int. Scientific Conference of Information Communication. ICEST 2004"*, 16÷19 June 2004, Bitola, Macedonia, pp. 737÷738.

[13] http://www.mathcad.com/

[14] http://model.exponenta.ru/

[15] http://www.tu-varna.acad.bg/l1/l2/fe/RTT/eng.html

[16] Brochures of DYNAUDIO, Morel, Fokal and Audax.

Production of TV Multimedia Content: Modelling in Problem Space

Aleksandar Spasic[1], Jovica Bogdanovic[2] and Miloje Nesic[3]

Abstract – **Model of problem space (MOPS) deals with creating an understanding of the problem that the potential user of the system is facing. While usually it is the business problem that is being described, even a technical problem can be described at the user level in MOPS. Aim of this paper is to investigate the functional as well as dynamical behaviour of the user in system of software-intensive television production using the methodology of modelling in problem space and set of tools provided by Unified Modelling Language (UML).**

***Keywords* – Modelling in problem space, Software-intensive television production, Multimedia content life cycle**

I. INTRODUCTION

The new technology induced some fundamental changes in the basic workflows and business models of the content creation in the television industry. The dividing line between offline and online editing is eroded and the linear workflows of tape-based production are fragmented. Producers can now perform multiple tasks in parallel including media creation, editing, and compositing. The sequence in which production and post-production tasks occur is less important than it used to be. Production processes are changed, each department is involved and processes are coming closer. These changes are placing unprecedented strain on traditional production workflows and many of them collapsed under the pressure.

The new model of production and post-production is based upon: digital formats, the centralized management of media and metadata, non-linear assembly of media elements, high-speed networks, format agnostic distribution and automated processes.

Program makers in search of a solution quickly discover that there is no existing model within the broadcast and production industry to which they can turn. Today's "off-the-shelf" digital production solutions rarely do everything needed by the typical media enterprise. Ultimately, what is needed is a complete re-thinking of the way technology can be applied to the art and business of program making.

The main goal of this paper is to analyze this area of interest in a systematic way and to discuss underlying organizational and technical issues.

[1]Aleksandar Spasic is with Agency for Computer Engineering "String", Bore Stankovica 26, 18300 Pirot, Serbia & Montenegro. E-mail: aspasic@string.co.yu

[2]Jovica Bogdanovic is with "Sistel" Nis, Ktitor 19, 18000 Nis, Serbia & Montenegro. E-mail: sistel@ptt.yu

[3]Miloje Nesic is with Radio Television 5, Bulevar Zorana Djindjica 19/19, 18000 Nis, Serbia & Montenegro. E-mail: miloje.nesic@rtv5.co.yu

II. THE METHOD

A model, by its very nature, is an abstraction of the reality. The modeller, depending on his/her needs, keeps parts of the reality that are important to him/her in a particular situation and leaves out others which may be considered less important. Therefore, the model is not a complete representation of the reality.

Modelling raises abstraction to a level where only the core essentials matter. The resultant advantage is twofold: easier understanding of the reality that exists and efficient creation of a new reality [1].

Software projects use modelling throughout the entire life cycle. Subsequently, modelling is used not only to create the software solution but also to understand the problem. As a result, modelling occurs in the problem, solution and background (architectural) spaces.

Successful modelling needs to consider the areas in which modelling needs to take place. These modelling spaces have been formally considered and discussed by [1]. The three distinct yet related modelling spaces are defined: problem, solution and background.

In UML projects, model of problem space (MOPS) deals with creating an understanding of the problem, primarily the problem that the potential user of the system is facing. While usually it is the business problem that is being described, even a technical problem can be described at the user level in MOPS. In any case, the problem space deals with all the work that takes place in understanding the problem in the context of the software system before any solution or development is attempted.

Typical activities that take place in MOPS include documenting and understanding the requirements, analyzing requirements, investigating the problem in detail, and perhaps optional prototyping and understanding the flow of the process within the business. Thus the problem space would focus entirely on what is happening with the business or the user.

As a description of what is happening with the user or the business, the problem space will need the UML diagrams that help the modeller understand the problem without going into technological detail. The UML diagrams that help express what is expected of the system, rather than how the system will be implemented, are: use case diagrams, activity diagrams, class diagrams, sequence and state machine diagrams, interaction overview diagrams and package diagrams.

The UML diagrams in the problem space that are of interest here are:

Use case diagrams—provide the overall view and scope of functionality. The use cases within these diagrams contain the behavioural (or functional) description of the system.

State machine diagrams—occasionally used to help us understand the dynamicity and behaviour of the problem better.

III. Model of TV Production Workflow

Basic production stages are defined here as follows: development, planning, acquisition, processing, control, archiving and publication. These stages are shown on Fig. 1 as well as the production processes consisting of. At each step in the production workflow we can collect, and possibly re-use the metadata.

A programme's life traditionally begins with a need to fill a slot in a schedule. New skeleton schedule is produced from the analyses of the audience numbers and reactions. This schedule has to encompass details of the programme categories, possibilities for re-using (repeat) of the programmes as well as outline budgets of the programmes required to fit into slots.

During the development stage, programme ideas are investigated and a commission results when the producer persuades the TV company to finance the conversion of an idea into a real programme. The commission is very important for production as it gathers some key information like the 'working' title, producer's identity, possibly contributor's names, genre and possibly initial scripts. It could well have financial decisions which subsequently apply to the rest of the programme making process.

When a commission has been accepted research was doing and archives and other databases are examined for potential contributors, locations, facilities and material that can be re-used.

On the end of the planning stage a production order may be produced. The planning encompasses the staffing, resources and also the creation of the artistic description in the form of a storyboard and script.

During the acquisition stage, video shoots, audio clips and other programme items are created, pre-selected, ingested into production system and logged.

The obvious capture device is the camera, but equally, sound effects, graphics, stills, captions and music may all be added. At all points in capture there is an opportunity for metadata collection. Some of the metadata, like producer's comments and annotation, can only be captured by direct entry at the time of shooting. The metadata at this point in the chain should be viewed as 'portable', carried along with the essence as a link directly to a central.

The importance of the ingestion process is emphasized in [2] and noticed that *"crucial problem of Content Management Systems (CMS) is constituted by the ingestion of new content. As we cannot realistically expect that all the aspects of a production/archive environment are under the rules of a CMS, we need to set up gateways through which the content must pass when migrating from a non-managed environment to a CMS. The role of these gateways, that we call Ingestion Systems, is that of collecting and organizing as many relevant information (metadata) on the item as possible and that of generating all the content versions required by the CMS, including low resolution replicas of the essence, that can be exploited to economically implement browsing and offline editing functionalities, in such a way that time relations between the various versions are maintained."*

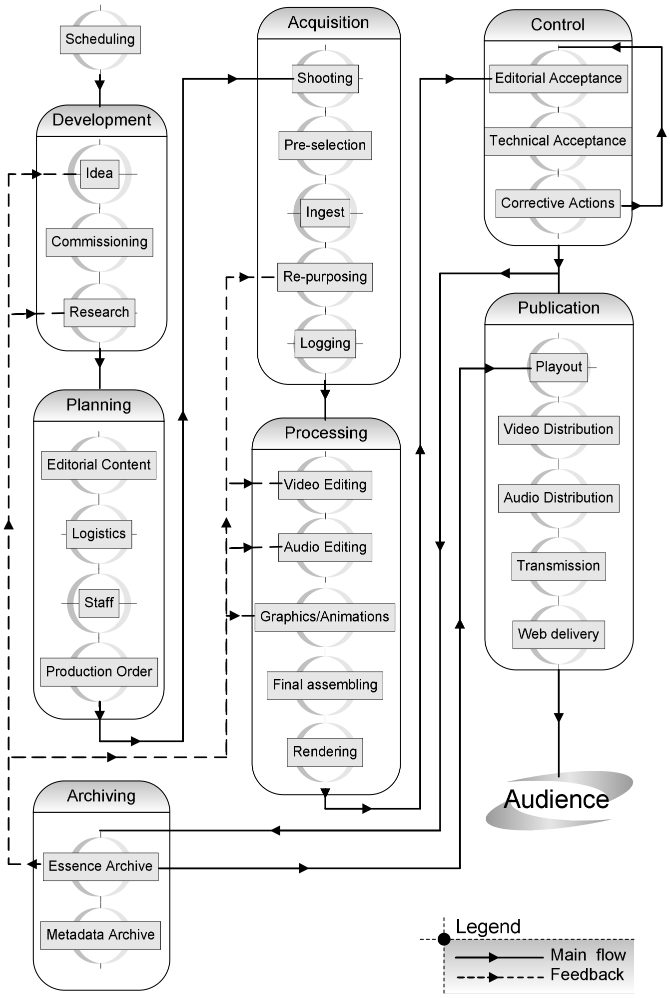


Fig. 1. Content Production Workflow

During the ingest we take all the content collected during a shoot, as well as new metadata, and transfer it into the production environment. We assume that the planning and commissioning metadata is already in the system. More metadata can be generated at ingest and this can either be directly entered, for example by an operator marking technically poor sections, or regions for special processing, or it can be extracted automatically.

Logging is where the producers review what they have, and mark down its possible use. It is expected that all the metadata capture that has taken place up until this stage will greatly reduce this overhead.

Processing stage represents a craftsman work where the shoots, clips, sounds and already assembled items are put into an order. Whole editing process, which is consisting of video and audio editing, has to be concentrated on capturing the composition metadata, so called Edit Decision List, in order to accurately represent the artistic composition of the programme from its constituents. Different graphics, subtitles as well as animations are produced and added to the essence.

Editorial and technical acceptances, which are the constituent parts of the recurrent control stage, approve the use of the produced programme material. If the corrections are needed, corrective action must be undertaken until editorial and/or technical approval is received.

Approved final product is catalogued and stored in archive. Archiving is one of the most important and most demanding organizational and technical processes in whole television production. Over time, media-rich organizations realized the value of their media assets. For instance, BBC Archive system has more than 750000 hours of television programmes in the archive, receives over 2000 enquires each weak and loans 45000 items per month [3]. Archival system is usually consisted of different servers such as workgroup media servers for short term storage and deep archive media servers for long term storage. Among the other things, archival systems can contain and manage metadata archives, low resolution archives as well as archives of still images, effects, sounds and other media related data. Archival in any form requires metadata to be captured and archiving is a prime candidate for metadata re-use, as the metadata is the basis for a comprehensive search. The capture of metadata not only enhances the search, but also removes some of the overhead and uncertainty that archivists can have in cataloguing the material.

Publication is the last but not least stage in the new production workflow. Playout process allows scheduled showing of the program produced at earlier stages. Programs, whether live or played from archive, are sent to the delivery point (transmitter chain, web etc.).

IV. FUNCTIONAL DESCRIPTION: USE CASE DIAGRAM

The main objective of a use case diagram is to visualize how the user (represented by the actor) will interact with and use the system. This is done by showing the actor associating with one or more use cases and, additionally, by drawing many use case diagrams.

Use case diagrams can be used by the project manager to scope the requirements. A comprehensive list of use cases in a use case diagram helps the users, together with the business analyst and the project manager, to decide which use case(s) to include in the initial iteration of the development cycle.

One of the important strengths of a use case diagram is its ability to model the actor (role). The actor demonstrates to the user who is involved in specifying requirements and where he exists in the context of the software system. In addition, the actor helps users to express their requirements in greater detail.

Use cases and use case diagrams help to organize the requirements i.e. use cases document complete functional requirements.

Use case diagram of the TV production as well as the description of the use cases in TV production are shown on Figs. 2 and 3, respectively.

V. DYNAMIC-BEHAVIOURAL DESCRIPTION: STATE MACHINE DIAGRAM

In terms of modern communications, business models need to account for the vital resources of production and distribution technologies, content creation or acquisition, and recovery of costs for creating, assembling and presenting the content [4].

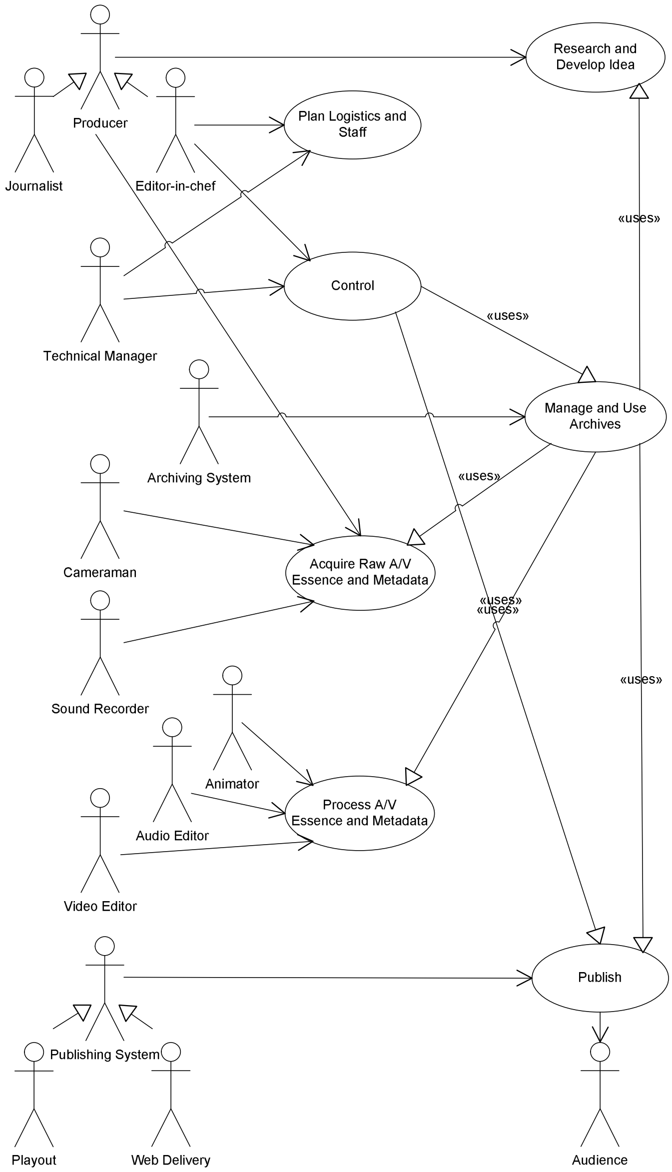


Fig. 2. Use Case Diagram of TV Program Production

The traditional emphasis of the media business has been the creation, bundling and distribution of content consisted of information and entertainment. In publishing and media, content is information and experiences created by individuals, institutions and technology to benefit audiences in venues that they value [5]. The creation of the content that is of interest to users is the basic issue in the broadcasting business model.

The nature of the state machine diagram is considered dynamic-behavioural. The state machine diagram of UML has the ability to represent time precisely and in a real-time fashion. “What happens at a certain point in time?” is a question that is answered by this diagram.

Use Case: Develop and Research Idea
Short Description: The producer research and develop an idea for production of a programme item
Actors: Producer, which is often the journalist or editor
Pre-Conditions: New programme schedule is produced and need for fill the slot exists
Post-Conditions: Programme ideas are investigated, commission results and initial script is made
Main Flow:
(1) Producer suggests an idea for program production.
(2) Producer persuades the TV company to finance the conversion of an idea into a real programme.
(3) Steps (1) to (2) are repeating until commission has been accepted
(4) Research is doing, initial script is made and Use Case terminates.

Use Case: Plan Logistics and Staff
Short Description: The managers plan logistics and staff needed for production of programme item
Actors: Editor-in-chef, Technical manager
Pre-Conditions: Programme idea is investigated, commission results and initial script is made
Post-Conditions: Production order is issued
Main Flow:
(1) Editor defines the editorial content based on inital script.
(2) Technical manager plans the logistics (objects, vehicles, production equipment...) needed for production.
(3) Editor and tehnical manager plan the staffing, journalists and production crew, respectively.
(4) Production order is issuing for all members of the production team.

Use Case: Acquire A/V Essence and Metadata
Short Description: Video shoots, audio clips and other programme items are creating, pre-selecting, ingesting into production system and logging
Actors: Producer, Cameraman, Sound Recorder
Pre-Conditions: Production order is issued
Post-Conditions: All essence materials, as well as related metadata, are ingested into production system and logged
Main Flow:
(1) Cameraman takes the shots in studio or terrain.
(2) Sound Recorder records the sounds in studio or terrain.
(3) Producer chooses the raw material or previously produced essence from archives and repurposes it.
(4) During the ingest, all the content collected during a shoot, recording and repurposing, as well as new metadata is taken and transfered into the production environment.
(5) Producer reviews what he/she has, and marks down its possible use. Use Case terminates.

Use Case: Process A/V Essence and Metadata
Short Description: Craftsman work where the shoots, clips, sounds and already assembled items are put into an order
Actors: Video Editor, Sound Editor, Animator
Pre-Conditions: All essence materials, as well as related metadata, are ingested into production system and logged
Post-Conditions: Essence material, as well as related metadata, are finalized
Main Flow:
(1) Video Editor edits the video essence
(2) Sound Editor edits the audio essence
(3) Animator makes the animations, graphics and subtitles
(4) Video Editor assembles and renders edited essence.
(5) Metadata are created in all phases of the editing process. Use Case terminates.

Use Case: Control
Short Description: Editorial and technical acceptances approve the use of the produced programme material
Actors: Editor-in-chef, Technical Manager
Pre-Conditions: Essence material, as well as related metadata, are finalizedingested into production system and logged
Post-Conditions: Essence material, as well as related metadata, are approved
Main Flow:
(1) Editor-in-chef checks and aproves/disaproves the editorial quality of the produced material.
(2) Technical Manager checks and aproves/disaproves the technical quality of the produced material.
(3) If the corrections are needed, corrective actions must be undertaken [A1]
(4) Steps (1) to (3) are repeated until the produced material is accepted and Use Case terminates.
Alternate Flow:
(A1) No need for corrections. Programme material is approved. Use Case terminates

Use Case: Manage and Use Archives
Short Description: Cataloguing, storing, search and retreiving programme material to/from archives
Actors: Archiving System
Pre-Conditions: Approved access to archives
Post-Conditions: Cataloged and stored essence, as well as related metadata. Retreived essence/metadata
Main Flow:
(1) After ingestion, system catalogues and stores raw material and related metadata and supports search and retreiving.
(2) After processing and control, system catalogues and stores final essence and related metadata and supports search and retreiving.

Use Case: Publish
Short Description: Playout, distribution and transmission of programme
Actors: Publishing System
Pre-Conditions: Programme prepared and approved for publishing
Post-Conditions: Programme transmited/delivered
Main Flow:
(1) Programme playout.
(2) Programme distribution, both audio and video components, as well as data component.
(3) Programme transmition using terrestrial, cable or satellite transmission[A1].Use Case terminates.
Alternative flow:
(A1) Programme delivering using web services. Use Case terminates.

Fig. 3. Description of the Use cases in TV Program Production

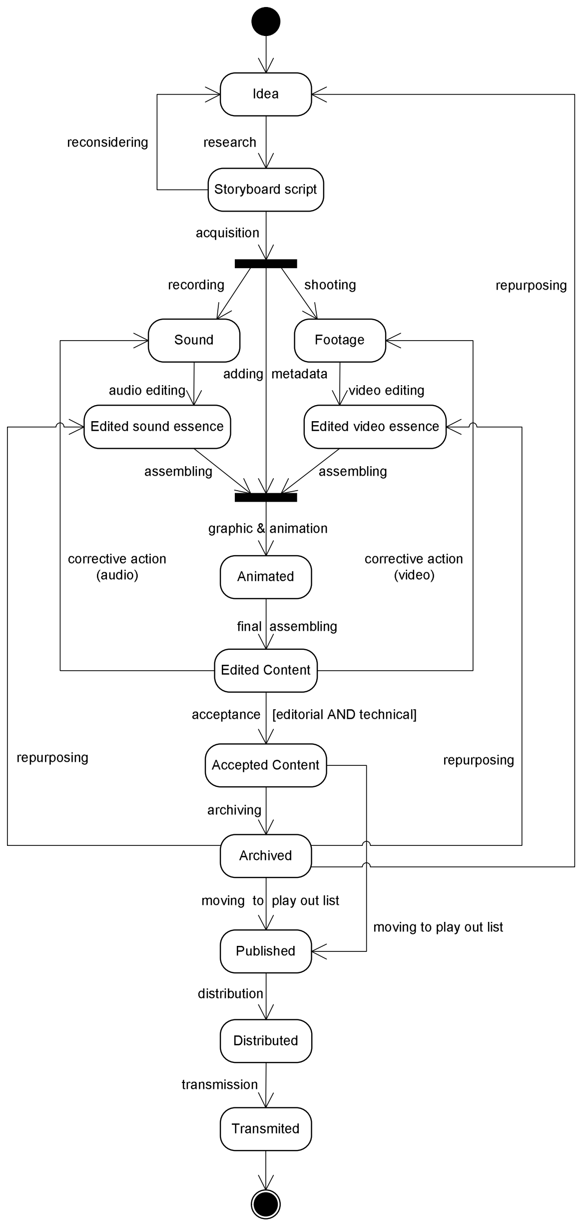


Fig. 4. Content State Machine Diagram

VI. CONCLUSION

As content is one of the most valuable assets for broadcasting companies, ingesting, archiving, accessing, managing, delivering and security of digital content assets become basic requirements in the everyday life of multimedia producers and providers; at the same time, it becomes ever important the way the company structures the processes involved and how it chooses the technologies that best adhere to the purpose related to content handling.

Model of problem space (MOPS) deals with creating an understanding of the problem, primarily the problem that the potential user of the system is facing.

Partial model of problem space related to the production of television content is suggested and analyzed in this paper.

REFERENCES

[1] B. Unhelkar, *Verification and Validation for Quality of UML 2.0 Models*. Hoboken, New Jersey: John Wiley & Sons, Inc., 2005
[2] D. Airola, L. Boch, G. Dimino, *Automated Ingestion of Audiovisual Content*, 2002, http://www.broadcastpapers.com/asset/IBCRAIAutoIngestAVContent.pdf
[3] J. Evans, *The Future of Video Indexing in the BBC*, 2003, http://www-nlpir.nist.gov/projects/ tvpubs/papers/ bbc.slides.pdf
[4] R.G. Picard, Changing Business Models of Online Content Services - Their Implications for Multimedia and other Content Producers, The International Journal on Media Management, 2 (2), 60-68.
[5] http://en.wikipedia.org/wiki/Content

Possibilities for incidence of SMIL based multimedia applications

Kalin H. Monov[1]

Abstract – **Some different possibilities for incidence of SMIL multimedia applications are proposed in this report. A variant for transferring of SMIL clips using Helix Universal Server 9.0 is discussed and the basic server characteristics and advantages are described.**

Keywords – **HUServer, SMIL, multimedia applications.**

I. Introduction

The necessity of developing and incidence of multimedia broadcast clips with comparatively small size and good proportion quality – using video compression degree is result of digital networks evolution for information transfer like global network Internet. On-line incidence of multimedia clips, realized with help of rather using SMIL language v2.0, is provocation to companies are developing platforms for digital stream information transfer and safe-keeping. The company Real Networks Inc. is one of leader on the software market for digital media transfer and control with Helix Universal Server.

The purpose of this paper is to represent the platform Helix Universal Server v9.0 and exemplary approach for SMIL based clips incidence.

II. Helix universal server v9.0 – basic performances

It is possible to construct flexible and sensible decision in respect of stream incidence of multimedia clips using stream technologies of Real Networks. The Helix server give powerful administrative tools for stream control.

The Helix Server can distribute stream clips and on-line transmissions in different file formats [5]. The Helix server can operate with file format of different companies that are presented in table I. The platform use as well as with represented formats also with other by means of plug-in.

The Helix server run properly as using operation systems MS Windows also systems based on UNIX. This is the reason to be possible using of desire file formats from desire operation system. The Helix servers that using different operating systems are completely compatible in various network environment.

This server is property for using in NGN (Next Generation Network) [1]. The NGN services that can be realized by means of Helix are Video on Demand, VoIP and Data.

[1]Kalin H. Monov is with the Faculty of Electrical Engineering and Electronics, Technical University – Gabrovo, 4 H. Dimitar St., 5300 Gabrovo, Bulgaria, E-mail: monov@tugab.bg

TABLE I
USED FILE FORMATS IN HELIX SERVER

Company	File Format
RealNetworks	RealAudio (.rm), RealVideo (.rm, .rmvb), RealPix (.rp), RealText (.rt)
Macromedia	Flash (.swf)
Microsoft	Windows Media (.asf, .wma, .wmv)
Apple	QuickTime (.mov)
Standards-Based	MPEG-1, MPEG-2, MPEG-4, MP3
Image Formats	GIF (.gif), JPEG (.jpg), PNG (.png)
Other	AU (.au), AIFF (.aif, .ief), WAV (.wav)

The Helix server distribute stream clips and on-line transmissions, but it not support into tools for their creation.

They are three basic steps for clips stream incidence: clip codding with corresponding code tool, its incidence by means of given server and its reproduction by client side [5]. They are code tools that receive on-line transmission on their input, code them like stream string and they incidence them without a necessity to be safed. An example scheme of incidence stream clips process is given on figure 1.

The possibility for reproducing programs like QuickTime, RealPlayer, Windows Media Player, and Web browser using is typical for the incidence process. That using of great number of reproducing programs give a possibility for server platform independence by client size.

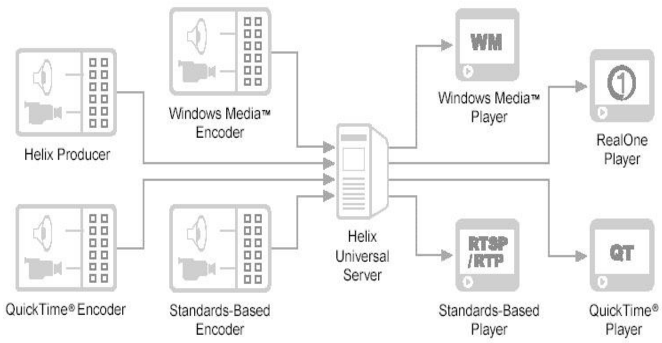


Fig. 1. Universal scheme for stream media incidence

III. Used protocols

The Helix Server can transfer stream information in LAN or Internet. Although the server has possibility for HTML pages delivery it usually is used along with distant Web server as code transfering of HTML pages and stream media. The protocols that are using for stream media transfer are of interest. These protocols are:

- Real Time Streaming Protocol (RTSP). RTSP is a standard based protocol for stream data transfer and it is recommended from Internet Engineering Task Force [6]. This protocol provide for server a possibility to communicate with different reproducing programs and also with RTSP MPEG players.

- Progressive Networks Audio (PNA). PNA is an older private protocol that is used in earlier version of RealSystem Server and RealPlayer.

- Microsoft Media Services (MMS). MMS is a private protocol that is used from Helix Server for communication with Windows Media Player [Microsoft].

- Hyper Text Transfer Protocol (HTTP). Although HTTP is not a protocol for stream media files delivery, the Helix Server utilize its for Helix Administrator HTML pages transfer that allow server configuration.

The Helix Server possess additional options like access authentication and control, observation and report creation.

IV. Stream transmission types

The on-line stream information (clips) transmission is one of the used and the applicable possibility of Helix Server. The Server provide for some possibilities for on-line stream information transmission:

- single transmission (Unicasting). Unicasting is the simple method for on-line transmission. It is typical that every one reproducing program supply one's own stream. Single transmission is limited of license client links number.

- multiple radiating (Multicasting). Multicasting curtly decrease a broadband width and it gives a possibility for more users including. In this case the reproducing programs give not one' own streams. Instead all of them are connected to one general stream. There is a disadvantage that Multicasting require a network with high bandwidth.

The Helix Server operate with three types of multiple transfer: multiple transfer with back channel (back-channel multicasts), scaling multiple transfer (without control channel), multiple transmissions for Windows Media (scaling transmissions only for Windows Media Player).

- Splitting. Transfer the stream between two or more Helix Servers. When the number of servers that radiate given stream increase then the number of users using such type servers and a possibility for multiple transmission behind fire walls increase too.

V. Interaction between SMIL and Helix server

SMIL is well-known format for multimedia presentations codding that are using mainly in LAN and Internet. SMIL is a creation of W3C and it is an abbreviation from Synchronized Multimedia Integration Language [2].

The SMIL documents can be created by different methods. The often method is SMIL editors using. They exist many of programs for SMIL clips creation [3]. The primary SMIL advantage is the possibilities for multimedia objects synchronization [2]. They exist two basic ways to incidence of SMIL clips using Helix universal server v9.0:

- using web browser;

- by means of reproduction programs like: Real Player of Real Networks Inc. and QuickTime.

VI. Experimental results

The results that are achieved at the moment can be present in following directions:

- Helix Server experimentation and configuration;

Helix Universal Server v9 configuration is possible to be realized on operation systems like WindowsNT, 2000, XP, 2003. Installation procedure is not difficult because an installation interface is good realized. A server capacity is tested using Helix Administrator console in laboratory.

- creation of multimedia clips, based on SMIL v2.0;

It is used a TagFree2000 editor of Dasan Technology for example SMIL clips creation. For experimental server test in example clips were included as static also dynamic media objects and they are used all three methods for synchronization in SMIL [6]. These methods are subsequent (seq), parallel (par) and synchronization on exception (excl).

- integration of SMIL clips into Helix Server environment;

Helix Server and SMIL using is suitable combination for realization of different learning tools or some part of them, multimedia presentation clips, also for on-line transmissions radiating. For taken experiments is observing an essential disadvantage – a necessity of Real Player G2 installation, but it is free. It is possible to use web browser – HTTP based incidence of corresponding clips.

At this moment interaction between SMIL clips and Helix server is realized only in laboratory using LAN at Technical University of Gabrovo. They were used as simple also multiple clips transmission. When multiple transmission is used there is some delay in respect of simple transmission. Because users number of server free license file is limited maximum connections to it are only 15. From permitted 15 connections when the server was tested were used only 10.

VII. Conclusion and future work

The material represented in this report is new and interesting approach about multimedia presentation incidence. Using of platforms for stream media incidence like Helix Universal Server in combination with SMIL has the following advantages and disadvantages:

– possibilities of different file formats integration;

– a possibility of presentations incidence in networks with different bandwidth;

– a possibility of settings of received media information from users size according to his requirements;

– a good possibility for synchronization to last user;

– a possibility for on-line transmissions realizing;

– limitations in respect of reproducing programs.

References

[1] J.C. Crimi, *Next Generation Network (NGN) Services*, Telcordia Technologies, 2002.
[2] www.w3.org/TR/1998/REC-smil-19980615
[3] Dick, C., A., Bulterman, *SMIL 2.0*, IEEE 2001, Oct-Dec 2001.
[4] K. Monov, R. Ilarionov, “SMIL – Conception and Possibilities for Construction of Multimedia Presentation. Synchronization”, *International Scientific Conference UNITECH'03, Conference Proceedings*, vol I, pp.270-273, Gabrovo, Bulgaria, 2003.
[5] www.real.com.
[6] www.ietf.org.

Software Tools for Network Modelling and Simulation

Marek Huczala

Abstract – **The presented paper introduces Java SSFNet platform for computer network modelling and simulations. The SSFNet platform is a Java-based software interface that might be implemented on its own or with a co-operation with an upper level software kit. First, we discuss the library classes for the network scheme development and the ways of launching the SSFNet network simulation process. Later chapters provide an overview to building an upper level software interface to interact with SSFNet models.**

Keywords – **Java, SSFNet platform, NetSim, network, simulation, modelling.**

I. Introduction

Network modelling and simulation plays an important role in today's computer network design. The main intention of the following paper is, however, to outline the "software paths" that would lead into an improved network model design and a better overview of the simulation results. The paper could be a developer's key to organizing and running a new simulation process under the SSFNet platform.

Since all the SSFNet software has been written in Java programming language, here mentioned software solutions are also based on Java. Using the solution, any developer will be now able to easily define a new network model and launch the simulation by calling SSFNet.

II. Overview of Java SSFNet Platform

The SSFNet is a collection of Java SSF-based components for modelling and simulation of Internet protocols and IP based computer networks. By default, the SSFNet components are represented by Java pronsipal classes that were later united into the following two main software frameworks:

- SSF.OS is used for modeling of the host and operating system components. Network and transport layer protocol such as SSF.Net.IP and SSF.Net.TCP are laid on top of SSF.OS class.
- SSF.Net is used for modeling network connectivity, creating nodes and link configurations. It loads all the model's configuration file and controls the orderly instantiation of the entire model: hosts and routers with their protocols, links connecting hosts and routers, as well as traffic scenarios and multiple random number streams.

Marek Huczala, Ing., PhD student at Telecommunications Department, Faculty of Electrical Engineering and Communication Technologies, Brno University of Technology; Purkyňova 118, 612 00 Brno, Czech Republic. E-mail: huczala@kn.vutbr.cz.

III. Network Model Design

The network configuration is stored in DML scheme definition file. New DML file uniquely describes the complete network architecture from both hardware and software aspect. The configuration file follows the DML syntax structures that allow keyword, value specifications. DML syntax grammar is based on standardized well-known XML structure.

Network file scheme definition begins with keywords scheme and Net as it follows:

```
schemas [
  Net [
       frequency  simulation_runtime
```

The frequency attribut specifies the total time of simulation process. Network, always defined by the Net keyword, represents a set of hosts, routers, links and when modelling more complex network environments even subnets.

Here is an example of a subnet definition using the Net keyword:

```
Net [
  id id_no
  idrange id_no from --- to ---
  .
  .
  .
  ip net_mask_def
  _extends .schemas.Net
]
```

The id and idrange attributes are used for subnet identification while ip attribute passes on network mask definition. The _extends attribute specifies the higher level syntax used - in current example it follows the default Net scheme. The SSF.Net.host intruduces the following host configuration scheme:

```
host [
  id id_no
  idrange id_no from --- to ---

  ]
```

Router definition very likely follows the host definition. Interfaces of the implemented network elements are described by a local keyword interface followed by a set of attributes such as bitrate, latency or virtual. Links connecting network nodes are set up by link keyword whereas the traffic flow between nodes is defined by traffic keyword and its attributes. Every fragment of the DML definition file is processed by a corresponding class. For example host definition is being handled by SSF.Net.host class while links by SSF.Net.link principal class.

By default, the network definition scheme is interpreted by SSF.Net.Net class. It loads the complete network model, network elements such as routers, hosts, protocols and links specifications as well as network traffic scenario.

IV. Overview of SSFNet Pre-simulation Process

Pre-simulation process is usually invoked by the main function of SSF.Net.Net class. The process itself comprises of several stages. The first stage involves schemantical DML scheme check:

```
if (doSchemaCheck)
  netconfig.check();
  Configuration netcfg =
(Configuration)netconfig.findSingle(".Net");
 if (netcfg == null) {
      System.exit(-1);
      }
if (null != netconfig.findSingle(".link") ||
      null != netconfig.findSingle(".router") ||
      null != netconfig.findSingle(".host")) {
                System.exit(-1);
            }
```

Within this pre-simulation stage some essential parameters are being checked, such as link, router or host. The IP address space of all networks and subnets is being allocated in the 2nd stage. Routing informations are added subsequently. Links and connections between network nodes are being checked in the end of the pre-simulation process.

At programmer's view, the SSF.Net.Net object will use the services of the DML library to load the content of the configuration files and net.dml (by default) into a runtime Configuration database object. After that, SSF.Net.Net will systematically instantiate and configure all simulation objects such as hosts, routers, protocols, and network links. Once all simulation objects have been instatiated, the initialization phase begins by calling init() methods of all the entity subclasses. Finally SSF.Net.Net invokes its method startAll(), and the simulation begins with simulation time value equal to 0 sec. There is a lot of verbose output, including the automatically generated IP address blocks etc., that may be suppressed by command line options to SSF.Net.Net.

The simulation runtime may be specified as a command line argument to SSF.Net.Net class or directly inside the configuration scheme file. The Java Development Kit 1.3 or higher is required for running all the SSFNet simulations.

V. Developing SSFNet Applications

The simulation results as they come out of the SSFNet simulation process can only be viewed in a text mode. This was the main reason for me to start building a new application that would provide a graphical interface to SSFNet simulation. Knowing this and with a knowledge of Java programming language, now anyone himself could build a new graphical application interface that would show the SSFNet simulation results in a windows-like style frame.

This chapter briefly describes a newly build graphical interface to SSFNet Simulation, called NetSim as well as other ways to extend the SSFNet simulation framework.

Only small adjustments to the SSFNet source code were made while building NetSim. First of all, the pre-simulation phase needed to be a little more transparent. This required adding small code fragments to the SSF.Net.Net class. The purpose of the change was to enable a possibility to monitor the current state of the pre-simulation process. Secondly, I added the dimensions to the host and router elements in network scheme DML file structure which resulted in a contructors' change of SSF.Net.Host and SSF.Net.Router classes. The dimension parameter is now being used by JGraph (see www.jgraph.com) graphical environment to draw the picture of network scheme. The network scheme picture is shown in an internal frame, see Fig. 1.

NetSim is fully open for further adjustments and improvements. Adding the possibility to trace the simulation process could be one of them.

The NetSim application window consists of a network scheme internal frame, state bar showing the current simulation state and the simulation results frame. Figure 1 shows network scheme and results of a simple client-server simulation running on the SSFNet platform. Simple client-server communication model is used in this example. The picture of a client-server scheme is drawn during the model initialization process in the upper frame. The simulation results are shown in the bottom of the main window.

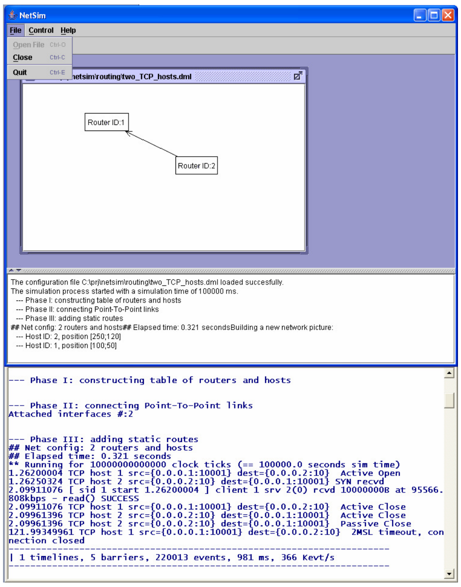


Fig. 1. Sample graphical user interface (GUI) application showing results of the SSFNet network simulation.

The advantages and disadvantages of using NetSim come from the properties of Java programming language. The object-oriented program code is simple to read and easy to change and simulations can be run under different operating systems, including Windows or Linux. The application NetSim and simulation platform SSFNet, however, require a huge memory space when modelling complex network environments.

VI. CONCLUSION

The paper introduces a Java-based platform for network modelling and simulations SSFNet. SSFNet kernel and its source code can be easily downloaded and installed from www.sffnet.org.

NetSim is a Java-based graphical user interface that was developed to graphically demonstrate the results of SSFNet simulation process. The Jgraph (www.jgraph.com) graphics library is used to picture the detailed network scheme.

Only small adjustments to the SSFNet source code were made while building NetSim. The application and the SSFNet platform is fully open for further adjustments and improvements. Adding the possibility to trace the simulation process could be one of them.

The NetSim package and the installation instructions can be downloaded from the authors web site (http://hawk.cis.vutbr.cz/~huczala/vizualizace).

REFERENCES

[1] HUCZALA, M. Vizualization of routing algorithms in TCP/IP network environments, final report to the FRVŠ grant project, Brno 2005.

[2] SSFNet Community. SSFNet 1.3 DML Reference, www.ssfnet.org.

[3] SSFNet Community. Implementation and Validation Tests, http://www.ssfnet.org/Exchange/tcp/index.html.

[4] SSFNet Community. SSFNet software exchange, Package overview, http://www.ssfnet.org/exchangePage.html.

Using String Comparing Algorithms for Serbian Names

Petar J. Rajković[1], Dragan S. Janković[2], Dušan M. Vučković[3]

Abstract – **String matching algorithms are widely used in many areas. Some of them are adapted for special languages and special type of words, as for example, person names or company names. We have researched the possibilities of using string matching algorithms on Serbian names. Report about it is presented in this paper. As we expected and as experiments show, some of phonetic algorithms are not suitable for Serbian names. On other side, distance measure based algorithms can be applied. Our results are good starting point for modification of existing or constructing of a new algorithm suitable for Serbian language.**

Keywords – **String matching algorithm, code resulting, similarity resulting, Serbian names, comparing strategies.**

I. Introduction

Government and commercial organizations are increasingly required to store, maintain, search and match identity data from many nations and numerous languages. A variety of algorithms have been published to allow approximate verbal matches to be found in documents or databases. In each case, the user specifies a word and the system retrieves records containing similar ones.

The reasons for wrong typed words in some text in general are typing errors and spelling errors. The most frequent typing errors are: deleted letter, inserted letter, replaced two letters, added letters, removed letters, used abbreviations, split words, joint words, etc. For different kinds of typing errors has been developed different kinds of algorithms. Some of them detect a set of above numbered errors.

Two main classes of algorithms can be distinguished: those that determine word similarity by examining the order of the letters, and those that rely principally on phonetics. The second class of the algorithms depends strongly of the chosen language. Also, there are the combinations of above two approaches (for example editex algorithm [1]).

The most of the phonetic algorithms are constructed and adapted for English names. So it is interesting to investigate the applicability of these algorithms on Serbian languages and especially on names.

So we developed program for the most known string matching algorithms from both kind phonetic and distance matching (Jaro – Winker, Levenstein, NYSIIS, Metaphone, Double metaphone, different SoundEx algorithms, etc) and apply them on Serbian names. Results are reported.

Our future work will be upgrading of presented algorithms in order to become more suitable for Serbian language, as well as, enlarge our test base.

II. Used algorithms

For implementing our string matching application we used two different kinds of algorithms – similarity based and code resulting algorithms. From the group of similarity based we have tested Jaro – Winkler and Levenstein, and from group of code resulting methods we have used NYSIIS, Metaphone, and different implementation of SoundEx algorithms (Daitch Mokotof, and four standard modifications – Miracode, Simplified, SQLServer, and Knuth Ed2).

Jaro Winkler algorithm is a kind of a measure of similarity between two strings. The Jaro measure [2] is the weighted sum of percentage of matched characters from each file and transposed characters. Winkler increased this measure for matching initial characters, and then rescaled it by a piecewise function, whose intervals and weights depend on the type of string (first name, last name, street, etc.). This is an extension of the Jaro distance metric, from the work of Winkler in 1991 to 1999 [3].

Levenshtein algorithm is based on calculating distance that is obtained by finding the simplest way to transform one string into another [4]. Transformations are the one-step operations of (single-phone) insertion, deletion and substitution. In the simplest versions substitutions cost two units except when the source and target are identical, in which case the cost is zero. Insertions and deletions costs half that of substitutions. On the base of these values similarity is computed according the length of the source string.

NYSIIS is a member of group phonetic coding algorithms. Basically, it has been used to convert a name to a phonetic coding of up to six characters [5]. Now, NYSIIS codes can be larger than six characters. NYSIIS is the short form of the *New York State Identification and Intelligence System* Phonetic Code. It features an accuracy increase of 2.7% over the traditional SoundEx algorithm. It is a pretty simple algorithm described in *Name Search Techniques*, New York State Identification and Intelligence System Special Report No. 1, by Robert L. Taft, is and it has some seven steps that converts word to string that represents its code.

Metaphone (we use its *double metaphone* variant) is an algorithm to code English words (and foreign words often heard in the United States) phonetically by reducing them to 12 consonant sounds [6]. This reduces matching problems from wrong spelling in English language.

Soundex is a phonetic algorithm for indexing names by their sound when pronounced in English [7]. The basic aim is for names with the same pronunciation to be encoded to the

[1] Petar J. Rajković is from Faculty of Electronic Engineering, Beogradska 14, 18000 Niš, Serbia & Montenegro, E-mail: rajkovicp@elfak.ni.ac.yu
[2] Dragan S. Janković is from Faculty of Electronic Engineering, Beogradska 14, 18000 Niš, Serbia & Montenegro, E-mail: gaga@elfak.ni.ac.yu
[2] Dušan M. Vučković is from Faculty of Electronic Engineering, Beogradska 14, 18000 Niš, Serbia & Montenegro, E-mail: dvuckovic@elfak.ni.ac.yu

same string so that matching can occur despite minor differences in spelling. Soundex is the most widely known of all phonetic algorithms and is often used (incorrectly) as a synonym for "phonetic algorithm". Soundex was developed by Robert Russell and Margaret Odell and patented in 1918 and 1922. A variation called American Soundex (U.S. SoundEx) was used in the 1930s for a retrospective analysis of the US censuses from 1890 through 1920. The Soundex code for a name consists of a letter followed by three numbers: the letter is the first letter of the name, and the numbers encode the remaining consonants. Similar sounding consonants share the same number so, for example, the labial B, F, P and V are all encoded as 1. Vowels can affect the coding, but are never coded directly unless they appear at the start of the name.

The one of the latest significant improvement of basic SoundEx is the Daitch-Mokotoff algorithm. In 1985, this author indexed the names of some 28,000 persons who legally changed their names while living in Palestine from 1921 to 1948, most of whom were Jews with Germanic or Slavic surnames. It was obvious there were numerous spelling variants of the same basic surname and the list should be soundexed. It is a modification to U.S. SoundEx.

III. TESTING

In order to test all previously described algorithms we have develop simple Windows based application in Visual Studio 2003, using C#.Net named Word matcher. We have implemented Jaro – Winkler and Levenstein similarity algorithms, as well as following code resulting methods: NYSIIS, Metaphone (as Double Metaphone), Caverphone, Daitch Mokotoff SoundEx, and four variants of standard SounEx algorithms (Knuth Ed2, Simplified, Miracode, and SQLServer SoundEx).

This application provides us possibility to test two words or two sentences (Figure 1), or one word (or sentence) with strings from some source file (Figure 2). Source files only have to be placed in the same folder with executable file and they will be loaded. The comparison results can be saved in the text file and processed later.

For example of usability of previously described algorithms we will present results of testing similarity of last name ***Jankovic***. All testing results are presented by tables that are consisted of three columns – the count of found similar words, minimal similarity (in percents), and duration of this operation in seconds. All tests are done on computer with Pentium Mobile processor on 1.8 GHz with 512MB of RAM.

The source for this testing is text file that contains list of 22505 different words. The list members are first names, last names, names of settlements and other commonly used words that are collected by different organization in Serbia and Montenegro. Some of them are written using Serbian alphabet specific characters (š, č, ć, đ, ž) and some of them are written using English alphabet. The large number of the collected lat names is presented on both two ways (e.g. you can find both ***Rajkovic*** and ***Rajković*** in the list).

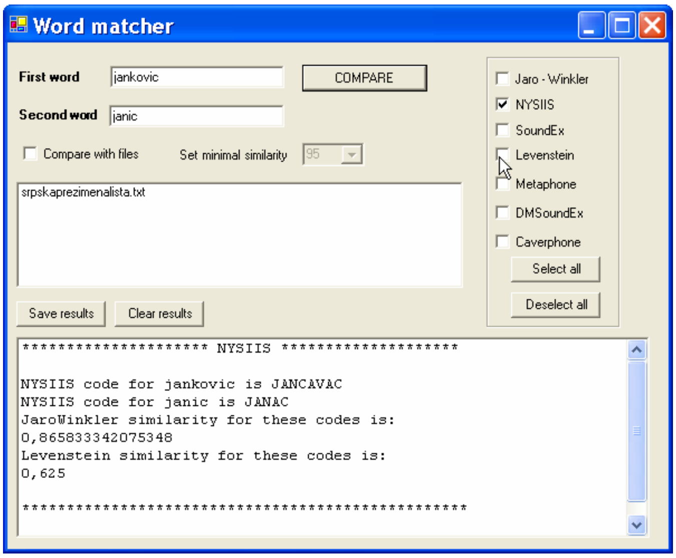


Fig. 1. Comparing two words by selected method

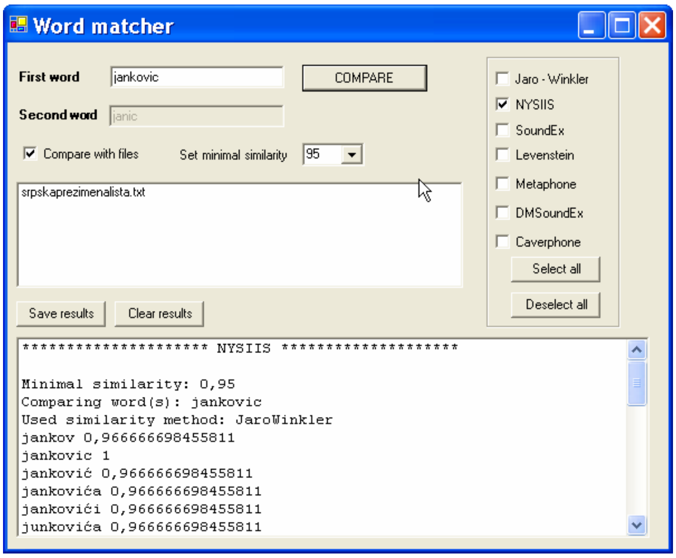


Fig. 2. Comparing certain word with all words from specific file by selected method

In the following text different comparing strategies are tested and some of testing results are presented. Generally, there are four strategies for testing similarity between two strings:

- Exact matching – comparing two strings in order to determine if they are equal.
- Using similarity method – comparing two strings using some algorithm that will return us some value (between 0 and 1) that will be information about similarity level.
- Using code resulting method – comparing codes that are generated by using some code resulting algorithm, in order to determine if they are equal.
- Using similarity for generated codes – apply similarity method to determine level of similarity of passed strings' phonetic codes.

When comparing one word (test word) with the list of words we will obtain, as a result, the list of the words with similarities corresponding to test word. In the case when comparing list has 2000 words, list of results will have 2000 results. In order to reduce the size of resulting list we should determine some kind of threshold for placing certain word to the list of the results. And this threshold is the level of minimal similarity. If we want to use matching strings by similarity for some spelling helper we need result list with not more than 10 to 12 words. If we want to solve this kind of problem we have to use matching string by similarity, otherwise exact matching will provide us only one solution – the searched word. The results of using Jaro – Winkler and Levenstein similarity methods are presented in following two tables.

TABLE 1. THE RESULTS OF COMPARISON OF WORD ***JANKOVIC*** WITH REPRESENTATIVE LIST OF NAMES USING ***JARO – WINKLER*** SIMILARITY METHOD

Found similar words (count)	Minimal similarity level (percent)	Duration (seconds)
1270	70	0.375
416	75	0.21875
164	80	0.203125
36	85	0.1875
13	90	0.171875
6	95	0.171875
1	99	0.171875

TABLE 2. THE RESULTS OF COMPARISON OF WORD ***JANKOVIC*** WITH REPRESENTATIVE LIST OF NAMES USING ***LEVENSTEIN*** SIMILARITY METHOD

Found similar words (count)	Minimal similarity level (percent)	Duration (seconds)
24	70	0.171875
7	75	0.15625
3	80	0.15625
3	85	0.21875
1	90	0.15625
1	95	0.15625
1	99	0.15625

As it has been discussed in previous part of this paper critical measure for discussed problems is setting of minimal similarity in order to obtain list of synonyms that has reasonable number of members (less then 10). The Jaro – Winkler algorithm returns following words when minimal similarity is set on 95%: *jankov* 0.967, *jankovic* 1, *janković* 0.967, *jankovica* 0.954, *jankovići* 0.954, and *jankovi* 0.983. Each word is followed by its similarity level with word ***jankovic***. Comparing word ***rajkovic*** with mentioned list of names will return six words for minimal similarity of 95% (*rajkov* 0.967, *rajkovica* 0.954, *rajkovići* 0.954, *rajkovac* 0.967, *rajković* 0.967, and *rajkovci* 0.983) and 14 words for 90%. On the base of comparing of many other Serbian last names we could say that "reasonable" threshold for minimal similarity level for Jaro – Winkler method should be set on some value between 90 and 95%. By example, the level of 92% will return list of 10 similar words with word ***rajkovic*** and 8 similar words with word ***jankovic***.

In the case of Levenstein alghorithm minimal similarity level could be set on lower percent number – 70 to 80 percent. For word ***jankovic*** and threshold of 75% our testbench application returns seven words: *jankovic* 1, *stankovic* 0.778, *janković* 0.875, *jankovica* 0.778, *jankovići* 0.778, *janaković* 0.778, *jankovi* 0.875. For similarity of 70% the number of similar words is 24. Testing Levenstein algorithm on word rajkovic returns 3 words (*rajkovac* 0.875, *rajković* 0.875, and *rajkovic* 1) for 80% threshold, 10 words (brajković 0.778, brajkovac 0.778, raškovice 0.778, rašković 0.778, rajkovic 1, ranjković 0.778, rajkovac 0.875, rajković 0.875, rajčković 0.778, trajković 0,778) for 75% and 35 words for 70%.

When using code resulting algorithms the results are little bit different. Next table (table 3) shows number of matching for different code resulting algorithms. The main idea in this kind of matching is find specified phonetic code for supplied word and get all words from list that has exact code.

TABLE 3. THE RESULTS OF COMPARISON OF WORD ***JANKOVIC*** WITH REPRESENTATIVE LIST OF NAMES USING DIFFERENT CODE RESULTING METHODS

Used algorithm	Number of matching words	Duration (seconds)
NYSIIS	1	0.625
Metaphone	17	0.1875
Daitch Mokotof SoundEx	4	49.421875
Knuth Ed2 SoundEx	19	0.15625

The NYSIIS code returns only one word (*jankovic*), but it is not so desirable result. The Daitch – Mokotof SoundEx algorithm has found more reasonable number of similar words (4 – *jankovic*, *smokvice*, *smokovac*, *smokovica*), but it took too much time for this operation and found words that are not adequate for Serbian language. Knuth Ed2 Soundex (found 19 words – *jankov*, *jankovic*, *janković*, *johanesburg*, *jankovica, janjevići, jankovići, junkovica, janjušević, junković, janjevica, janjuševica, janaković, janićijević, jankovi, junaković, junuzovci, jankovići, janjević*) and Metaphone (17 words – *jankov, jankovic, janković, anković, jankovica, jankovići, junkovica, junković, inković, inkovići, janaković, jankovi, junaković, onković, jankovci, unkašević, unković*) are fastest and produce more reliable results than NYSIIS and DMSoundEx. The only problem here is that Knuth Ed2 Soundex and Metaphone return more word that users usually expect.

On the base of larger number of examples we have found that the most suitable code resulting algorithm for Serbian words is Metaphone. Metaphone is little bit slower than Knuth Ed2 SoundEx, but it is able to provide the most acceptable lists of similar words. Also, in some cases, union between Metaphone's and Knuth Ed2 SoundEx's resulting lists can be best solution. The problem with large lists of synonyms also remains. When comparing word ***rajkovic*** the results are: one synonym for NYSIIS, 8 for Metaphone, 12 for DMSoundEx, and 18 for Knuth Ed2 SoundEx.

The usage of the NYSIIS algorithm for comparing Serbian names (and the other words) can be improved if strategy "compare codes by similarity" is used. This kind of comparing strategy introduces two – level comparing technique. When one wants to compare two words, he can determine codes for these words (on the first level), and, after that he can calculate similarity between codes. By this way, one can enlarge set of found words. Following tables proves this claim.

TABLE 4. THE RESULTS OF COMPARISON OF WORD ***JANKOVIC*** WITH REPRESENTATIVE LIST OF NAMES USING COMBINATION OF ***NYSIIS*** CODE RESULTING METHOD AND ***JARO – WINKLER*** SIMILARITY METHOD

Found similar words (count)	Minimal similarity level (percent)	Duration (seconds)
2390	70	1.15625
1008	75	1.015625
447	80	0.90625
135	85	0.875
9	90	0.90625
8	95	0.875
1	99	0.828125

TABLE 5. THE RESULTS OF COMPARISON OF WORD ***JANKOVIC*** WITH REPRESENTATIVE LIST OF NAMES USING COMBINATION OF ***NYSIIS*** CODE RESULTING METHOD AND ***LEVENSTEIN*** SIMILARITY METHOD

Found similar words (count)	Minimal similarity level (percent)	Duration (seconds)
134	70	0.8125
12	75	0.796875
7	80	0.78125
7	85	0.828125
1	90	0.828125
1	95	0.8125
1	99	0.8125

As one can notice, the minimal similarity level for using Jaro – Winkler similarity method in combination with NYSIIS coding algorithm is 90% and above (for 95% results is *jankov, jankovic, janković, jankovica, jankovići, junkovica, junković, jankovi*). When using Levenstein instead of Jaro – Winkler similarity metric mentioned threshold is about 75% (for 80% resulting words are *jankovic, janković, jankovica, jankovići, junkovica, junković, benkovac*). Using Levenstein similarity, also, gives to us a better time based result. The time performance of this calculation can be improved if could all words from list of Serbian terms be placed in some kind of Hash – table, which hash – key would be NYSIIS code.

According these tables and previous discussion we can assume that combination NYSIIS + Levenstein gives the most appropriate results for Serbian names comparison.

V. CONCLUSION

This paper presents an overview of well known string matching algorithms (that are generally divided in two groups – similarity and code resulting) and, in the same time, explores their possible application for Serbian and other Slavic names. At this point no solution like this could be find in Serbia. For testing purposes we've used some demo base that contains about 2500 Serbian last names and names of settlements. Different comparing strategies are tested: comparing names by similarity (using Jaro – Winker or Levenstein algorithms), comparing names by their phonetic codes directly (NYSIIS, Metaphone, different SoundEx codes) or calculating similarity between codes in order to enlarge set of results. By all previously presented information we can agree that the most of technique are suitable for desired application. In the further work, we will try to upgrade presented algorithms in order to become more suitable for Serbian language, as well as, enlarge our test base.

REFERENCES

[1] Justin Zobel, Philip Dart, *"Phonetic String Matching: Lesson from Information retrieva"* , In Proc. 19th Inter. Conf. on Research and Development in Information Retrieval (SIGIR'96), pages 166--172, Aug. 1996.

[2] Matthew A. Jaro, *Advances in Record-linkage Methodology a Applied to Matching the 1985 Census of Tampa, Florida*, Journal of the American Statistical Association, 89:414-420.

[3] William E. Winkler, Yves Thibaudeau, *An Application of the Fellegi-Sunter Model of Record Linkage to the 1990 U.S. Decennial Census*, Statistical Research Report Series RR91/09, U.S. Bureau of the Census, Washington, D.C., 1991

[4] Peter Kleiweg, „*Implementation and Visualization of Levenstein Algorithm*", the article taken from url http://www.let.rug.nl/~kleiweg/lev/levenshtein.html

[5] Paul E. Black, "NYSIIS", from Dictionary of Algorithms and Data Structures, Paul E. Black, ed., NIST. http://www.nist.gov/dads/HTML/nysiis.html

[6] *Lawrence Philips Metaphone algorithm*, the article taken from http://aspell.net/metaphone/

[7] *How To: Understanding Classic SoundEx Algorithms*, the article taken from http://www.creativyst.com /Doc/Articles/SoundEx1/SoundEx1.htm

Evolution of the workflow management systems

Krasimira P. Stoilova[1] and Todor A. Stoilov[2]

Abstract – **In this paper a short chronology of evolution of workflow management systems is given. Classifications according to different criteria of software tools supporting modeling and execution functionalities of workflow management systems are presented.**

Keywords – **Automation systems, business processes, web services**

I. Introduction

Workflow Management Systems are a mature technology for automating and controlling business processes [1], [2]. One widely accepted definition of workflow comes from the Workflow Management Coalition [3]: "Workflow is the computerized facilitation or automation of a business process, in whole or part". With the rise of the Web as the major platform for making data and services available for both, humans and applications, a new challenge has become prevalent requiring support of workflows within and crossing organizational boundaries [4], [5].

A general task of the workflow system in the current business activities is the implementation of principles of the automatic control in business systems. The lasts do not consist pure technical components, but they integrate both human and human-computer activities and non-automatic interactions. The paper presents an overview about the evolution of the workflow systems. A critical study is performed, addressing workflow standards and modeling languages. Software products, supporting modeling and execution functionalities of workflow management systems are discussed.

II. Importance of workflow automation

To implement automation in the business processes it is necessary to apply modeling techniques for the non-technical, organizational systems. Over the last decade there has been increasing interest in information systems that are used to control, and/or monitor business processes. Examples of them are Enterprise Resource Planning (ERP) systems, Work Flow Management Systems (WFMS) and Customer Relationship Management (CRM) systems. These systems are implemented to specific business processes. A set of formal languages have been worked out in the context of web services (BPEL4WS, BPML, WSCI, etc). The support of leading firms as IBM, Microsoft, HP and SAP for a language like BPEL4WS (Business Process Execution Language for Web Services), [6] proves that workflows have become important for development. As a result, workflow systems are well addressed in standards: BPEL4WS, XPDL, WfMC [7].

III. The evolution of business process management

In [8] a history review of the workflow technology is given. Till nineties a more fundamental approach is missing. "The aim of workflow management technology is the separation of process logic from application logic in order to enable flexible and highly configurable applications" [9]. In [10] seven fields of importance of workflow management technology are given: office automation, database management, e-mail, document management, software process management, business process modeling. The office automation targets "to reduce the complexity of the user's interface to the (office information) system, control the flow of information, and enhance the overall efficiency of the office" [11]. An overview of the historical development of office automation systems and workflows is given in [9]. The understanding of workflow management has been determined by the terminology of the Workflow Management Coalition [3]. The workflow system consists of a modeling component, functionality for the creation of workflow instances from these workflow models, and functionality for the execution of the workflow instances. The products, which implement workflow technologies, have this functional consistence.

In the sixties information systems were built for small operating system with limited functionality. These systems mainly consisted of particular applications. New software tools added new functionalities like database management. This trend leads to the emphasis from programming to assembling of complex software systems. The coding of individual modules is an old approach now [8]. The challenge now is orchestrating and combining pieces of software [8].

A. Workflow Management Systems

Workflow systems use a variety of languages based on different concepts. Most of the products use a proprietary language rather than independent one. Some workflow systems are based on Petri nets but typically add product specific extensions and restrictions. The differences between the various tools are considerable. The reasons for the lack of consensus of what constitutes a workflow specification are the variety of the business processes. The absence of standard business process modelling concepts is the reason for the diversity in workflow languages. Respectively the comparison of different workflow products looks to be more as a dissemination of products then a critique of workflow language capabilities. An example of model of a collaborative workflow system is given in Fig. 1.

[1]Krasimira P. Stoilova is with the Institute of Computer and Communication Systems – Bulgarian Academy of Sciences, Acad. G. Bonchev str., bl.2, 1113 Sofia, Bulgaria, e-mail: k.stoilova@hsi.iccs.bas.bg

[2]Todor A. Stoilov is with the Institute of Computer and Communication Systems – Bulgarian Academy of Sciences, Acad. G. Bonchev str., bl.2, 1113 Sofia, Bulgaria, e-mail: todor@hsi.iccs.bas.bg

B. *Web Services Composition Languages*

There are two trends in the world of E-business that are creating opportunities to automate business processes across organizations. The first is the technology taking XML-based standards and the Internet as a starting point. The other trend is the need to improve the efficiency of processes from a business perspective. There is a need to utilize the potential of Internet for automating business processes. The goal of web services is to exploit XML technology and to integrate applications that can be invoked over the Web.

Developments of web services composition languages have been mainly driven by software vendors like IBM, Microsoft, Sun, BEA, SAP. This resulted in an abundance of standards with overlapping functionality. The efforts of specialists are directed to narrow the software workflow tools by ignoring standardization proposals that are not using well-established process modeling techniques.

The first standardization effort was the Workflow Management Coalition (WfMC), http://wfmc.org/. The reference model, which is applied, interfaces between a workflow management system and other actors are defined. The standardization efforts try to extend the application area of e-Business solutions.

IV. Suits for workflow software

Workflow models identify: how tasks are structured, who performs them, what their relative order is, how they are synchronized, how information flows support the tasks and how tasks are tracked. Workflows can be modeled using Petri nets [12], [13]. Distinction can be made between "scientific" and "business" workflow models. The “scientific” is mostly concerned with throughput of data through algorithms, applications and services, the business concentrates on scheduling and task executions.

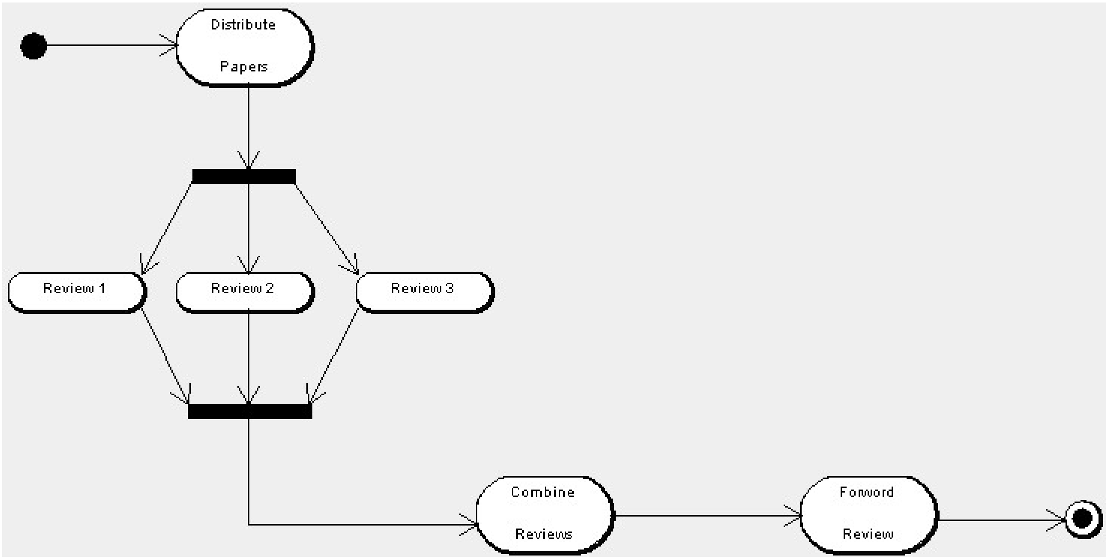


Fig. 1. Collaborative workflow

V. Classifications of software tools for business and scientific workflows

The list of products for modeling and execution of workflow systems is long. The products now have different level of maturity and they are presented as market available an open source software suits. Here are presented classifycations of these software tools:

- Software for business and scientific workflows.
- Tools according to their software language design.
- Tools according to their supported standard.
- Open Source and Commercial tools.

The scientific Workflows systems receive wide acceptance particularly in bioinformatics in 2000s. They are successful to perform interconnections between tools, to handle different data formats and volumes http://en.wikipedia.org/wiki/ Workflow.

The business workflow represents the set of business tasks in the organizations, perform the time scheduling between the processes, coordinate software applications, and manage the paper flow documentation. Workflow are “systems that help organization to specify, execute, monitor, and coordinate the works within a distributed office” http://en.wikipedia.org/wiki/Workflow. The Workflow diagrams use standardized graphical notations to describe workflow.

The software tools for business and scientific workflow systems are presented below:

- Tools for business workflow

@enterprise	Aegeanet System
Agentflow	Amazonas Worflow
Bonita	Captaris
Business Process Management	Business Integration Engine
COLOSA	CoMo-Kit
EmeriCon	Enhydra Shark
EventStudio	FlowRunner
infoRouter	iKE
Ils/process	IngTech Corporation
Intella	Interstage BPM
MyControl Workflow Server	K2.net Enterprise Workflow
Jboss	OpenFlow
OpenSymphony	OpenWFE
OracleBPEL Process Manager.	Skelta Workflow.NET
Ring Pro	PL/FLOW
VivTek	W4
Web and Flo Kontinuum	WebSphere MQ Workflow
YAWL	

- Tools for scientific workflow

Taverna	Kepler
GridNexus	SPA
Triana	Jopera

VI. Tools according to the software language design

A review of an open source workflow project is www.manageability.org/blog/stuff/workflow_in_java/view. A classification is performed according to program language, used for the design of the software tool: Java based; other Java; non-Java; for specific application servers or environment.

A. Open Source Workflow Engines Written in Java

The review list consists:

ActiveBPEL	Antflow
Apache Agila	Beexee
Bigbross Bossa	Codehaus Workflow
con:cern	Dalma
Enhydra Shark	Freefluo
jBpm	Jfolder
MidOffice BPEL Engine	Micro-Workflow
ObjectWeb Bonita	OFBiz
OpenWFE	OpenSymphony OSWorkflow
Pi Calculus for SOA	PXE
Syrup	Taverna
Twister	wfmOpen
Xflow2	YAWL
Zbuilder	Zebra

An extended description of each product is available from http://www.gripopprocessen.nl/index.php?id=35&no_cache=1. Currently available tools are:

- Generic J2EE

Apache Agila	Bonita
Imixs	Jboss jBpm
Jfolder	Open Business Engine
wfmOpen	XFLow

- Other Java based products

Bossa	Enhydra Shark
OpenWFE	Syrup
Werkflow	Twister
Open Symphony Workflow	Yet Another Workflow Language engine
Zebra	

- Other: Non-Java tools

wftk (workflow toolkit)	

- For specific application servers or environment

ActionWorks	BizFlow
Galaxia Tikiwiki	OpenFlow
Grid-based computing workflow – Taverna	Open for Business (OfBiz)
WebAsyst Issue Tracking	WebAsyst Workflow Management Software
webMethods Business Process Management	

VII. Software tools according to their supported standard

- XPDL based : http://www.wfmc.org/standards/docs/TC-1025_10_xpdl_102502.pdf

Aspose.Workflow	Open Business Engine
Agentflow	Aspose Workflow
Enhydra JaWE	ObjectWeb BONITA
Enhydra Shark	Fuego BPM
Interstage BPM	Newgen
OFBiz Workflow Engine	Vignette Process Workflow Modeler
wfmOpen	YAWL

- BPEL based: http://www.oasis-open.org/committees/download.php/16024/wsbpel-specification-draft-Dec-22-2005.htm

iGrafx BPEL	Apache Agila
ActiveBPEL Engine	Bexee
ActiveWebflow Standard	Cape Clear
Biztalk Server	IBM BPWS4J
Oracle BPEL Process Manager	JOpera
PXE	MidOffice
Parasoft BPEL Maestro	Twister
SAP NetWeaver Exchange Infrastructure	

• BPMN based: http://www.bpmi.org

Borland Together Desidner	ITpearls Process Modeler
IntalioDesigner	AXway Process Manager™
Fujitsu: Interstage Business Process Manager 7.1	Kaisha-Tec: ActiveModeler
Lanner: Witness™	Mega International: Mega Suite™

VIII. OPEN SOURCE AND COMMERCIAL TOOLS

• Open Source Workflow Tools

Additional information can be found in http://java-source.net/open-source/workflow-engines

Twister	jBPM
Enhydra Shark	OpenSymphony OSWorkflow
con:cern	Codehaus Werkflow
ObjectWeb Bonita	Bigbross Bossa
Open Business Engine	The Open for Business Workflow Engine
OpenWFE	WfMOpen
XFlow	Jfolder
Taverna	Freefluo
Micro-Flow	Jflower
YAWL	Syrup
PXE	ActiveBPEL
Antflow	Swish

• Commercial Workflow Tools

Active Endpoints ActiveWebflow Server	ActiveWebflow Designer
ADONIS	Biztalk Server
Cape Clear Orchestrator	Digité Process Composer
Fiorano SOA Platform	FiveSight PXE
FuegoBPM	IBM BPWS4J
IBM WebSphere Business Integration Server Foundation	OpenLink Virtuoso Universal Server
OpenStorm ChoreoServer	Oracle BPEL Process Manager
Parasoft BPEL Maestro	PolarLake Integration Suite
SAP NetWeaver Exchange Infrastructure	SeeBeyond eInsight BPM

This work is partly supported by EC Project VISP, IST-2004-027178

IX. CONCLUSION

This survey of products, available for workflow modeling and execution describes a huge amount of products. It is difficult for a common user to make a choice of appropriate software tool, which have to be applied in user applications. Thus the problem of evaluation and assessing this class of software becomes quite important. It is firmly related to the functionalities of the products to support common standards, related to the workflow management. But the implementation of the workflow management systems is assumed to be the most prospective domain where automation and information technology overlap in business management applications.

REFERENCES

[1] G. Kappel, S. Rausch-Schott, W. Retschitzegger. "A Framework for Workflow Management Systems Based on Objects, Rules and Roles", *ACM Computing Surveys Electronic Symposium on Object-Oriented Application Frameworks*, 2000.

[2] F. Leymann and D. Roller, *Production workflow: concepts and techniques,* Prentice-Hall, 2000.

[3] Glossary, Terminology and Glossary, 3rd Edition. Document No WFMC-TC-1011. Workflow Management Coalition. Winchester, 1999.

[4] W. Aalst, "Process-Oriented Architectures for Electronic Commerce and Interorganizational Workflow", *Inf. Systems*, Vol. 24/8, 1999.

[5] U. Dayal, M. Hsu, R. Ladin, *Business Process Coordination - State of the Art, Trends, and Open Issues*, Proc. of the 27th VLDB Conference, 2001.

[6] T. Andrews, F. Curbera, H. Dholakia, Y. Goland, J. Klein, F. Leymann, K. Liu, D. Roller, D. Smith, S. Thatte, I. Trickovic, S. Weerawarana, Business Process Execution Language for Web Services Version 1.1, 2003. Technical report, http://xml.coverpages.org/ BPELv11-May052003Final.pdf

[7] WfMC, Workflow Process Definition Interface –XML Process Definition Language. Technical Report Document Number WFMC-TC-1025, Workflow Management Coalition, 2002, http://www.wfmc.org/standards/docs.htm

[8] W. M. P. van der Aalst, A. H. M. ter Hofstede, B. Kiepuszewski, and A. P. Barros, "Workflow patterns". *Distributed and Parallel Databases*, 14(1), 5-51, 2003.

[9] Michael zur Muehlen, *Workflow-based Process Controlling. Foundation, Design, and Application of Workflow-driven Process Information Systems*. Logos Verlag Berlin, 2004, ISBN 3-8325-0388-9.

[10] S. Jablonski, C. Bussler, *Workflow Management: Modeling Concepts, Architecture and Implementation,* Int. Thom. Comp. Press, London, 1996.

[11] C. A. Ellis, G. J. Nutt, "Office Information Systems and Computer Science", *ACM Computing Surveys*, 12 (1980) 1, pp. 27-60, 1980.

[12] J. L. Peterson, *Petri Net Theory and the Modeling of Systems*, Prentice Hall. ISBN 0136619835, p.288, 1981.

[13] W. Reisig, *A Primer in Petri Net Design*, Springer-Verlag, ISBN 3-540-52044-9, 1992.

New genetic selection strategy for Genetic Parallel Algorithm

Milena N. Karova[1], Vassil J. Smarkov[2]

Abstract – **However, the major disadvantage of GA is that the algorithm uses a tremendous amount of processing time. This paper proposes a method that can reduce the processing time by using Parallel Genetic Algorithm (PGA). The new genetic selection strategies are presented.**

Keywords – **Parallel Genetic Algorithm, selection, population, crossover, mutation, mutation rate, crossover rate**

I. INTRODUCTION

The major disadvantage of GA is that the algorithm uses a very large amount of processing time. A parallel algorithm approach can be applied to the classical GA for reducing processing time. Parallel Genetic Algorithm (PGA) can be classified into three different models: Master-slave PGA, Coarse-grained PGA and Fine-grained PGA.

Master-Slave parallelization is machine dependent [1]. A shared memory multiprocessor computer is not easy available when compared with a clustered computer, which consists of network of workstations. This paper present selection schemes of Grained PGA. The aim is apply Genetic Algorithm to find an optimal solution, which satisfies nontrivial constraints of timetable problem.

II. GENETIC REPRESENTATION

A. Chromosome

Our work use alternative chromosome representation where each position in a chromosome represents the period which the classes take place.

B. Genetic Operators

The genetic operators are defined as follows:

Selection: The selection Scheme use Tournament Selection, random Walk Selection and Spatially Orientated selection.

Crossover: The algorithm proposes 3 variants of crossover: uniform crossover, one point and two point crossover.

Mutation: mutation operator swaps two values at random mutation positions. It is direct mutation (Fig. 1).

Genetic parameters: Table 1 shows the algorithm parameters. The set of constraints are generated as an input file. The node number varies from 1 to 9. The efficiency of Parallel GA is measured in terms of computational effort, defined as the number of individuals that must be processed to get solution, compared to the serial algorithm. The optimal solution is defined like as solution that satisfied all hard constraints. The good solution is defined like as satisfied 95% of all hard constraints.

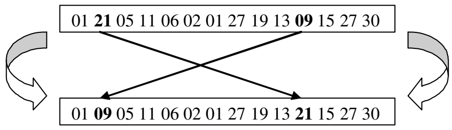


Fig. 1. Direct Mutation

Table 1

Population Size	16 (32)
Crossover probability	90%
Mutation probability	0,4%
Migration size	10%
Maximum generation	500

Fitness function: The fitness function (1) is a sum of penalties.

$$f(ch) = \sum_{i=1}^{events} \sum_{j \in C} WjPj \,, \tag{1}$$

where $f(ch)$ is fitness function, events is length of chromosome (number of lectures), C is set of constraints, Wj is weight of constraints j and Pj is function that returns 1 (violated constraint j) and 0 (not violated constraint j). The target value is 0.

III. PARALLEL GENETIC ALGORITHM (PGA)

GA has the ability to be parallelized because an algorithm work with a set of population, not only an individual. The evolution of an individual is separated from each other. The concept of PGA is to divide the task of the classical GA and distributes on different processors. [2, 3] There are three main types of PGA:

Master-Slave PGA: This model uses a single global population and the fitness evaluation is done on different processors. Furthermore, genetic operations may also be done in parallel. The nature of GA is not changed because an algorithm still works with the whole population. A global population is suitable for a shared memory computer.

[1]Milena N. Karova is of the Department of Computer Science, Studentska 1, Technical University Varna Email: mkarova@ieee.bg

[2]Vassil J. Smarkov is with the Department of Computer Science, Studentska 1, Technical University Varna Email: smarkov@ieee.bg

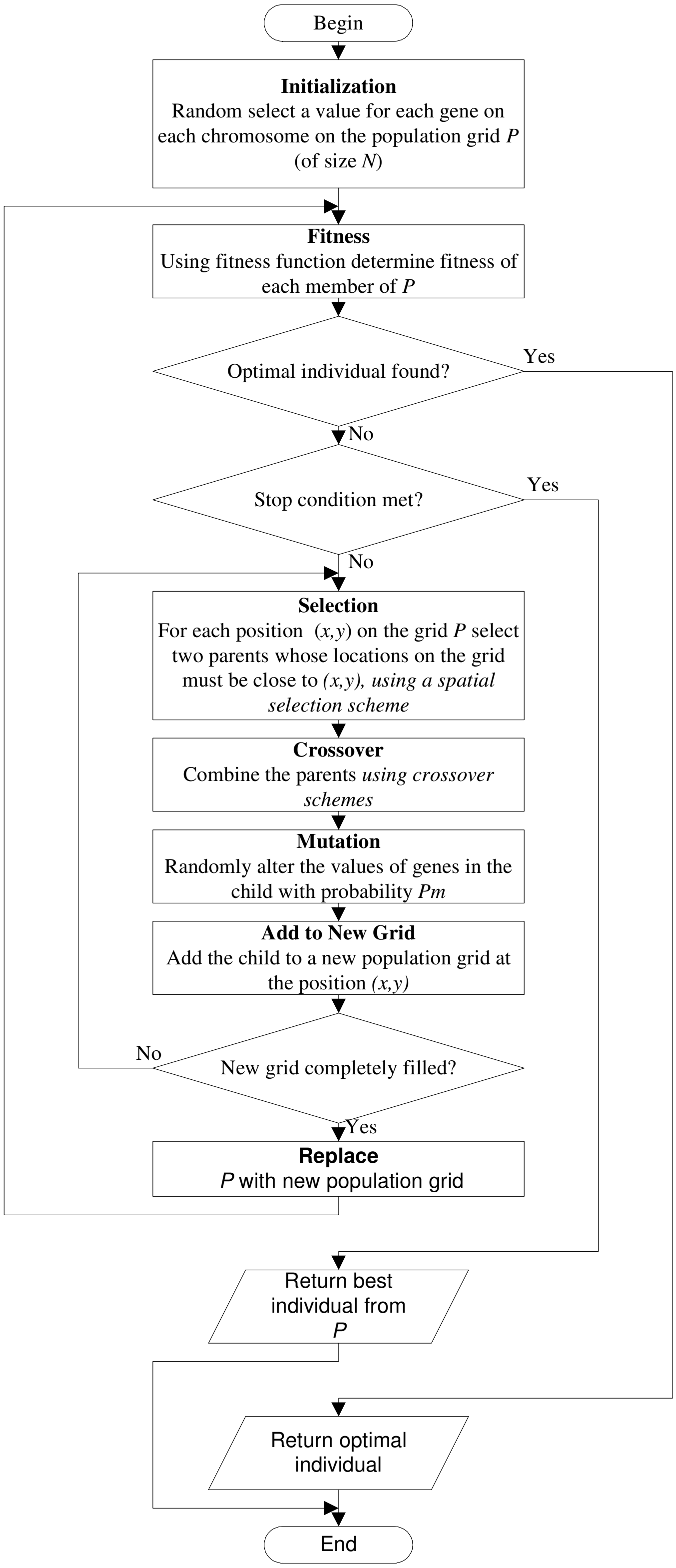


Fig. 2. PGA

Coarse-grained PGA: The population is divided into a few large subpopulations. Each of these subpopulations are maintained by different processors and some selected individuals are exchangeable via a migration operator. The model is known as Island model or distributed PGA and subpopulation called deme. Island model [4] is a popular and effective parallel genetic algorithm and also reduces probability of premature convergence – finding the local instead of the global optimum.

Fine-grained PGA: The population is separated into a large number of very small subpopulations, which are maintained by different processors. The subpopulation may be only an individual. This model is suitable for massively parallel architectures – machines consisting of a huge number of basic processors and connected with a specific high speed topology. The computer structure limits an interaction between individuals. This model is machine dependent like Master-Slave PGA.

The flowchart of PGA is shown at Figure 2. The population is thought of as occupying a two dimensional grid, each individual occupying its own position on the grid. The grid is taken to be toroidal, which is to say it wraps round itself so that the position to the right of a cell in the last column of the grid is a cell in the first column and the same row.

IV. SELECTION SCHEMES

The question of how parents are selected from the population is fundamental to the operation of genetic algorithms. As with many other aspects of GAs there is no hard and fast rule regarding the choice of selection scheme. Selection schemes must strike a balance between the stochasticity needed to maintain diversity ('exploration' in Holland's analogy) and the determinism needed to propagate fit schemata ('exploitation'). Some methods are of particular use in granular PGAs as they make use of the spatial orientation of individuals in the population.

A. Random Walk Selection

Random walk is a local selection method for use in granular PGAs. A random path of length *k* is mapped out starting from the position on the grid for which a parent is needed (Fig. 3). Of the k individuals encountered along this path the fittest is chosen for breeding. The stochasticity of random walk is provided by the random choice of direction:

right: `currentIndex-=(int)sqrt(POPULATION)-1` or `currentIndex+=1`

left: `currentIndex+=(int)sqrt(POPULATION)-1` or `currentIndex-=1`

up: `currentIndex+=POPULATION-(int)sqrt(POPULATION)` or `currentIndex - = (int)sqrt(POPULATION)`

down: `currentIndex-=POPULATION-(int)sqrt(POPULATION)` or `currentIndex+=(int)sqrt(POPULATION)`

`POPULATION` is the size of population.

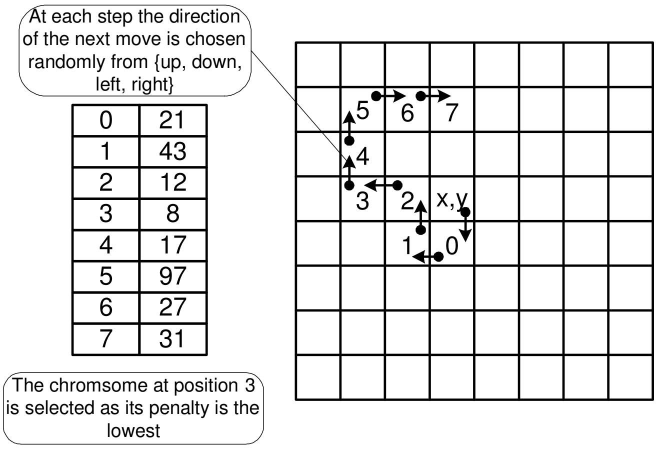


Fig. 3. Random Walk selection

B. Spatially selection

Spatially selection is orientated selection scheme for granular PGAs. Here (Fig. 4) the idea is to move a square 'window' or 'mask' over the grid such that the position for which parents are sought lies at its centre. The fittest of the chromosomes located at positions within the window is chosen as one parent and the chromosome at the centre of the window as the other.

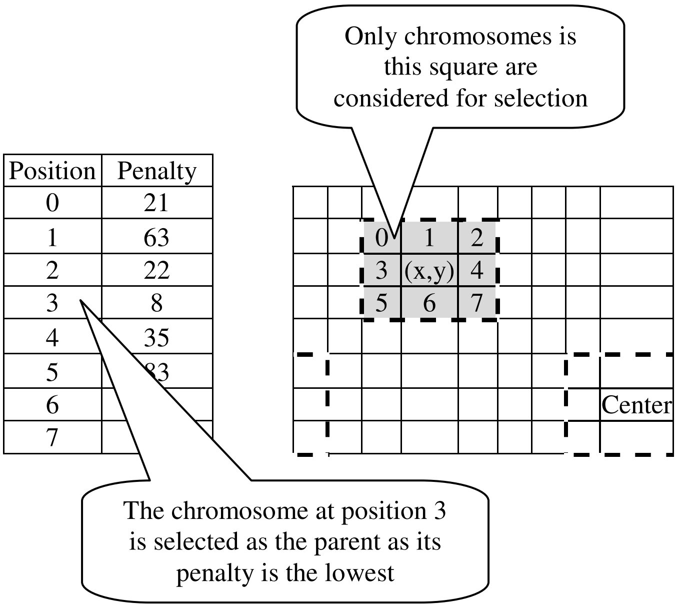


Fig. 4. Spatially Selection

Compared to Random Walk Spatially selection is a more explicitly spatially orientated and local selection scheme. Whereas Random walk may sometimes end up choosing as a parent an individual located at some distance (for example if at each iteration the same direction is chosen) Spatially selection restricts the choice of parents to within a clearly defined area. Spatially selection can implemented with a 3x3 window. The intention was to contrast a highly localised spatial selection method with a less localised one (Random walk).

The comparison of two selection schemes Tournament Selection and Random walk Selection is shown at Figure 5. Evolution process of GA using Random walk selection convergences previously using Tournament selection.

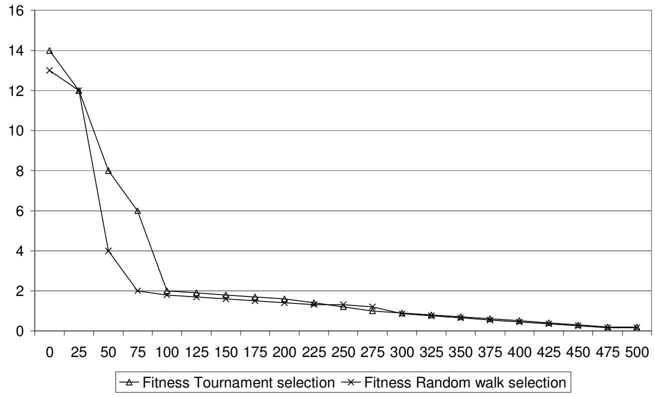


Fig. 5. Comparison of two Selection schemes

V. CONCLUSION

A little effort is needed to converting classical GA to grained PGA [5]. Each processor performs simple GA and periodically exchanges some population by migration operators. Using Random walk selection and Spatially selection provide for fast optimal individual find. Granular PGAs are able to evolve sub-populations in relative isolation possibly leading to improved overall performance. They can be usefully implemented on single processor or serial hardware. The possible efficiencies they offer are not solely related to implementation on parallel hardware.

A great deal of further work will be needed to better understand and exploit Spatially selection with different mutation schemes in evolutionary timetabling. Of immediate interest would be a study of transitions in convergence rate over the course of search.

REFERENCES

[1] Corcoran A.L., A Parallel Island Model Genetic Algorithm for the Multiprocessor Scheduling Problem, Proceeding of the 1994 ACM/SIGAPP Symposium on Applied Computing, ACM Press , March 7-8, 1994, pp.483-487.

[2] Nowostawski M., Poli R., Parallel genetic algorithm taxonomy, University of Otago, New Zeland, 1999.

[3] Ross P., Ballinger G., PGA-Parallel Genetic Algorithm Tested, Department of Artificial Intelligence, University of Edinburgh, 1993

[4] Vertanen K., Genetic adventures in parallel: towards a good island model under PVM, Oregon state university, 1997

[5] Gong Y., Nakamura M., Tamaki S., Parallel genetic algorithms on line topology of heterogeneous computing resources, GECCO'05, Washington, DC, 2005

Intelligent Discovery and Retrieval of Geoinformation using Semantic Mediation and Ontologies

Aleksandar Stanimirović[1], Leonid Stoimenov[2], Slobodanka Đorđević-Kajan[3]

Abstract-**In this paper we proposed architecture, based on modified hybrid ontology approach and semantic mediation, for solving the problems of semantic heterogeneity. Process of semantic mediation is based on ontologies that provide abstract description of the content of information sources. It also uses ontologies to provide semantic consistency and geodata exchange with terminology translation. Proposed architecture provides means for building intelligent systems that can integrate a number of different information sources.**

Keywords-**semantic heterogeneity, ontologies, semantic mediation, geoinformation retrieval, geoinformation discovery**

I. INTRODUCTION

In many areas of life, geographic information is key factor in planning and decision-making. In such environment, accurate and simple methods for identification and retrieval of, often distributed and heterogeneous, geodata is very crucial. Interoperability of GIS (Geographic Information System) is a possible solution for these problems.

Interoperability of information systems relies on bases of agreement that describe what is shared among information sources. Interoperability means openness in the software industry, because open publication of internal data structures allows GIS users to build applications that integrate software components from different developers. Interoperability also means the ability to exchange data freely between systems, because each system would have knowledge of other systems formats [1].

Nowadays, providing a user with huge amount of data from different sources across networks or inter-networks in a short amount of time is done by interoperable systems. Users of spatial data expect information systems to help them to search and process the data as well as supporting them with necessary information and knowledge about the data – i.e., metadata. They also expect to have homogenous interfaces for managing, modeling and processing the data from different sources. Though interoperability has to overcome complexity of conversion and integration process, there is a long way from data transfer and data format exchange (or conversion) to system interoperability. Interoperability issues not only refer to different structures and models of data sets, but also to the different methods and operations applying to the data.

In the domain of GIS interoperability, differences in data sources, disciplines, tools and repositories can cause heterogeneity. Interoperability helps to reuse geodata and avoid waste of our assets. It has the potential to offer a prompt reaction to obtain suitable data set hen dealing with geographic information analysis (such as natural disaster management). In order to achieve the sharing and exchange of existing data between different departments, semantic heterogeneity between data from different component systems must be taken into account. Without knowing the semantics of data (i.e. their meaning), it is impossible to integrate and understand data appropriately. Unfortunately, there are only few integration approaches that consider the semantics of data.

In this paper we focus on our approach for discovering and interpretation of geographic information based on ontologies. The rest of paper is structured as follows. In section II we discuss heterogeneity problems during discovery and retrieval of geographic information. Section III describes GeoNis generic architecture for interoperability and the role of semantic mediators in it. Section IV describes our architecture for semantic mediation based on ontology approach.

II. RELATED WORK

Geographic data set integration is the process of establishing relationships between corresponding object instances in different, autonomously produced, geographic data sets of a certain domain. The purpose of geographic data set integration is to share information between different geographic information sources. Geographic data set integration gets more and more attention nowadays since the digitizing of traditional map series has ended. In these map series, corresponding object instances were only linked implicitly by a common spatial reference system. In order to make these relationships explicit geo-science researchers and computer scientists have developed various strategies.

One important initiative to achieve GIS interoperability is the OpenGIS Consortium [2]. This is an association looking to define a set of requirements, standards, and specifications that will support GIS interoperability. The objective is technology that will enable an application developer to use any geodata and any geoprocessing function or process available on ‘the net’ within a single environment and a single workflow. But, data standardization cannot solve the whole problem. The interoperability problem would go away if every system always uses the same data model to represent the same information (identical names, structure, and representations). OpenGIS standards will only partially solve this problem.

Mediator-based system is important for spatial data interoperability architecture [3]. The 3-level architecture of mediator-based systems is constructed from an application layer, and large number of relatively autonomous information sources (heterogeneous data sources with wrappers), communicating with each other over a standard protocol [4]. A wrapper is a program that is specific to every data source [5]. Wrapper extracts a set of records from source file and

[1]Aleksandar Stanimirović is with the Faculty of Electronic Engineering, Beogradska 14, 18000 Serbian and Montenegro, E-mail: alex@elfak.ni.ac.yu
[2]Leonid Stoimenov is with the Faculty of Electronic Engineering, Beogradska 14, 18000 Serbian and Montenegro, E-mail: leni@elfak.ni.ac.yu
[3]Slobodanka Đorđević-Kajan is with the Faculty of Electronic Engineering, Beogradska 14, 18000 Serbian and Montenegro, E-mail: sdjordjevic@elfak.ni.ac.yu

performs translation in the data source format. The most important advantage is the fact that data integration system allows users to focus on specifying what they want rather than thinking about how to obtain the answers. As a result, it frees them of combining data from multiple sources, interacting with each source and finding the relevant sources. Nowadays, mediation concept is a part of the ARPA I3 (Intelligent Information Integration) reference architecture [6].

The uses of ontologies as semantic translators is approach that can possible overcome the problem of naming conflict and semantic heterogeneity. Research on ontology is becoming increasingly widespread in the computer science community, and its importance is being recognized in a multiplicity of research fields and application areas, including knowledge engineering, database design and integration, information retrieval and extraction, and information systems [7].

During past several years many different solutions for problems of semantic heterogeneity have been proposed [8, 9, 10]. In this paper we proposed our approach for discovering and retrieving geodata based on ontologies and semantic mediation.

III. GeoNis Framework for semantic interoperability

GeoNis is a generalized framework for GIS interoperability. It provides infrastructure for data interchange in the local community environment [3, 11]. The basic architecture of GeoNis framework is shown on Fig. 1.

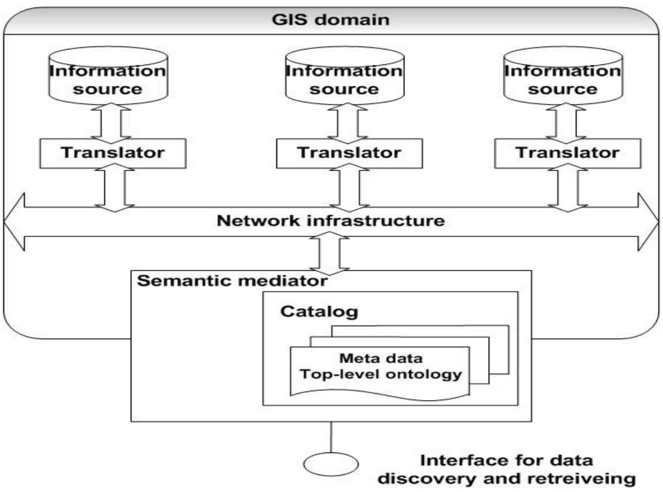


Fig. 1. GeoNis – framework for semantic interoperability [12].

Generic architecture of GeoNis recognizes several different components that have important role in geoinformation discovering and retrieving process [13]:

- *Information source* – in each node of GeoNis framework there exist GIS application and corresponding (spatial and non-spatial) data sources. Data in these local data sources are accessible according to user privileges.
- *Translator* - component that translates information flow between information source and GeoNis system.
- *Semantic mediator* - requests for specific data set are forward through.
- *Catalog* – maintains metadata and all shared/common geographic data as addition to domain oriented GIS applications.

When user (human or GIS application) issues a request, system have to decide which source (or sources) is able to deliver the requested information. The simplest way to do that is to query each information source that is registered. If the request succeeds, the results are delivered to the requester. In the sense of intelligent information discovery and retrieval such behavior is not aspired. We want the system to be aware which sources are worth of querying and which are not. This goal can be reached by providing an abstract description of the content of an information sources using ontologies.

Architecture that we propose in this paper (Fig. 2) solves the problems of geoinformation discovery in distributed and heterogeneous environment. Proposed architecture is based on generic Semantic Mediator architecture given in [10]. Proposed architecture must provide techniques for realization of domain oriented mediator and translator chaining in a specific domain. This architecture also provides means that support interpretation of retrieved geoinformation. Ontologies have a central role in resolving semantic conflicts.

All access to geoinformation in local community environment goes through *Web Feature Service* defined by proposed OGC technology standards [2]. We enhanced this interface with additional functionality in order to support user profiles and privileges. This interface is implemented by *Semantic Mediator* component. This component acts as an access point for a number of independent geoinformation sources and allows integration of their information bridging over the semantic differences among them. *Semantic Mediator* mediator enables users to access multiple information systems as though they were a single system with a uniform way to retrieve information and perform computations. It accepts high-level requests from users and automatically translates them into a series of lower-level requests for different GICs. Results of this low-level request are then combined into a result for a high-level request. In order to accomplish this task Semantic Manager must use metadata information, provided by *Ontotology Manager* services, in order to discover every concept (top-level or local) in environment that can be targeted with user request.

Ontology Manager is a component that provides access to the shared metadata that resides on the common server. This component implements interfaces that provide means for discovery, access and management of metadata to the rest of the community. Interface *Publish service* allows GIC nodes to register their data and services and to create relations between their local ontologies and top-level ontologies. Interface *Query service provides* functionality for locating and accessing requested metadata.

Intelligent UI component dynamically generates web-based user interface according to user privileges and user profiles. This interface allows users to query local community for metadata and geodata according to their privileges.

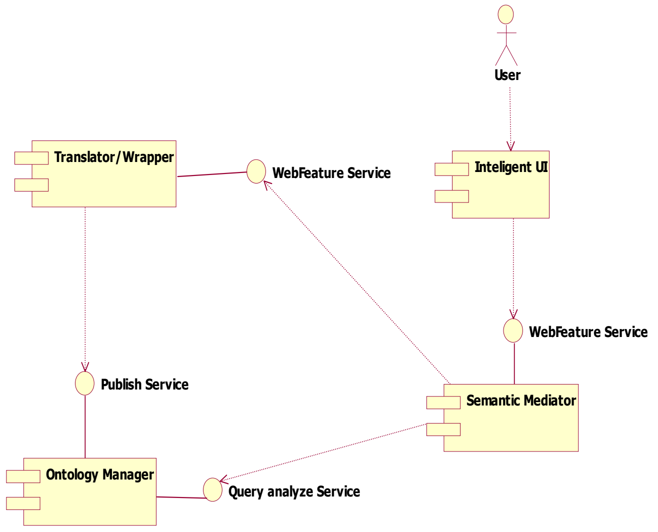


Fig. 2. Components and interfaces for semantic mediation

IV. SEMANTIC MEDIATION

In GeoNis Semantic Mediator we propose a semantic based integration approach that uses multiple ontologies, instead of an integrated view. In this context, ontologies are virtually linked by inter-ontology relationships, which are then used to indirectly support query processing. Our Semantic mediator uses hybrid ontology approach (Fig. 3). Meaning of the terminology of each community is specified in the local ontology. Semantic Mediator provides a methodology and software support for semantic mismatches (conflicts) resolving between terminologies. This methodology uses the ontology mappings between each community terminologies and a top-level ontology or the common data model (reference ontology).

All meta-data information (top-level ontologies, inter-ontology relations, ontology mapping rules and additional metadata information) in local community environment are hosted by Shared server. This metadata describes the properties of geo-data sources that can be queried through mediators and translators. Using meta-data from Shared server, as respond to a user query, Semantic Mediator can return sets of geo-objects (features) from different data sources. Contained public information can be classified as follows:

- *User metadata*:
  - *User privileges* – user rights for accessing data in local GIC nodes or in common GeoNis server.
  - *User profiles* – description of customized, intelligent, web-based user interface that is dynamically generated whenever user access data in local community environment.
- *Virtual organization metadata (Ginis Catalog)* – local community structure description. Every new GIC who wants to participate in exchanging data must register with common GIS server in order to allow access to his public available data and local ontology. After that, users from registered GIC have access to all available data from other public GIC databases and access to shared data owned by shared GIS server (with possible given rights for access).
- *Top-level ontologies* – domain shared vocabulary and description of available data sets
- *Ontology mappings* – describes mapping between top-level and local (application) ontologies.

When user wants to retrieve some information from local community environment, first, he has to logon on the system. According to his user profile and privileges *Intelligent UI* component builds appropriate Web based interface. This interface also includes tools that can help user to build query using top-level concepts obtained from local community environment.

After logon procedure, user can query environment for geoinformation that he is interested in. In order to provide requested data, *Semantic Mediator* must analyze user query and discover every concept (feature) in environment that fits to user request.

This analysis is done by *Ontology Manager*. User query is first matched against top-level ontologies (description of the available datasets in information sources). In a highly heterogeneous system it might happen that the representations of queries are incompatible with some top-level ontologies, because they use a different terminology or different structuring principles for the domain. In such case rules for conflict resolving must be used (ontology mappings).

Ontology Manager also has to deal with distributed geospatial datasets. That means that user request can hit several information sources at the same time and a result data set is not located in one data source. Columns that forms result data set can be distributed in great number of data sources. Also, result data is not necessary in relational databases. Spatial data can be stored in great number of heterogeneous data sources. *Ontology Manager* builds *FeatureSchema*, metadata information that contain description and location of every concept (feature) in environment that can be treated as a result of a user query [14]. Result of user query is treated as a collection of feature tables and spatial reference systems definitions. Every feature table description contains basic table meta-data information and a collection of different connections. Feature table connections describe information sources that have to be used to build up feature table data. These connections also contains information of means for retrieving and rules for integrating data from different sources. Using this schema information, *Semantic Mediator* retrieves data from local GIC nodes and builds result dataset that is forwarded to the user, as a response to his query, in the form of a GML document.

V. CONCLUSION

In this paper we introduced generic architecture and framework for ontology based discovery and retrieval of geoinformation. This architecture provides us with means for building intelligent system that can integrate a number of different information sources bridging over syntactic and semantic differences between them. Intelligence of this system is based on techniques for semantic mediation.

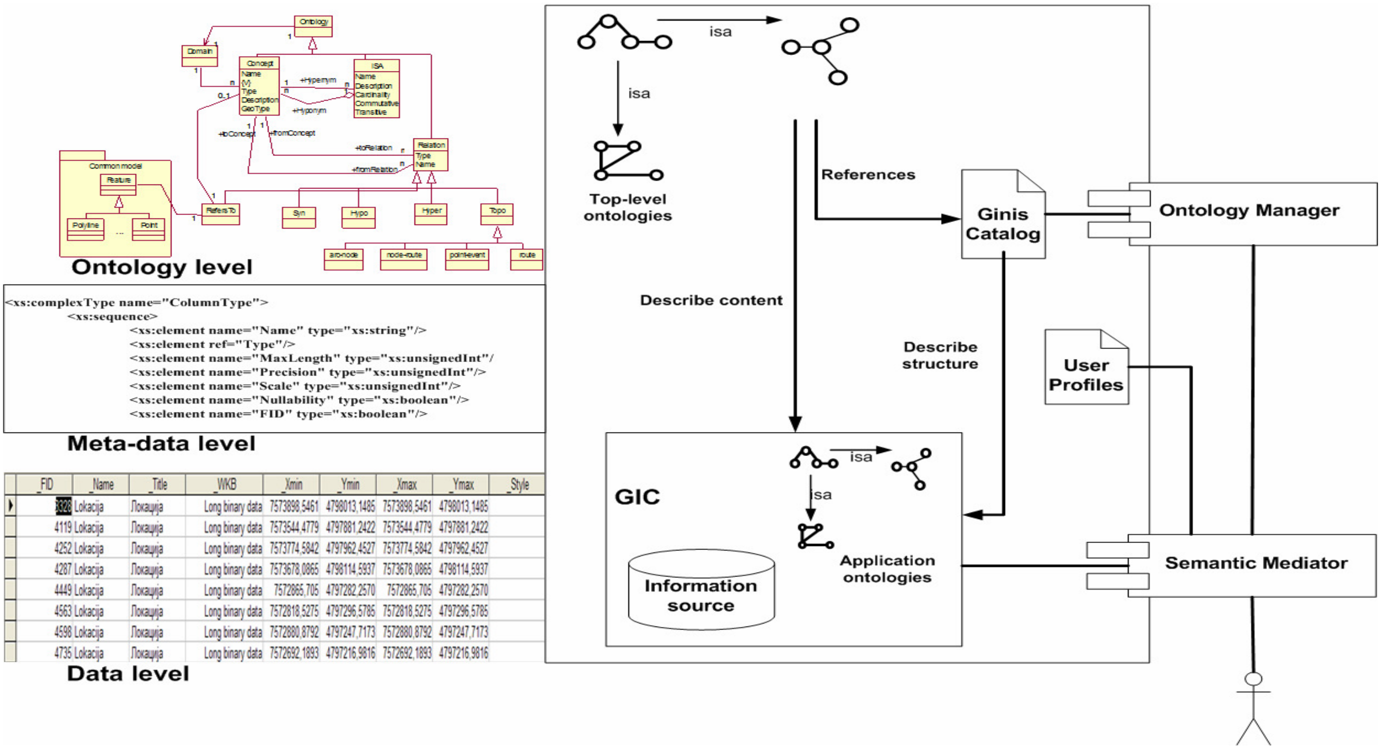


Fig. 3. Hybrid ontology approach

Semantic Mediator acts as an access point for a federation of independent geoinformation sources. When processing user requests, Semantic Mediator must be aware, in advance, which information sources contain data that fulfill user requests? This goal can be reached by providing an abstract description of the content of an information sources using ontologies. It also uses domain ontologies in order to provide semantic consistency for data from that domain and geodata exchange with terminology translation.

Proposed architecture is a component-based with strongly defined interfaces between components. In this way architecture can be easily extended in various directions. Introduced architecture solves the problems of semantic heterogeneity in a single GIC environment domain.

REFERENCES

[1] Buehler R., and McKee L., 1998, *The Open GIS Guide*, Third Edition, OpenGIS Consortium, Inc,http://www.OpenGIS.org/techno/specs.htm , 1998

[2] Open Geospatial Consortium, http://www.opengeospatial.org , 2005

[3] Stoimenov L., Djordjević-Kajan S., Stojanović D., 2000, "Integration of GIS Data Sources over the Internet Using Mediator and Wrapper Technology", *MELECON 2000, 10th Mediterranean Electrotechnical Conference*, May 29-31 2000, Cyprus, Proceedings Vol. 1, pp 334-336.

[4] Wiederhold G., "Interoperation, Mediation and Ontologies", *Int. Symp., on Fifth Generation Computer Systems*, Tokyo, Japan, 1994.

[5] Stoimenov L., Mitrovic A.,Djordjevic-Kajan S., Mitrovic D., "Bridging objects and relations: a mediator for an OO front-end to RDBMSs", *Information and Software Technology*, Elsevier, 1999, Vol 41, No. 2, pp. 59-68.

[6] ARPA I3 - *Reference Architecture for the Intelligent Integration of Information*, Prepared by the Program on Intelligence Integration of Information (I3), version 1.0.1, May 19, 1995.

[7] Guarino N., "Formal Ontology in Information Systems", *In N. Guarino (ed.) Formal Ontology in Information Systems. Proceedings of FOIS'98*, Trento, Italy, IOS Press, Amsterdam, 1998, pp.3-15.

[8] Bernard L., Einspanier U., Haubrock S., Hübner S., Kuhn W., Lessing R., LutzM., Visser U., 2003, "Ontologies for Intelligent Search and Semantic Translation in Spatial Data Infrastructures", *Photogram-metrie - Fernerkundung - Geoinformation*, 2003, pp. 451-462.

[9] Klien E., Einspanier U., Lutz M., Hübner S., 2004, "An Architecture for Ontology-Based Discovery and Retrieval of Geographic Information", *Proceedings of the 7th Agile Conference on Geographic Information Science*, Heraklion, Crete, 29 April-1 May, 2004

[10] Stoimenov L., Stanimirović A., Đorđević-Kajan S., "Semantic Interoperability using multiple ontologies", *Proceedings of 8th AGILE Conference on GIScience,* Estoril, Portugal, May 26-28, 2005, pp.261-269.

[11] Stoimenov L., Djordjević-Kajan S., 2002, "Framework for semantic GIS interoperability", *Journal FACTA Universitatis, Series Mathematics and Informatics*, 17(2002), pp.107-124.

[12] Stoimenov L., Đorđević-Kajan S., 2003, "Realization of GIS Semantic Interoperability in Local Community Environment", *Proceedings of 6th AGILE conference on Geographic Information Science: "The Science behind the Infrastructure"*, ISBN 2-88074-541-1, AGILE 2003, Lion, France, April 20-23.2003. pp.73-80

[13] Stoimenov L., Stanimirović A., Đorđević-Kajan S., "Development of GIS Interoperability Infrastructure in Local Community Environment", *From Pharaohs to Geoinformatics, FIG Working Week 2005 and GSDI-8 Cairo*, Egypt April 16-21, 2005, TS41.2.

[14] Stanimirovic A., Stoimenov L., Đorđević-Kajan S., "Geodata integration in distributed environment using OLE DB data provider", *ICEST 2004*, Bitola, Jun 2004, pp. 329-332

Evolutionary Theories and Genetic Algorithms

Hristo I. Toshev[1], Stefan L. Koynov[2] and Chavdar D. Korsemov[3]

Abstract: - **Some of the basic theories of evolution are analyzed in this paper and the possibilities for different applications of their ideas are introduced. Of special interest are the characteristics and the peculiarities of genetic algorithms (GA) that distinguish them from traditional optimization methods and procedures for searching, the stages of preliminary preparation for solving practical problems by GA included.**

Keywords – **genetic algorithms, evolution, selection, mutation, optimization**

I. INTRODUCTION

The word *evolution* (*evolutio* in Latin) means *change*, *development* (slow and gradual). Evolution is usually the opposite of revolution which concerns fast and vast changes that sometimes lead to unpredictable results. At the end of XVIII century Bonnet, a Swiss scholar, uses the term evolution for the first time in biology in the context of a slow and gradual qualitative and quantitative change of the object when each state of the object must be at a higher level of development and organization compared to the previous one.

Modern time is marked by a great number of versions of different concepts for evolution [1], [2], [7], [12]. The unifying factor in their development is the type of changeability that they accept as a basis of evolution, a determinate direction of adaptability or indeterminate when adaptability is just random.

Development of theories of evolution lead to big chances for development of new scientific trends.

Evolution in biology is defined by hereditary changeability, struggle for existence, natural and artificial selection [1].

During early 40-s in 20-th century, in classical genetics D. Morgan postulated the following evolutionary approaches: (i) genes (elementary units of hereditary information in chromosomes) are mutable, they mutate and mutations of particular chromosomes lead to changes of single elementary features; (ii) all features of the organism are controlled by genes.

Later based on elaborations of M. Watson and F. Crick a new scientific direction appeared, molecular genetics. Genetic code turns out to be common for all natural systems on our planet and biological specifics of living creatures are inherited according to definite rules (discovered by Mendel).

New scientific branches of knowledge originated, e.g. bionics (*bion* in Greek means *element of life*) that joints together biology, genetics and technique aimed at solving engineering-technical problems on the basis of genetics and analysis of vitality of organisms.

[1] Hristo I. Toshev, [2] Stefan L. Koynov and [3] Chavdar D. Korsemov are with the Institute of Information Technologies, Bulgarian Academy of Sciences, Acad. G. Bonchev str., bl. 29A, 1113 Sofia, Bulgaria, E-mail: toshev@iinf.bas.bg, slk@iinf.bas.bg, chkorsemov@iinf.bas.bg

At the end of the 50s scientists from different countries [7], [8], [10] explored in detail evolutionary systems and independently came to the conclusion that they could use the theory of evolution as an instrument for optimization in the process of solution for problems of different nature with the main goal creating a population of eventual solutions using some of the most characteristic peculiarities of nature – heredity, changeability, selection and so on.

Genetic algorithms are a method for search based on the selection of the best species in the population in analogy to the theory of evolution of Ch. Darwin [7].

Their origin is based on the model of biological evolution and the methods of random search. From the bibliographical sources [3], [7], [8], [10] it is evident that the random search appeared as a realization of the simplest evolutionary model when the random mutations are modelled during random phases of searching the optimal solution and the selection is modeled as “removal” of the unfeasible versions.

The main goal of GA-s is twofold:

- abstract and formal explanation of the adaptation
- modelling natural evolutionary processes for efficient solution of determined class of optimization and other problems.

During the last years a new paradigm is applied to solve optimization problems GA-based and modifications of GA. GA realize searching a balance between efficiency and quality of solutions at the expense of selecting the strongest alternative solution from undetermined and fuzzy solutions [12].

II. EVOLUTIONARY THEORIES

2.1. Lamarck’s Theory

Lamarck in the beginning of 19-th century (in 1809) [2] presented the first group of concepts and hypotheses of evolution. He assumed that all living creatures adapt expediently to the environmental conditions. He accepted that the reasons for evolution are: the trend towards a progress, a development from simpler to more sophisticated (gradation) and also a modification and an adaptation of organisms to their surroundings.

These changes are provoked directly by the external conditions thus leading to an improvement of the organs and also to an inheritance of the acquired features. These statements are considered a basis of the theological theory of evolution; according to it the ability of creatures to adapt to their environment is their inborn property.

Lamarck explains one of the peculiarities of the organic world with the adaptability. The progressive evolution, the origin of new species (more sophisticated and more perfect) he explains with laws of gradation, the tendency of living creatures to complicate their own structure:

- first law. An improvement (drill) of organs leads to their progressive development. The absence of improvement of organs leads to a reduction.
- second law. The results of improvements and their absence for long periods of stimulations are accepted by the organism as new qualities and consequently they are handed down from one generation to another depending on the influence of the environment.

Though that biologists reject Lamarck's theory, it is a powerful concept for the artificial evolution that is used in science and technique because it allows local searching in cases of complex practical problems with big dimensionalities for a research of local domains of the population.

2.2. Charles Darwin's Theory

Charles Darwin is the first scientist who determined a common principle in living nature; this principle is the natural selection [1]. There are two aspects in his theory: the materials of evolution, and its factors and rootive powers. Strictly speaking, the rootive power of evolution is the natural selection. Selection is used to achieve the best results of development of species. As a science it is created by Charles Darwin; he chose three forms:

- natural selection. It provokes changes connected with an adaptation of the population to the new conditions.
- unconscious selection. The best species (exemplars) are saved in the population.
- methodic selection. The changes in the population are purposeful.

The rootive powers of evolution are the following:

- indefinite changeability, i.e. a hereditarily determined variety of organisms of every population;
- struggle for existence when less adapted organisms are eliminated from reproduction;
- natural selection. The exemplars that are better adapted survive. As a result, useful hereditary changes are accumulated and integrated thus producing new adaptations.

Based on his theory, Darwin proved that the main rootive power of evolution is the selection of the best species. Their realization means that the following conditions must be observed: a correct selection of the original material, a strict determination of the goal, an implementation of the selection in sufficiently wide ranges and a quick removal of the rejected material for the selection of a single basic feature.

2.3. Fries's Theory

In his research [5], [6] he does not take into consideration the integrating role of the natural selection. Instead he bases his platform on the basis of manifestation of mutations as direct factors for the formation of species.

He develops the ideas of saltacionism where the basic way for the origin of species are sudden jumps (quick changes) without any preliminary accumulation of qualitative changes in the evolutionary process. This type of evolutionary mechanism is called an evolutionary failure. It is manifested once per several thousand generations. The basic idea lies in the imported global changes in the contents of the genes on catastrophic occasions. The origin of species in moments of sudden leaps is a real process.

Similar research in this direction was performed also by scientists like Jan Potszki, U. Batton, S. Goldschmidt, W. Kordlam, etc.

2.4. Popper's Theory

The concept of K. Popper [4] belongs to the basic types of discussed theories. He presents an evolutionary epistemology interpreted as a triad deductivism – selection – critical removal of errors.

The exposure is in the form of theses the basic of which are the following:

- all evolutionary systems solve definite problems;
- problems are always solved using the method of trials and errors;
- errors are removed by complete removing of unsuitable species or by a modification;
- the population uses a mechanism that is elaborated in the process of evolution of the species;
- the population is a basic factor of the evolutionary species to which it belongs. It is a trial solution analyzed in the evolutionary process that chooses the environment and its transformation.

The scheme of the evolutionary sequence of events is the following:

$$F1 \rightarrow TS \rightarrow EE \rightarrow F2$$

where F1 is the source problem, TS is the set of trial solutions, EE is the removal of the errors and F2 is the basic problem.

The difference from Darwin's theory with its single problem – survival of the strongest – and the case with Popper is the presence of other problems with the latter: reproduction, removal of the extra generation and so on; besides here the difference between F1 and F2 is important because it explains the creative intellectual evolution.

According to [4] the information does not penetrate in the system, rather the scientists explore the environment and they actively receive information from it.

Briefly said, according to Popper the evolution is a process of random changeability and the selective memory is in the basis of increase of knowledge and growth of adaptability to the surroundings. Three main parts are evident in this process (triads): changeability mechanism (that joints the selection process), saving and spreading of selected versions.

2.5. Dubinin's Theory

Dubinin's theory [5] has united some of the basic concepts from the theories of Darwin, Lamarck, de Fries and the principles of genetics of populations. The basic concepts are the following:

- evolution is impossible without an adaptation of the organisms to the conditions of the environment. The factor formulating the adaptivity is the natural selection; it uses random mutations and recombinations;
- natural selection based on a transform of genetics of populations creates complex genetic systems. Their modifycations are strengthened by a stabilizing factor;
- hereditary changeability in populations has a massive character. The origin of random mutations is inherent only to single exemplars;

- most adapted exemplars produce a great number of successors;
- species originate by an evolution of populations.
- the contradictions between the random character of hereditary changeability and the requirements of selection define the uniqueness of the types of genetic systems.

Processes that comprise the results from realizations of Dubinin's evolutionary theory include:

- random variations in the size of separate populations, periodic fluctuations, unsteady changes and fluctuations in the size of the population that are determined by the migration process.

N. Dubinin separates four basic forms of realization of the inner unity of the evolutionary process:

- microevolution (processes of intraspecies evolution);
- phase of an increasing evolutionary improvement;
- turning points in evolution;
- evolutions of basic specifics in the organization of the natural system.

III. Genetic Algorithms

Genetic algorithms (GA) are a new research domain. They originated as a result from the research of D. Holland and colleagues in the University of Michigan, USA. GA that he described borrow their terminology to a great extent from natural genetics. They appeared as a combination of models of biological evolution and also of methods of random search.

We shall present some concepts and definitions from the theories of GA. All GA function on the basis of some initial information; the population of alternative solutions P covers this information. The population $P = \{ p_1 , p_2 ,\dots p_i , \dots, p_{Np} \}$ presents a set of elements p_i . Here N_p is the size of the population. Every element of this population p_1 as a rule presents one or several chromosomes. Chromosomes consist of genes. The positions of genes in the chromosome are called *locii* (*locus* per a single chromosome), i.e. the gene is a subelement (an element in the chromosome), *locus* is position in the chromosome, *allel* is gene's meaning.

Genes may have numeric and functional values. Usually these numeric values are from some alphabet. The genetic material of the elements usually is coded in the binary alphabet {0,1} thought that it is possible to use alphabetic and also decimal, etc. alphabets.

The elements in GA are usually named *parents*. The parents are selected from the population on the basis of predefined rules that are mixed (cross-fertilized) to produce *children* (*successors*). The *generation* (process of realization of a single iteration of the algorithm) is called an *offspring*.

The evolution of the population according to [8], [9] is a succession of generations where the chromosomes change their values in such manner that every new generation adapts to its environment in the best way.

Each element in the population possesses some definite level of quality that is marked by the value of the *goal function* (GF) (sometimes referenced to as *function of utility* or *fitness function*). GF is used in GA to compare solutions and to choose the best one of them. The basic problem of genetic algorithms is the optimization of GF. With other words GA analyze populations of chromosomes that are combinations of elements of some set and after that they optimize GF estimating every chromosome. GA manipulate populations of chromosomes based on the mechanism of natural evolution.

In simple GA created by Goldberg [8] the realization is carried out by three basic genetic operators; selection, crossover and mutation [8].

There exist plenty of different types of selection that we can conditionally divide into probabilistic, determined and combined. We must note that explorers of GA use more and more combined selections with a predefined knowledge about the problem to be solved.

Crossover methods are: simple (single point), paired, sequential, serial, partially corresponding, cyclic, universal, etc. It is important to note that the search of the optimal crossover operator is still going on.

Mutation operators are necessary to avoid losses of important genetic material. More important mutation operators are: single point, paired, of inversions, of translocation, of segregation, of removal and adding.

We can show more precisely the real necessity to use GA-s if we try to present some of their most characteristic peculiarities that distinguish them from traditional optimization methods for search and learning [10], [12], [15]:

- GA operate basically not with the problem parameters, but with a coded set of parameters;
- they realize the search not by improving the solution, but by using several alternatives of the defined set of solutions;
- they use a goal function, not different increments for evaluation of quality of the accepted solutions;
- the accepted rules for analysis of the optimization problem are not determined, but probabilistic.

In operational mode, the GA chooses a set of natural parameters of the optimization problem and it codes them into a sequence with a finite length in some alphabet. The GA operates until a given number of iterations is fulfilled (algorithmic iterations) or during some iteration a solution with a definite quality is obtained or a local optimum is found, i.e. a premature convergence is observed and it is impossible to find an exit from this state [11], [13], [14].

GA-s provide a series of advantages for solving problems in practice. One advantage is the adaptation to the changing environment. In real life the problem that is postulated for solution can undergo vast modifications during the process of its solving. The usage of traditional methods requires that all calculations must be performed from scratch and this leads to big losses of machine time. In the case of evolutionary approaches the population can be analyzed, supplemented and modified in accordance with the modified conditions. For this reason the complete selection is an option, not obligation.

Another important peculiarity in the process of solving practical problems supported by GA is the necessity to execute some preliminary stages:

- *Choose a way for solution presentation* (stage 1). The chosen structure must allow coding of all possible solutions and it also must admit and ensure their estimate [11].
- *Choose random operators for posterity generation* (stage 2). Besides the usage of the two basic types of reproduction,

sexual and asexual (cloning), it is possible to use also operations that do not exist in real nature: for example, the usage of material from three and even more parents that is accompanied by voting for the choice of parents. There are no limitations for the usage of different operators and also it is senseless to copy laws of nature and their limitations mechanically and blindly [3], [7,], [8], [11];

- *Define DM rules for creation and selection of posterities* (stage 3). The simplest choice is the case to accept only the best solutions and to ignore the rest of the decisions. Very often this rule turns out to be little effective and that good decisions may come also from bad ones, i.e. not just from the best ones. Therefore this leads to the logical assumption that the possibility for choice of the best decision must be the most important principle at this stage;

- *Create an initial population* (stage 4). If there is a lack of knowledge for the considered problem then solutions are accessible according to some random choice from the set of possible solutions. This points to a process of generation of random problems when every single problem itself is a definite solution. On the other hand during the process of creation of the primary population it is possible to use data received during the experiment of solving the same problem with other GA-s. If these solutions are really valuable then they will outlast and produce a posterity if they do not perish together with other weak individuals meanwhile [3], [8].

V. Conclusions

The performed analysis of development of some basic evolutionary theories forms the image of a continuous ascending progress of science in this trend. It includes wide possibilities for applications of ideas for evolutionary approaches in cases of solving vast ranges of problems. In order to provide a rough idea of the sort of applications tha are being tackled in the current literature, we will classify the applications in three large groups [17]: engineering, industrial and scientific. A representative sample of engineering applications of the following: electrical engineering, robotics and control [20], structural engineering, hydraulic engineering.

Industrial applications are the following: design and manufacture, scheduling [18], management.

Finally, we have a variety of scientific applications: chemistry [21], physics, medicine, computer science.

GA are a new and effective optimization methodology that is used to solve different optimization problems based on an analogy of natural processes of selection and genetic transformations taking place in nature.

The strong interest for using GA in so many different disciplines reinforces the idea of the multi-objective nature of many real world problems. However some applications domains have received relatively little attention from researchers are represent areas of opportunity. For example [19]: cellular automata, pattern recognition, data mining, bioinformatics and financial applications.

References

[1] Ch.Darwin, *"The Origin of Species by Natural Selection*", Works, V. 3, Moscow-Leningrad, Academia, 1939 (in Russian).

[2] J.-B.Lamarck, "*Philosophy of Zoology"*, V. 1, 2, Moscow-Leningrad, Academia, 1939 (in Russian).

[3] V.V.Emelianov, V.M.Kureichik, V.V.Kureichik, *"Theory and Practice of Evolutionary Modelling"*, Moscow, 2003 (in Russian)л

[4] Evolutionary Epistemology and Logic of Social Sciences: Karl Popper and His Critics // Compilers D.G.Lahuti, N.V. Sadovsky, K.Z.Finn, Moscow, Editorial URSS, 2000 (in Russian).

[5] N.P.Dubinin, *"Selected Works"*. V. 1, "*Problems of the Gene and Evolution"*, Moscow, Nauka, 2000 (in Russian).

[6] E.Meyer, "*Populations, Species and Evolution"*, Moscow, Mir, 1974 (in Russian).

[7] J.H.Holland, *"Adaptation in Natural and Artificial Systems"*, the MIT Press, 1992.

[8] D.L.Goldberg, *"Genetic Algorithms in Search, Optimization and Machine Learning"*, Addison-Wesley, 1989.

[9] L.Davis (Ed.), Handbook of Genetic Algorithms, N.Y., Van Nostrand Reinhold, 1991.

[10] M.Mitchell, *"An Introduction to Genetic Algorithms"*, MA: MIT Press, 1996.

[11] J.R.Koza, *"Genetic Programming 2"*, Cambridge MA: MIT Press, 1994.

[12] D.Goldberg, *"Web Courses"*, Available at: http://www.engr.uiuc.edu/OCCE, 2000.

[13] C.M.Fonseca, P.J.Fleming, Genetic Algorithm for Multiobjective Optimization: Formulation, Discussion and Generalization. In: *Proceedings of the Third International Conference on Genetic Algorithms* (S. Forest, Ed.) Morgan Kaufmann Publishers, pp. 416-423, 1993.

[14] A.Jaszkiewicz, Genetic Local Search for Multi-Objective Combinatorial Optimization, *European Journal of Operational Research*, 137, pp. 50-71, 2002.

[15] M. Vassileva, A Generalized Scalarizing Problem of Multicriteria Optimization. *Comptes Rendus de l'Academie Bulgare des Sciences*, 58, 5, pp. 537 – 544, 2005.

[16] C. A. Coello, Evolutionary Multi-Objective Optimization: A Historical View of the Field, *IEEE Computational Intelligence Magazine*, February 2006, pp. 29-36, 2006.

[17] C. A. Coello and G. B. Lamont, editors, Applications of Multi-Objective Evolutionary Algorithm, World Scientific, Singapore, 2004, ISBN 981-256-106-4, 2004.

[18] T. Hanne and S. Nickel, "A mutiobjective evolutionary algorithm for scheduling and inspection planning in software development projects", European Journal of Operational Research, vol. 167, no. 3, pp. 663-678, Dec. 2005.

[19] R. M. Hubley, E. Zizler and J. C. Roach, "Evolutionary algorithm for the selection of single nucleonide polymorfizm", BMC Bioinformatics, vol. 4, no. 30, July 2003.

[20] A. Molina-Cristobal, I. A. Grifin, P. J. Fleming and D. H. Owen, "Multiobjective controller structure with genetic algorithms", In Proceedings of the 2005 IFAC World Congres of Automatic Control, Prague, Czech Republic, July, 2005.

[21] S. Majumdar, K. Mitra and G. Sardar, "Kinetic analysis and optimization for the caralync estenfication step of PPT polymerization", Macromolecular Theory and Simulation, vol. 14, pp. 49-59, 2005.

Technologies for Web Services Orchestration

Elena T. Ivanova[1]

Abstract – **Coordination of multiple services is considered. The paper surveys some aspects of main technologies for web services orchestration and focuses on two of them, BPEL and XPDL.**

Keywords – **Web services, Orchestration, BPEL, XPDL**

I. INTRODUCTION

Web service technologies sprang up from the need for different software applications to interoperate with each other. In general, Web services are modular applications or functions, which are generally independent and self-describing, that can be discovered and called across the Internet or an enterprise intranet. Web services are based on open standards: WSDL (Web Services Definition Language) for description, SOAP (Simple Object Access Protocol) for communication and UDDI (Universal Description, Discovery, and Integration) for register and discover services. Service orchestration is the next stage in developing and extending web services paradigm and addresses the need of composing several services in some business logic in order to achieve more complex and meaningful processes.

Term Orchestration refers to coordinating of multiple tasks. Orchestration of services means coordinating of several services (i.e. their relations and consequence of execution) so that the result is a complex service or a process composed by these services.

There is another set of technologies, relevant to composing of services – choreography. Choreography refers to global, multiparty, peer-to-peer collaborations where the component entities interact in long-lived stateful and coordinated way regardless of any programming model or supporting platform used. Choreography languages (examples are WSCI, WS-CDL, etc) are mainly descriptive and cannot be directly executed. These languages are necessary to be mapped to an orchestration language in order to be executed. Orchestration focuses on the behaviour of the composed services as a whole. Orchestration languages (e.g. BPLM, BPEL, XPDL, BPELJ, jPDL, etc) are executable languages and define a runtime environment for their execution.

Choreography Orchestration
UDDI WSDL
SOAP
HTTP
TCP/IP

Fig. 1. Web services protocol stack

[1]Elena T. Ivanova is with Institute of Computer and Communication Systems – Bulgarian Academy of Sciences, Acad. G. Bonchev bl.2, 1113 Sofia, Bulgaria, e-mail: e_ivanova@hsh.iccs.bas.bg

Choreography and orchestration languages are at the top of web services protocol stack (fig.1).

The paper surveys some aspects of main technologies for web services orchestration and focuses on two of them, BPEL and XPDL. Most of the software tools supporting orchestration are based on them.

II. LANGUAGES FOR ORCHESTRATION

There are several standardization bodies and organizations that aim at developing and adopting different orchestration technologies. Some of them are listed in the following Table I.

TABLE I

STANDARDIZATION BODIES FOR ORCHESTRATION TECHNOLOGIES

Standardization body	Short Description
BPMI (Business Process Management Initiative)	This is a non-profit organization that empowers companies of all sizes, across all industries, to develop and operate business processes that span multiple applications and business partners, behind the firewall and over the Internet. [1]
JSR (Java Specification Request)	This is an initiative of the Java Community Process (JCP) that standardizes how to automate business processes on a J2EE server. The JSR wants to standardize a set of XML meta tags that should be specified as meta data.[2]
OASIS (Organization for the Advancement of Structured Information Standards)	It is an international consortium working towards development, convergence, and adoption of common e-business standards.[3]
OMG (Object Management Group)	It is an open membership, not-for-profit consortium that produces and maintains computer industry specifications for interoperable enterprise applications.[4]
WfMC (The Workflow Management Coalition)	This is a consortium of about 300 organizations that defines a set of related standards based on an interesting reference model. The reference model describes the relation between a WFMS and other actors.[5]

There has been much effort in developing languages for web services orchestration during the last five-six years.

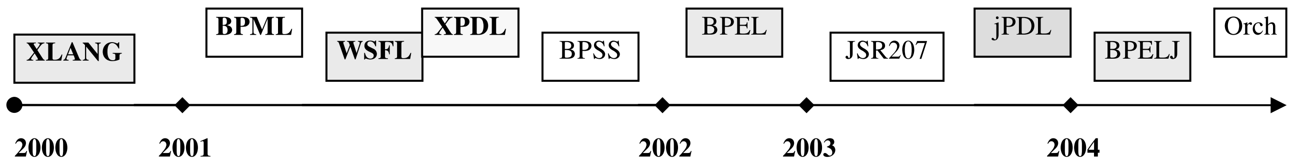


Fig. 2. Development of Orchestration Technologies

The chronological developments of these languages are tried to be represented on fig.2. Short descriptions of these languages are presented below in alphabetical order.

BEPLJ (BPEL for Java) [6] - It is a combination of BPEL and Java proposed by IBM and BEA that allows sections of Java code, called Java snippets, to be included in BPEL process definitions. Snippets are expressions or small blocks of Java code that can be used for things such as: loop conditions, branching conditions, variable initialization, Web service message preparation, logic of business functions etc. BPELJ introduces a few minor changes to BPEL as well as several extensions in order to fit BPEL and Java conveniently together. There is an effort to standardize BPELJ as part of JSR 207.

BPEL (Business Process Execution Language) [3] - This is a convergence of WSFL and XLANG, written by developers from BEA, IBM, SAP, Siebel and Microsoft. It is an XML-based language, depending on WSDL, XML-Schema and XPath, providing a language for the formal specification of business processes and interaction protocols. It supports both abstract and executable processes. In BPEL, a business process is composed of elements ("activities") that define activity behaviors, including the ability to invoke Web services and control flow, and to compensate when errors occur. The resulting business process is exposed as one or more Web services.

BPML (Business Process Modelling Language) [1] – This is a textual service-oriented process modelling language, based on XML, XML-NS, XML-S, XPath and WSDL. It provides an abstracted execution model for collaborative and transactional business processes. It is a strict superset of BPEL, although BPEL have been endorsed by for BPMI phase 2.0 instead of BPML. A process is viewed as a series of activities and an activity represents a component that performs a specific function. Activities can be composed into complex activities. Activities execute within a context which is transmitted from parent to child. In particular the context allows two activities to share properties. Properties constitute the data flow of BPML.

BPSS (Business Process Specification Schema) [7] - BPSS provides a standard framework (language) for business process specification. As such, it works with the ebXML Collaboration Protocol Profile (CPP) and Collaboration Protocol Agreement (CPA) specifications to bridge the gap between Business Process Modelling and the configuration of ebXML compliant e-commerce software.

jPDL (jBPM Process Definition Language) [8] - It is an XML based process execution language running in the Open Source JBoss jBPM (Java Business Process Management) workflow management system. Java actions can be triggered from the language.

JSR 207 (Java Specification Request) [2] is a draft JCP (Java Community Process) specification that is defining metadata, interfaces, and a runtime model that enables business processes to be easily and rapidly implemented using the Java language and deployed in J2EE containers. It will support tasks commonly encountered when programming business processes, e.g. parallel execution and asynchronous messaging.

Orch [9], [10] - This is a programming language and system for orchestrating distributed services. Orc has a strong theoretical foundation that supports modular composition and analysis of concurrent programs. The Orc model [9] assumes that basic services, like sequential computation and data manipulation, are implemented by primitive sites. Orc provides constructs to orchestrate the concurrent invocation of sites to achieve a goal – while managing time-outs, priorities, and failure of sites or communication.

XLANG [11] - It is an extension of WSDL, providing both the model of an orchestration of services and collaboration contracts between orchestrations. It has been standardized by Microsoft and has been used as a base by BPEL. XLANG, like BPML, were designed with an explicit -calculus theory foundation. WSFL and XLANG have converged to BPEL.

XPDL (XML Process Definition Language) [5] - It is an XML-based language from the Workflow Management Coalition (WfMC) for defining business processes. XPDL was based around a common set of functions for work distribution found in most workflow products. The XPDL's syntax is specified by an XML Schema.

WSFL (Web Services Flow Language) [6] - It is an XML based language, standardized by IBM, for the description of Web Services compositions as part of a business process definition. It relies and complements existing specifications like SOAP, WSDL, XMLP and UDDI. WSFL considers two types of Web Services compositions: the first type specifies an executable business process known as a flowModel; the second type specifies a business collaboration known as a globalModel.

Available orchestration technologies can be divided into three groups, according to their basic language: XML based, programming language based (mostly java based), and other based. This classification is represented on the following figure 3.

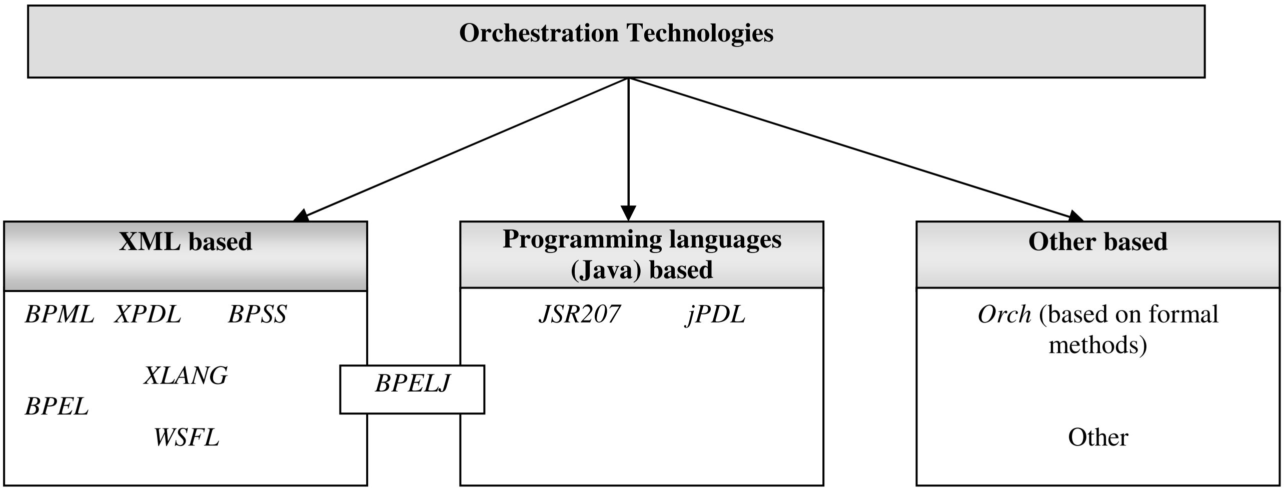


Fig. 3. Types of Orchestration Technologies

III. Requirements for Web Services Orchestration

In order for web services to be composed into a complex service or business process, there are a set of requirements which should be addressed. Several works and papers deal with such requirements [12, 13]. The main of them can be summarized as follows:

- ability to invoke services in an asynchronous manner – the system has to enable services to be invoked concurrently, not only sequentially;
- ability to manage transaction and compensation – this is crucial for long-running compound services – in case of a failure in a single component service, the system has to provide adequate compensation;
- exception handling – how the system will behave when some error occurs;
- security and reliable messaging – security is a key point in all the layers (see fig.1) in web service technologies;
- support the separation of abstract process logic and concrete web services used – it is necessary for building dynamic and flexible processes.

Almost all available languages for web services orchestration have mechanisms respecting the aforementioned requirements.

IV. BPEL and XPDL – main features

This section is devoted to the most used orchestration languages – BPEL and XPDL.

A. BPEL

BPEL is:

- is an XML-based language designed to enable task-sharing for a distributed computing;
- is written by developers from BEA Systems, IBM, and Microsoft;
- combines and replaces IBM's WebServices Flow Lan–guage (WSFL) and Microsoft's XLANG specification;
- depends on WSDL, XML Schema, and XPath;
- supports both abstract processes and executable processes:
  - Abstract processes are useful for specifying expected protocols and publicly visible behaviors without too much detail;
  - Executable processes contain enough detail to fully specify execution.
- BPEL is essentially a layer on top of WSDL, with WSDL defining the specific operations allowed and BPEL defining how the operations can be sequenced.

In BPEL, a business process is composed of elements ("activities") that define activity behaviors, including the ability to invoke Web services and control flow, and to compensate when errors occur. The resulting business process is exposed as one or more Web services. The BPEL elements are listed below.

- Partners / Partner Links – the different parties that interact with the business process
- Variables – Data variables used by the process; Persistent for long running interactions; Defined in WSDL types and messages
- Correlation Sets – Set of properties shared by all messages in a correlated group
- FaultHandlers – Activities that must be performed in response to a fault;
- CompensationHandlers - allowing the process designer to implement compensation actions for certain irreversible errors in business (wrapper for a compensation activity);
- EventHandlers – invoked concurrently if the corres–ponding event occurs;
- Activities (Flow Logic) – the actions that are being carried out within a business process. There are two types of activities:
  - basic activities (<receive> <reply> <invoke> <assign> <throw> <terminate> <wait> <empty>);
  - structured activities (<sequence> <switch> <while> <pick> <flow> <scope> <compensate>).

It is very hard to itemize all the software tools for service orchestration based on BPEL. The list is quite long. Some of these products are mentioned in the following Table II.

TABLE II
SOFTWARE TOOLS BASED ON BPEL

ActiveBPEL Engine	Bexee
ActiveBPEL Designer	Biztalk Server
ADONIS	Cape Clear Orchestrator
Apache Agila	iGrafx BPEL
BEA WebLogic	MidOffice
Oracle BPEL Process Manager	PXE
Parasoft BPEL Maestro	SAP NetWeaver Exchange Infrastructure

B. XPDL

It is an XML-based language from the Workflow Management Coalition (WfMC) for defining business processes. Whereas BPML and other business process languages are geared to Web services, the foundation of XPDL was based around a common set of functions for work distribution found in most workflow products. XPDL uses an XML-based syntax, specified by an XML schema.

The main elements of the language are [5]:

• Package – the container holding the other elements;

• Application – is used to specify the applications/tools invoked by the workflow processes defined in a package;

• Workflow-Process – is used to define workflow processes or parts of workflow processes; it is composed of elements of type Activity and Transition;

• Activity – the basic building block of a workflow process definition. There are three types of activities: Route, Implementation, and BlockActivity. Activities of type Route are dummy activities just used for routing purposes. Activities of type BlockActivity are used to execute sets of smaller activities. Element ActivitySet refers to a self contained set of activities and transitions. A BlockActivity executes such an ActivitySet. Activities of type Imple–mentation are steps in the process which are implemented by manual procedures (No), implemented by one of more applications (Tool), or implemented by another workflow process (Subflow).

• Transition – elements of type Activity are connected through elements of type Transition.

• Participant – is used to specify the participants in the workflow, i.e., the entities that can execute work. There are 6 types of participants: ResourceSet, Resource, Role, OrganizationalUnit, Human; System.

• DataField

• DataType.

The last two elements are used to specify workflow relevant data. Data is used to make decisions or to refer to data outside of the workflow, and is passed between activities and subflows.

Several orchestration software tools are listed in the Table III. The number of available orchestration products, based on XPDL, is quite less than these, based on BPEL.

TABLE III
SOFTWARE TOOLS BASED ON XPDL

Aspose.Workflow	ObjectWeb BONITA
Agentflow	OFBiz Workflow Engine
Enhydra JaWE	Open Business Engine
Enhydra Shark	Vignette Process Workflow Modeler
Newgen	wfmOpen

V. CONCLUDING REMARKS

Service orchestration is a natural growth and extension of web services paradigm. It refers to composing several services in a complex process. The paper has reviewed available technologies for web services orchestration and extracted some key features of theirs. The attention was paid on two of these technologies, BPEL and XPDL, because of their wide acceptance.

VI. REFERENCES

[1] http://www.bpmi.org
[2] http://www.jcp.org/en/jsr/
[3] http://www.oasis-open.org/home/index.php
[4] http://www.omg.org
[5] http://wfmc.org/
[6] http://www.ibm.com
[7] http:// www.ebxml.org
[8] http:// jbpm.org
[9] http://www.cs.utexas.edu/users/wcook/projects/orc/
[10] Misra J.and W. Cook, 2004, *Computation Orchestration: A Basis for Wide-Area Computing,* Lecture Notes for NATO summer school, held at Marktoberdorf in August 2004
[11] http://www.microsoft.com
[12] Peltz C., Web ser vices orchestration, available at http://devresource.hp.com/drc/technical_white_papers/WSOrch/WSOrchestration.pdf
[13] Yushi C., L. E. Wah and D. K. Limbu, Web Services Composition - An Overview of Standards, Singapore Institute of Manufacturing Technology

This work is partly supported by European Commission, project VISP №FP6-027178, and National Scientific Fund of Bulgaria, project № ВУ-МИ-108/2005.

Cost effective fourth generation network for small and medium enterprises

Tsvetan V. Filev[1]

Abstract – **In the current paper an approach will be presented for design of cost effective fourth generation network based on existing standards and software. The network propose number of services like multimedia, corporative services access: email, ERP, CRM, groupware, and also phone conversations between users. The used technologies are WiFi, skype and DHCP.**

Keywords – **WiFi, 4G, skype, small medium enterprises.**

I. Introduction

Working networks from this kind are extremely suitable for covering the area of different buildings of small and medium enterprises, universities, government institutions, schools etc. They allow authorized or unauthorized users to use free of charge conversations, email access, corporate web sites, groupware or web pages.

The construction is very easy and is accomplished by the use of existing software and hardware tools.

II. Review of different mobile technologies – evolution

1G – Wireless technologies from this kind appear in the beginning of the 70's of the previous century. As a whole they propose low mobility that concerns both coverage and speed of movement, low power, and two way wireless voice communications. The modulations and the process of signals are absolutely analog and frequency division is used to devide the channels so the users can be distinguished.

The speed of data transmission is extremely low and inappropriate for modern digital technologies including the internet

2G – The analog technology acquire great popularity and is wide spread fast and evolutes after few unsuccessful attempts into digital. In a result in Europe GSM system appears – Global Standard for Mobile, which uses GMSK – Gaussian Minimum Shift Keying.

In north America the analog is IS-54. In Japan PDC – Personal Digital Cellular shows up. The technologies are fully digital and even though the frequency band is very low to be applied to fast advancing existing digital technologies and computer wired networks and the services they offer.

3G – The great success of digital networks for connection leads to the appearance of the networks from third generation, which employs wideband modulation or CDMA – Code Division Multiple Access. This fully digital network offers high mobility, all current services as voice transfer, messages exchange as data exchange, video and files, internet connection etc.

4G – These networks are still experimental and because of the fact that there are no good expressed standards and also frequency resources. However there are multiple designs which allow to be adopted in small and medium enterprises with the aim to decrease expenses and increase communication abilities. The networks are built with the help of base stations covering one premises, floor or building.

III. Review of the exisiting working protocols

TCP/IP – This is a protocol designed exclusively for packet oriented networks like the Internet. It allows the construction of virtual sessions between separate users as every each one of them has its address called IP. The protocol allows statically and dynamically to point the separate packets between users and numerous of other operations.

DHCP – This is a protocol for dynamic assignment of the settings of a given network users. It allows dynamically or previously defined to allocate network addressess. For this purpose software is attention which works on centralized server. The settings can be assigned only for certain hardware devices which brings security in the network.

Routing, QoS, masquerade – The routing is a process of directing the packets to their proper destination. It can be statically predefined or dynamically, as it changes in working progress. QoS or Quality of Service is a service performed by the routers aimed to give priority of certain services like voice or text messages. Masquerading aims the translation of one real address to internal not real address to extend the network and for security.

WiFi – This is a wireless protocol on the IEEE 802.11g standard. It allows the construction of so called WLAN which mimics network from fourth generation. The protocol exploits widespread modulation and the last designs allow coverage of 100 meters with non directional antennas.
Therefore a floor or a building could be covered very easy with two or even one wireless access point.

IV. Review of the existing software

Linux – The operating system Linux is born in the beginning of nineties of the last century. It very fast gains popularity and transforms to whole area of the software. It inherits the functionality of the UNIX systems and simplicity and the readiness of Windows.

In practice on this platform can work every kind of software including one written for other systems. The open

[1]Tsvetan V. Filev is with the Faculty of Computer Systems and Control, Technical University, Kliment Ohridski 8, 1000, Sofia, Bulgaria, E-mail: tsvetan.filev@gmail.com

source of Linux makes it easy and fast for development. Also the price for adoption and support is very low and the operating system itself is free of charge.

DHCPD – This is a software application allowing dynamically assignment of addresses and settings of the users of a given network using the protocol DHCP.

Skype – This easy and fast program gets unseen success for very short period of time among the Internet users. The good encoding of the sound and the easy interface also the possibility to call subscribers outside Internet are indisputably leading components of now day’s communications software.

VoIP – There are numerous Voice over IP software solutions which can be used with the aim to construct internal communications and connection to other communication networks like the Internet, GSM, UMTS and others.

V. SCHEME AND A WAY OF WORK OF THE NETWORK

On Fig. 1 a scheme of the network for communication and connection is depicted. By using one router connected to fixed and wireless network simultaneously is accomplished routing, masquerading, address assignment, security, accounting and eventually VoIP. The presence of centralized server ensures the easy handover between the separate “base stations”. The Handover is reduced only to hardware transfer between two base stations while the DHCP server is concerned to save the settings of the users for the same hardware address with the aim to save the TCP sessions and security.

The conversations are formed through skype call which not only allows connection to other users but outside mobile and fixed networks. For higher quality of the conversations is good to use QoS on the centralized router and the routers realizing connection to the internet.

The radio part is absolutely independent and the appearance of new better networks wouldn't break the structure. This would only make better mobility and greater capacity and from there good quality in relatively low price for upgrade. For end devices PDA – Personal Digital Assistant are used.

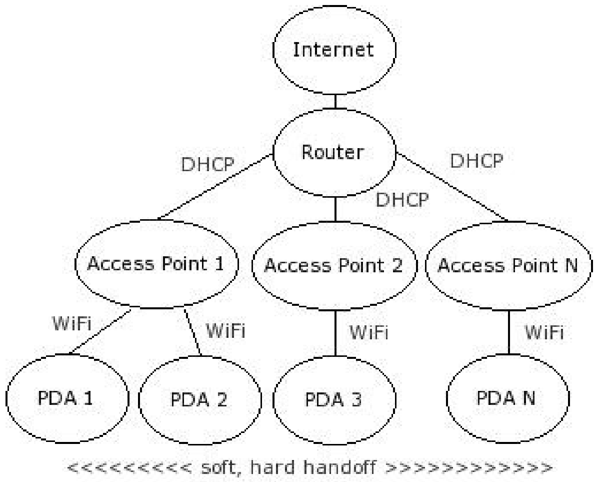


Fig. 1. A scheme of a fourth generation network.

VI. CONCLUSION

Such a network is extremely easy to be realized with the help of existing technologies while the price remains on very low levels. Under low price we understand not only the price for construction and adoption but the price for support and use of the network. The fast advance of wireless communications leads to the appearance of newer and faster, and better protocols for connection. Independently from this though the adoption of new radio network would not lead to whole reconstruction of the shown structure, but only to replacement of the radio part because of the standard protocols and services.

REFERENCES

[1] Boca Raton, *The communications handbook SE* – CRC Press, London, New York, 2002

[2] http://www.linux.org

[3] http://skype.com

[4] http://www.isc.org/index.pl?/sw/dhcp/

[5] http://www.suse.com

[6] http://qos.ittc.ku.edu

[8] http://lartc.org

A Flexible Architecture for Customizable Web Based Spreadsheet Engine[1]

Ivo Marinchev[2]

Abstract – **In this paper we introduce a flexible architecture for development of scalable, and customizable web based spreadsheet engines. Our architecture is geared towards new trends and technologies in application development as web services, rich clients, AJAX, web applications, etc. Finally we present our current implementation developed as AJAX based web application that uses relational database for persistence storage.**

Keywords – **spreadsheet engine, AJAX, web applications, web services, rich-clients, web 2.0.**

I. Introduction

In the last few years there is a trend to "web enabling" (applications that can be started and run through web browser) of the current desktop software applications. New ones are directly built as web application and many old one are re-implemented or extended as web based.

The first distributed applications with centralized business logic (backend) were rich-client applications. They require more sophisticated and responsive user interface (than what HTML, CSS, and JavaScript can offer at that time) that run on the user's computer and communicates with a business logic that is physically located on the centralized servers. Rich client applications and technologies appear about 10 years ago but could not become widespread. They remain in use mainly in intranet environment inside the organizations or shared between affiliate organizations. Their failure to become widespread was due to many factors some of which are:

- they require the users to install additional software on their systems;
- security concerns;
- not having enough support from big application vendors;
- high price tag.

The most prominent technologies in this area are Microsoft's ActiveX, Java Web Start (JWS), Eclipse Rich Client Platform (Eclipse RCP), and Macromedia Flex.

Recently widespread adoption of web applications became feasible with the introduction of several key technologies in practically all of the modern web browsers - Internet Explorer, Firefox, Mozilla, Opera, Safari, and Konqueror. These technologies are CSS [3, 4], JavaScript [5], DOM [6] and DHTML (dynamic screen re-flow). Although most of them had reliable implementations even in year 2000 the biggest boost started just recently with the introduction of so named XMLHttpRequest [2] object. It allows web pages (using JavaScript) to perform asynchronous request to their originating server and fetch updated data from it. On the next step these data is used to update part of the web page information in timely and responsive manner without requiring page reloads. As the XMLHttpRequest object has support for transferring data in the XML format (presentation/view neutral encoding) the corresponding technology was named AJAX [2] (Asynchronous JavaScript And XML). Hence it becomes possible to develop web applications that look and feel in a way very similar to the regular desktop ones and provide the user with similar usage experience and capabilities. Keeping the business logic on the centralized servers (usually cluster of servers), allows easier management, support maintenance, and upgrades. Also, new schemes of application delivery and usage becomes feasible as pay-per-use, application service providers, click-and-run (no installation is required), application delivered as a service, etc.

The first widespread web applications were Google Mail (http://www.gmail.com) and Flickr (http://www.flickr.com). Another application are web applications are Writely (http://www.writely.com) word processing application (recently March 2006 bought by Google), calendar – 30 boxes (http://www.30boxes.com), CalendarHub (http://www.calendarhub.com).

II. Motivation and Basic Requirements

In the field of web based spreadsheets applications the key players are NumSum (http://www.numsum.com), and iRows (http://www.irows.com), open source applications TrimSpreadsheet (http://trimpath.com/project/wiki/Trim Spreadsheet), WikiCalc (http://www.softwaregarden. com/wkcalpha). Unlike the rest of the web applications, at the time of this writing (March 2006), web spreadsheets are not feature complete, built for specific purposes and non-customizable. Open source ones are mostly unusable and are actually just a proof of concept that the real applications.

Above observations have motivated us to build web based spreadsheet engine to solve our specific needs and to be able to extend and customize it as needed. The requirements for our spreadsheet system are:

1. It must be web based - many different people can use it from any physical location provided internet connection is available.
2. It must be component based – consisting of logically separated software units with well-defined interfaces and behavior.

[1] The work report in this paper has been partially supported by Project IIT-010061 "Technologies of the Information Society for Knowledge Processing and Management".

[2] Ivo Marinchev is with the Institute of Information Technologies, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Bl. 29A, 1113 Sofia, Bulgaria. E-mail: ivo@iinf.bas.bg

3. It must be loosely coupled – every component must be easily replaceable, and must not depend on the rest of the components as much as possible.
4. It must be scalable – can scale up and down with minimal changes.
5. It must support different user roles (groups) - some users can design and edit spreadsheets' definitions (designers). Regular users can just fill the data in the cells that are not locked.
6. Entered data must be kept separated from the spreadsheet definitions so that it can be reused in different spreadsheets and/or other related applications.
7. Ajax version must be able to work on IE and Firefox. If it is possible Opera and Safari must be supported as well.
8. The system must be able to import MS Excel spreadsheets and convert them into its internal form – reuse already created spreadsheets and use MS Excel as a primary design tool until comparable web based spreadsheet designer is developed.

III. SYSTEM ARCHITECTURE

Fig.1 depicts the architecture of our system. It is compliant to the requirements enumerated above and, in practice, is 4-tiers (layers) system that comprises the following layers:

1. *User interface layer*. It consists of all of the software components that are transported through the network to the user system and executed on it. In this layer there are basically two type of implementations – Ajax based that run from the regular web browsers or rich client interface (for example JWS or Eclipse RCP). This layer communicates to the next layer using standard open communication protocols that have HTTP transport (so that the clients can be used behind firewall).

2. *Web layer*. It consists of software components that are executed in the web server environment. These components can be CGI scripts, PHP pages, Java servlets, JSP pages, etc. or even web services implemented in any programming language.

3. *Application layer*. It represents the main application business logic. It comprises all algorithms and internal data structures that are involved in spreadsheets management and processing

4. *Persistent storage layer*. It is a set of databases (Relational and/or XML) that organize the information on the persistent media.

In practice, in case of small deployment scenarios, it is possible the Web layer to be merged with the Applications layer. For example we have such situation when the application layer is implemented using web services (SOAP or REST) that are consumed directly from the user interface layer. This possibility of layer merging is a key feature of any systems that is projected to scale up and down.

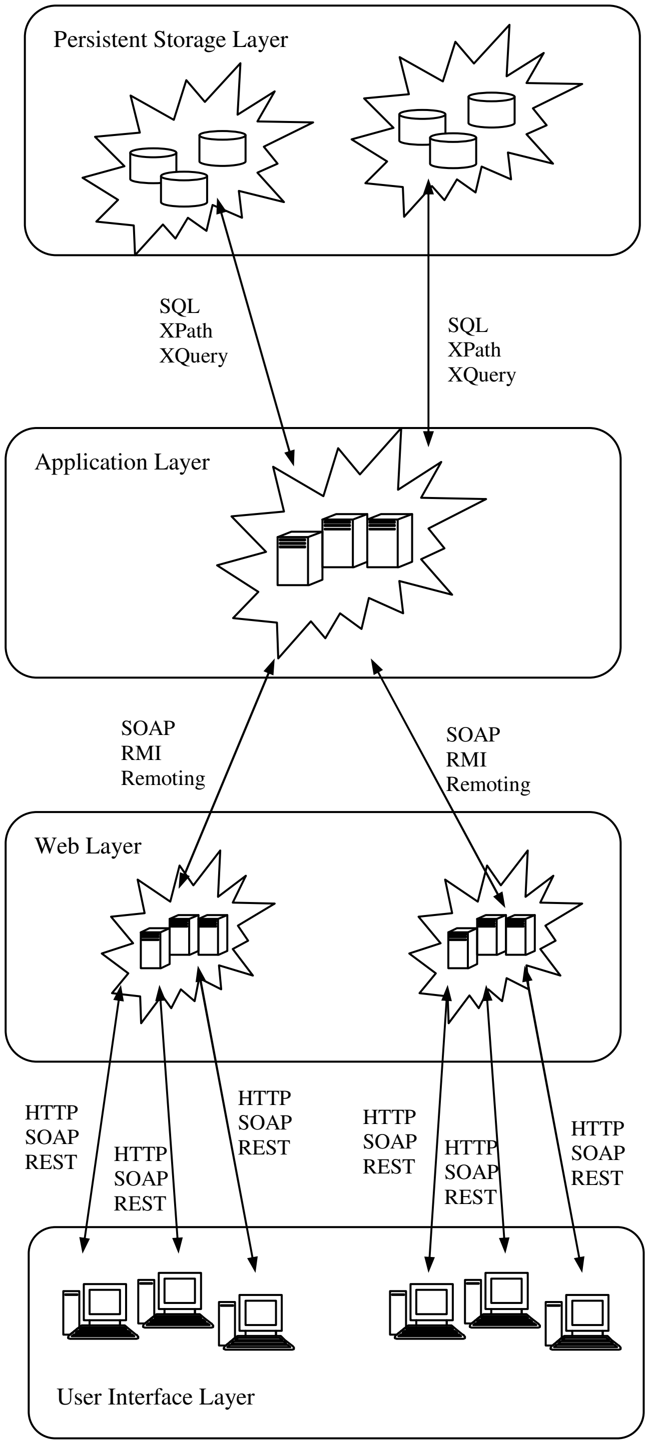


Fig. 1. System architecture.

As Fig.1 shows, all system's layers are very loosely coupled and communicate only with open and standard complaint protocols. This feature makes the system to be completely agnostic to hardware and software platforms that execute its components. In practice any client part and any server part can be executed on different operating system allowing almost any available hardware to be used as a client or a server.

IV. System Implementation

The current version of the system is implemented as a standard 3-tier web application. Technical specifications are as follows:

1. *User interface layer*. JavaScript library that uses Ajax requests to update and fetch data from the web layer. It works on IE (5.5, 6.0), Firefox (1.0, 1.5) and to some degree on Opera (8.5). It is possible to build Java Web Start client in the future.

2. *Web layer*. At the moment it consists of PHP pages (convenient only for stateless services) and Java servlets (better scalability, applicable for statefull services that use big data models). Current implementation uses REST [1] services. Although REST is not a standard but an architectural style, its light-weighted, requires fewer resources, and is simpler and faster for quick-and-dirty implementations. Later we can convert inter-layer communications to the complete SOAP, WSDL, WS-I stack if it is needed. This layer contains all of the algorithms and data structures that dynamically build user interface pages (screens) of the system. It is also responsible for getting and validating request parameters from the user interface layer and reformatting the response data if it is needed.

3. *Application layer*. In the current deployment scenario this layer is merged with the web layer because there is no need for separated web and application layers in the small size deployments. This layer contains all of the algorithms and data structures that implement the core business logic. Upon client requests (coming from the user interface layer) it fetches the data from the persistent storage layer and builds internal data model of the manipulated spreadsheets. Then it uses these data models to re-calculate the spreadsheets values based on the user changes and sends the updated data (user changes) back to the persistent storage layer and forth (recalculated fields values) to the user interface layer.

4. *Persistent storage layer*. Uses relational database storage. It was tested with MySQL and SQL Server, but as it uses standard SQL queries it should work with any complaint relational database system.

Figures 2, 3, and 4 depict a visual demonstration of our system. Fig. 3 shows a complex spreadsheet opened within MS Excel that we used for testing purposes. Fig. 4 shows a screenshot from the same spreadsheet converted to the format of our system loaded in it and then opened in Internet Explorer. Fig. 2 shows one of the unique features of our system its - ability to show the formula as tooltip when the mouse hovers over the corresponding cell.

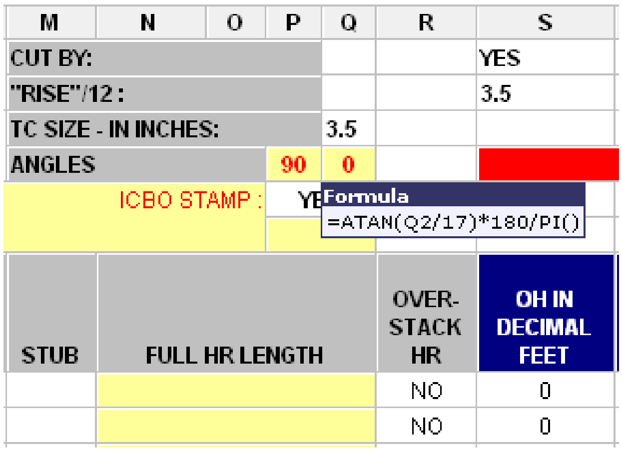


Fig. 2. Some of the new features in our system – shows formulas as tooltips

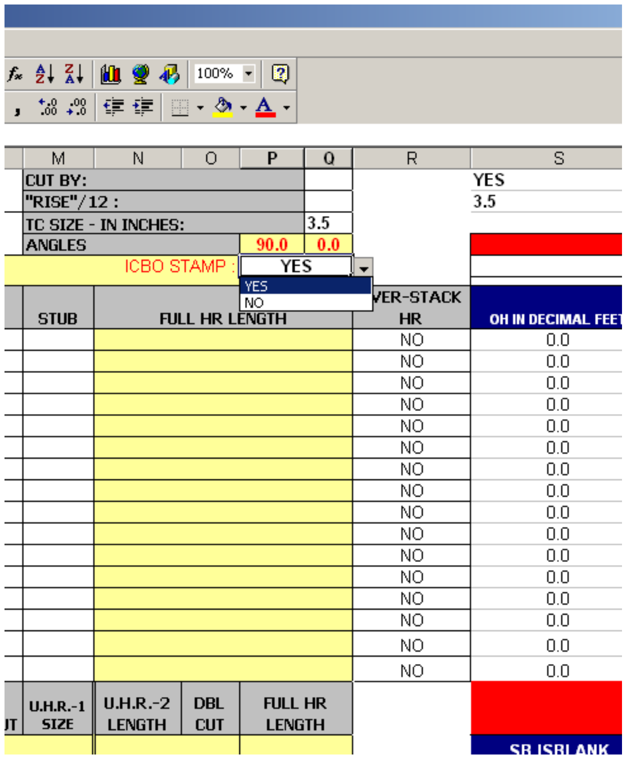


Fig. 3. Complex Spreadsheet in Excel

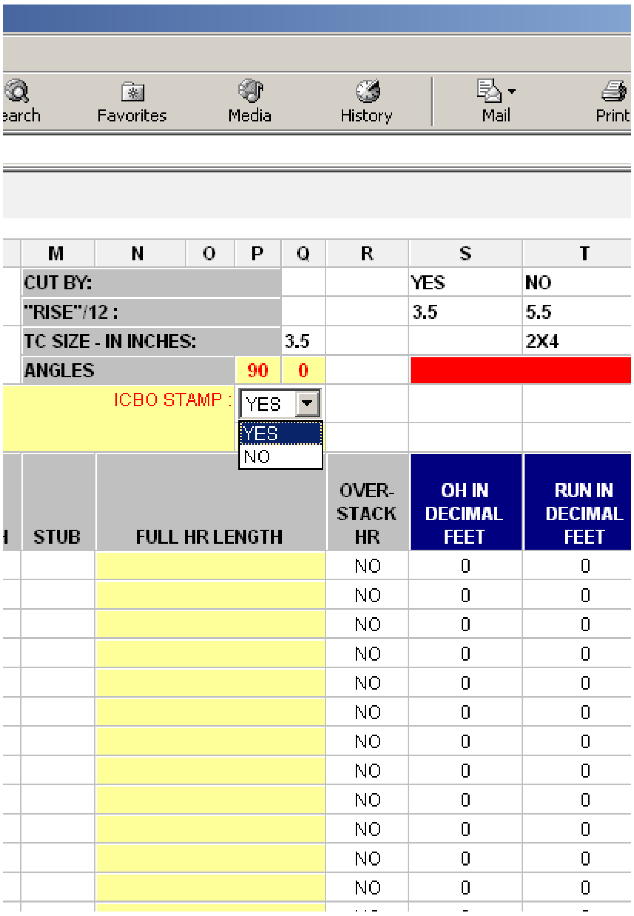


Fig. 4. The same spreadsheet from Fig. 2 in IE

V. Conclusion and Future Work

In this paper we presented our flexible architecture for development of scalable, and customizable web based spread-sheet engines. We presented also our concrete implementation of this architecture. Below we itemize some of the improvement and additions that have to be implemented in near future in order to make the system applicable in wider task scopes:

1. Improve the support of some of the alternative browsers as Opera and Safari that currently do not work due to bugs in their DOM implementation. Opera works with simple spreadsheets but not with our complex one.
2. Implement more of important MS Excel features that are still missing.
3. Add some unique features as the ability to use CGI scripts or web services in formulas. One example of this is getting weather conditions, exchange rates, interest rates, stock quotes, etc. in real-time.
4. Many other optimizations and improvements.

References

[1] R. Fielding. *Architectural Styles and the Design of Network-based Software Architectures*. PhD Thesis, University of California, Irvine, 2000

[2] AJAX, http://en.wikipedia.org/wiki/AJAX

[3] CSS1 Specification, http://www.w3.org/TR/REC-CSS1/

[4] CSS2 Specification, http://www.w3.org/TR/REC-CSS2/

[5] ECMA-262, ECMAScript (JavaScript Specification), http://www.ecma-international.org/publications/standards/Ecma-262.htm

[6] W3C Document Object Model, http://www.w3.org/DOM/

Demonstrating the Effect of the Fragmentation Process Regarding Selecting an Optimal Route

Delyan G. Genkov[1]

Abstract – **The heterogeneous nature of the Internet communications is transparent for the end user. The leading position in the system of the users' criteria is given to the access speed characteristics. The evolution in the communication technologies presents the problem of the adequate choice of metrics for estimating the route and the corresponding routing algorithm as decisive factors for setting the real access speed parameters. One of the factors determining the transit delays is the presence of fragmentation and reassembly of the IP datagrams.**

Keywords - **Internet, IP, Fragmentation, Reassembly, Routing**

I. Introduction

The fragmentation and reassembly function ensures compatibility of different network architectures, connected in an internetwork when the different networks in a route supports different maximal datagram size. Fragmentation of an Internet datagram is required when it comes from a network that permits a large-sized packet but it should pass through a network that limits the package size in order to reach its destination. Although the IP protocol requires a gateway to fragment a packet if it is too large to be transmitted, it can lead to poor performance or complete communication failure. [1]

The modern IP routing protocols aimed to select the best path for the user's datagrams, but in their implementations are not taken to consider the additional delays, caused by the process of fragmentation and reassembly. The term "best path" is given different mean in the various routing protocols, but in the end the ultimate goal is to minimize the overall delay for the user's packet.

Different routing protocols use various criteria to measure the routes and to select the best path, these criteria are called "metrics". After all we can define two basic types of routing protocols:

– *hop-based metric routing protocols* – the criteria for choosing a better route is actually the number of hops, or routers to be passed between the source and the destination. These protocols are quiet simple, but they do not always select the fastest path. In other words their criterion for a better path is actually the shortest path. The most common member of this class is RIP (Routing Information Protocol) [2], developed by Xerox in 1970 and is still widely used in many networks.

– *delay-based metric routing protocols* – their criteria for choosing a better path is often a combination of a different factors, like bandwidth or delay of the links, but in the end they tend to choose the fastest path. These protocols are often more sophisticated than the previous ones, but their decision is closer to the user's criteria for a better route. Typical members of this class are OSPF (Open Shortest Path First), described in [3], and the Cisco's EIGRP (Enhanced Interior Gateway Routing Protocol) [4].

The metrics of both types of routing protocols do not account for the necessity of datagrams' fragmentation in the route chosen by them, therefore, the route is set without taking into considerastion the time delay caused by that process while the data is being transmitted.

The present study aims at proposing a method for modification of the existing routing protocols to the effect that they should take into consideration the process of data fragmentation when the optimal route is fixed.

II. Experimental settings and results

Two experimental settings are chosen for the purposes of the present study. Each of them aims at demonstrating the differences between the time delays of the users' packets along the route chosen by the routing protocol that does not account for the process of fragmentation and reassembly of the packets and the better route that could be chosen if that process is taken into consideration. The routers used are Cisco 2501. All the described experiments are conducted sending 1000 packets with the size of 1500 bytes, which is usual for the Ethernet networks. The packets are sent and the time for each packet's transition to its destination and its way back to the source is measured. That is performed by a specially written C-language program, running on the first host.

Some of the packets (about 1 in 40) showed times that deviated significantly from the average result. Those are the packets that are detained by the router when its routing table is being actualized. Being a process of different nature, not related to the aims of the present study, these data are ignored.

For both the experimental settings are conducted experiments with three of the most frequently used in Internet routing protocols - RIP, EIGRP and OSPF. The first one belongs to the hop-based metric type; the other ones are delay-based metrics type.

The first experiment is presented on fig. 1. Two computers are connected with the corresponding router via 10 Mbit/s Ethernet interface. There are serial links with the speed of 4 Mbit/s between the routers. The maximum transmission unit (MTU) that can be transmitted through route 1 is 576 bytes – the official smaller MTU in Internet. Consequently, to pass along that route each packet of 1500 bytes will be separated in 3 smaller fragments. For route 2 the maximum size of the packet is 1500 bytes, which means that the packets will be transmitted unchanged.

[1]Delyan G. Genkov is with the Department of Computer Systems and Technologies,
Technical University of Gabrovo,
Hadji Dimitar 4, 5300 Gabrovo, Bulgaria,
Email: dgenkov@tugab.bg

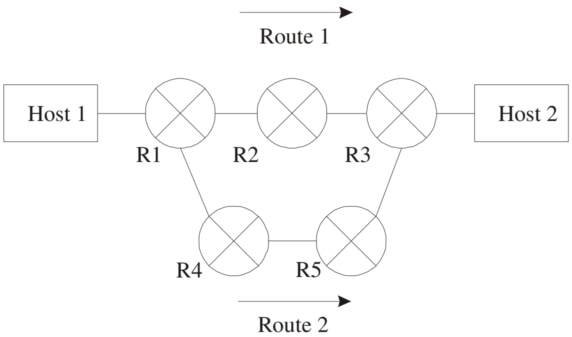


Fig. 1. Experiment 1.

For route 1 the packets pass through three routers and 2 serial lines with the speed of 4 Mbit/s each, for route 2 – they pass through 4 routers and three links with the speed of 4 Mbit/s.

Irrespective of the routing protocol the average values of the packets' transmission time are 30,8 ms for route 1 and 27,6 ms for route 2.

All the three routing protocols under examination select route 1. Actually that route would be the most convenient if there were no fragmentation/reassembly of the packets. The time delays, registered above, show that if they follow the route chosen by the protocol the user's packets will travel for a longer period of time which is determined by the additional delay caused by the process we are studying.

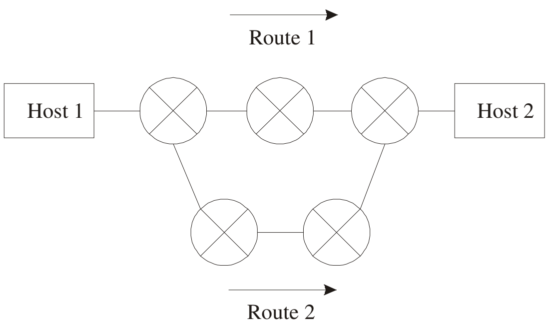


Fig. 2. Experiment 2.

The second experimental setting is shown on fig 2. There are serial connections at the speed of 2 Mbit/s and MTU = 1500 bytes by route 1 and another serial connections at the speed of 4 Mbit/s and MTU = 576 bytes by route 2. There should be no packets' fragmentation on the first route. However, fragmentation is expected on the second route.

The experiments show that the time delay on Route 1 is 41, 5 ms and on Route 2 it is 53,8 ms. The RIP protocol chooses route 1 but the other two protocols choose route 2 because they consider the speed of the linking lines. Route 2 would be the better choice if there were no processes of fragmentation/ reassembly of the packets.

The registered time delays indicate that if the user's packets follow the route chosen by the both protocols they will travel for a longer time, which is determined by the additional delay, caused by the process of fragmentation and reassembly.

III. Conclusions

The present paper aims at demonstrating that there are situations in which the current routing algorithms do not choose the best route for the users' packets because of the fact that they do not consider the time delays, generated by the process of fragmentation of the Internet packets. After estimating those time delays it is necessary to correct the algorithms used by the routing protocols so that they can select the best path. What the results show is that if these delays are taken into consideration the routing protocols could function more adequately to the operating Internet configurations. Since there are different types of routing protocols that use different metrics for setting the optimal route to the receiver of the message, the correction in the different protocols should have different value and dimension.

References

[1] Kent, C., J. Mogul, "Fragmentation Considered Harmful", Digital Western Research Laboratory, 1987
[2] Routing Information Protocol, rfc1058, IETF
[3] The OSPF Specification, rfc1058, IETF
[4] Cisco Systems, Enhanced Interior Gateway Routing Protocol (EIGRP)
[5] Cisco Systems, "Internetworking Technology Overview", 1998
[6] Cisco Systems, "Cisco 2500 Series Router Installation and Configuration Guide", 1996.

Advanced Current mode CMOS OTA Based Band Pass Filters for Detector Readout Front Ends

T. Noulis[1], C. Deradonis[2] and S. Siskos[3]

Abstract — **CMOS current mode band pass filters for front end electronics are proposed. Three semi Gaussian shaper topologies based on operational transcondactunce amplifiers (OTA) are designed using advanced filter design techniques which provide full integration. The implementations are compared in terms of noise performance, power consumption, total harmonic distortion (THD) and dynamic range (DR) in order to examine which is the most preferable in readout applications. Analysis is supported by simulations results in a 0.6µm process by Austria Mikro Systeme (AMS).**

Keywords — **operational transconductance amplifier, readout system, shaper, leapfrog, LC ladder.**

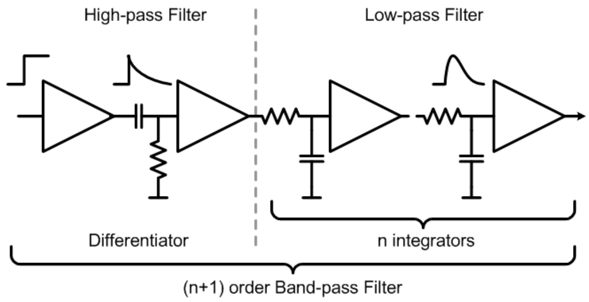


Fig. 1. Principal diagram of an *n* order voltage mode Semi-Gaussian shaper.

I. Introduction

Nuclear radiation detection has been developed in the last few years in various fields of radioactivity control, high energy physics, space science, medical applications and so on. Solid state detectors are gaining importance in a variety of X-rays detection applications that demand excellent resolution. These detectors require compact, low cost and low noise electronics with a high number of channels. Several motivations suggest that the most of these applications can benefit from the use of ASIC (single application specified integration circuit) readouts instead of discrete solutions. Continuous efforts were performed in order to implement readout systems in monolithic form. CMOS technologies have been chosen due to the high integration density, relatively low power consumption and capability to combine analog and digital circuits on the same chip [1]-[3].

The preamplifier – shaper structure is commonly adopted in the design of the above systems. Semi-Gaussian (S-G) shapers are the most common pulse shapers employed in readout electronics [4], [5]. The above typical voltage mode architecture was sufficiently studied (mainly in terms of the CSA input transistor for noise reduction [6]-[9]) but few studies have been performed on pulse shapers and especially on current mode designs. After all, many current mode preamplifiers were so far suggested [10], [11]. This is due to the fact that a current mode structure could be an attractive alternative to the more typical voltage mode one, since the signal is processed in the current domain, avoiding charging and discharging of the parasitic capacitance to high voltage levels and keeping the internal nodes of the circuit at low impedance values.

In this work, current mode S-G shaper designs based on operational transconductance amplifiers (OTA) are proposed. Advanced filter design techniques [12], [13] which provide full integration, are used and novel CR-RC^2 implementations are suggested. All the implementations are compared in terms of noise performance, power consumption, total harmonic distortion (THD) and dynamic range in order to conclude which the optimum one is.

II. Design of OTA Based Shapers

A semi-Gaussian shaper principal schema is shown in Fig.1. A high-pass filter (HPF) sets the duration of the pulse by introducing a decay time constant. The low-pass filter (LPF), which follows, increases the rise time to limit the noise bandwidth. Although pulse shapers are often more sophisticated and complicated, the CR-RC^n shaper contains the essential features of all pulse shapers, a lower frequency bound and an upper frequency bound and it is basically a (n+1) order band pass filter (BPF), where n is the integrators number (n is called shaper order). The transfer function of an S-G pulse shaper consisting of one CR differentiator and n integrators is given by:

[1] T. Noulis is with Electronics Laboratory of Physics Department, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece, e-mail: tnoul@physics.auth.gr.
[2] C. Deradonis is with Electronics Laboratory of Physics Department, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece.
[3] S. Siskos is with Electronics Laboratory of Physics Department, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece, e-mail: siskos@physics.auth.gr.

$$H(s) = \left(\frac{s\tau_d}{1+s\tau_d}\right)\left(\frac{A}{1+s\tau_i}\right)^n = H(s)_{BPF} \quad (1)$$

where τ_d is the time constant of the differentiator, τ_i of the integrators, and A is the integrators dc gain. The number n of the integrators is called shaper order. Peaking time is the time that shaper output signal reaches the peak amplitude and is defined by $\tau_s = n\tau_i$. The order n and peaking time τ_s, depending on the application, can be predefined by the design specifications or not.

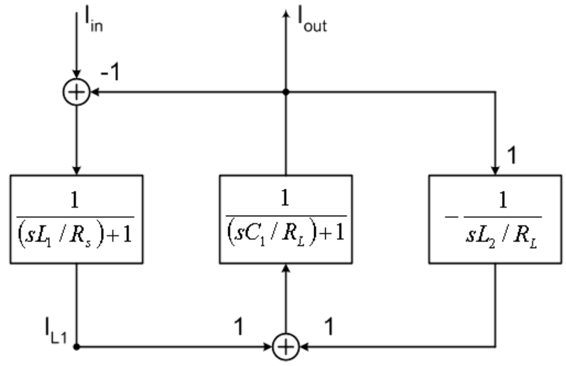


Fig. 2. Signal flow graph of a current mode 3rd order RLC filter.

TABLE I
SHAPERS PASSIVE ELEMENTS AND TRANSCONDUCTANCES

LC Ladder		**Leapfrog**		**Cascade**	
R_S	100 kΩ	C_1	11.8 pF	R_1	11 kΩ
R_L	100 kΩ	C_2	13.3 pF	C_1	4.2 pF
C_1	11.7 pF	C_3	3.7 pF	R_2	100 kΩ
C_2	11.7 pF	g_{m1}	11.4 μA/V	C_2	5 pF
C_3	11.7 pF	g_{m2}	24 μA/V	R_3	100 kΩ
g_{m1}	24 μA/V	g_{m3}	500.4 nF	C_3	5 pF
g_{m2}	500.4 nA/V			g_{m1}	37.8 μA/V
				g_{m2}	500.4 nF

Using the above shaper model and the respective passive RLC equivalent two port circuit, the signal flow graph (SFG) of a 2nd order current mode S-G shaper is extracted (Fig. 2). From the above SFG and using the Leapfrog (LF), the Ladder simulation method by element replacement and the typical cascade filter technique, three 2nd order shapers are designed. The basic circuit block of all three shapers is an operational transconductance amplifier (OTA). Figs. 3, 4 and 5 show the LC ladder shaper with simulation by element replacement, the leapfrog shaper and the OTA based cascade method shaper respectively. An OTA based inductance and a capacitor simulator are used in the cascade OTA shaper and the LF shaper respectively, in order to implement fully integrated systems [14].

The passive elements and the OTA transconductances of all the above configurations are given in Table I.

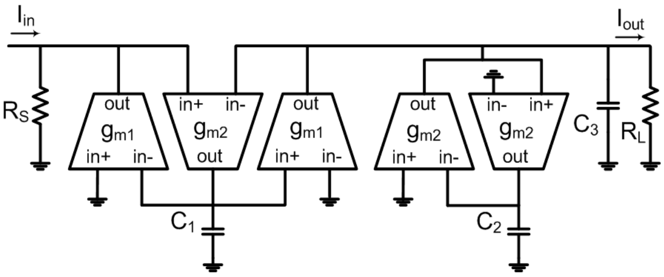


Fig. 3. OTA based LC ladder shaper using simulation by element replacement.

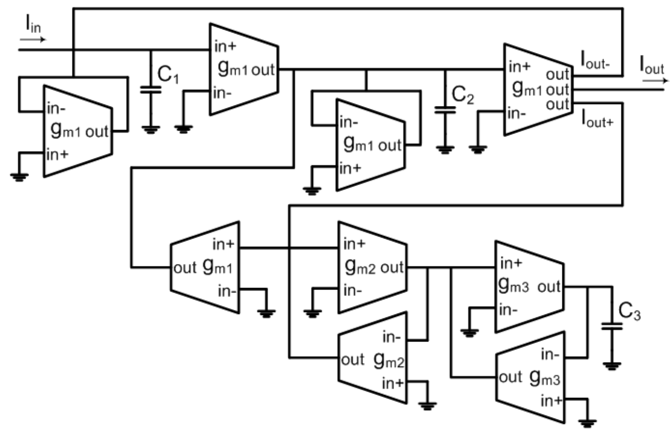


Fig. 4. OTA based leapfrog shaper with capacitance simulator.

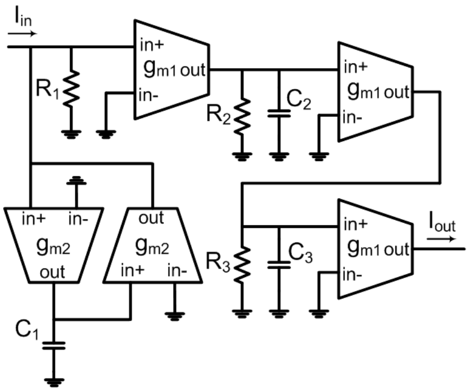


Fig. 5. OTA cascade method shaper with inductor simulator.

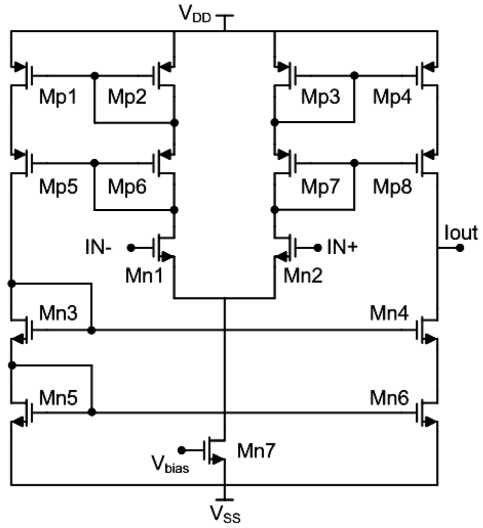


Fig. 6. CMOS operational transconductance amplifier.

TABLE II
OTA BASED SHAPERS PERFORMANCE CHARACTERISTICS

	Cascade	Leapfrog	LC Ladder
Maximum bias current	35.8 μA	4.7 μA	2.2 μA
Minimum bias current	83.6 nA	83 nA	83.5 nA
Power consumption	787 μWatt	648 μWatt	133 μWatt
$i_{in,peak\text{-}40dB}$ (THD = 1%)	1.3 μA	1.6 μA	8.4 μA
Output noise (rms)	77 nA	71 nA	71 nA
Dynamic Range (DR)	22 dB	22 dB	32 dB

III. SIMULATION RESULTS

A typical CMOS operational transconductance amplifier was designed in order to implement the above shaper structures. The respective OTA schematic is shown in Fig. 6. This OTA is implemented using a CMOS configuration with a cascode structure [15].

The OTA based shapers were designed in order to provide the same operating bandwidth (*BW*) at 230 kHz. Their frequency response is given in Fig. 7. The difference of the output current amplitude signal and in particular the lower gain of the LC Ladder and the Leapfrog structures is caused by the fact that the specific filter design methods reduce the output signal amplitude by half.

The total performance characteristics of each shaper system are listed in Table II.

The higher maximum bias current is observed in the OTA cascade shaper and the lower minimum in almost the same in all three configurations. All the shapers appear to be low power, but the LC Ladder shaper provide the optimum power consumption performance. Additionally, the cascade and the LF structure provide a dynamic range equal to 22 dB, far lower to the 32 dB value of the LC Ladder topology. Concerning the noise performance, all the structures appear to have low rms output noise, with the cascade shaper being slightly worse in comparison to the other two. The above characteristics, and in relation to the fact that the LC Ladder architecture is the optimum in terms of the total harmonic distortion, render it suitable for applications where the DR is required to be very high, and the noise and power consumption limits are the main factors that determine the application as in the readout front ends.

Fig. 8 shows the signal to noise ratio of the three OTA based shapers.

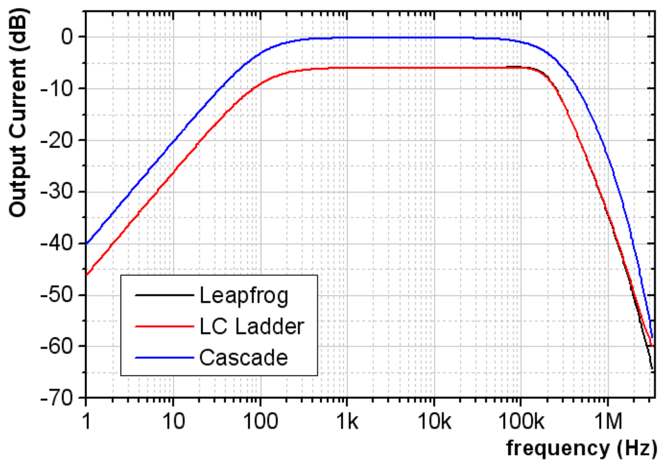


Fig. 7. Shapers frequency response.

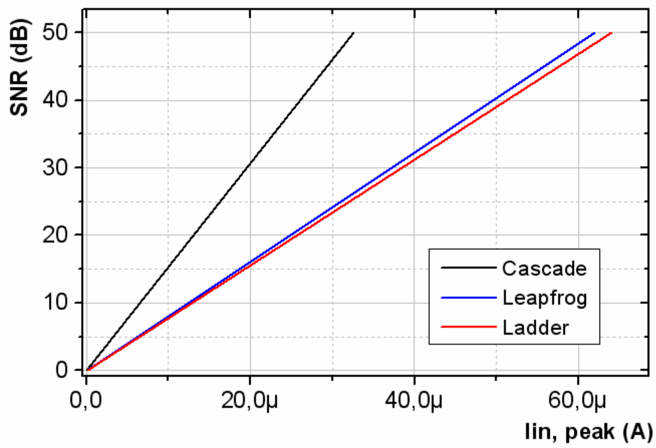


Fig. 8. Shapers signal to noise ratio.

IV. CONCLUSION

In this paper, a detailed examination of OTA based semi-Gaussian shapers suitable for readout applications is performed. Specifically, three different current mode shaper structures are designed using advanced filter design methods such as the Leapfrog and the Ladder LC technique by element replacement. All the methods used in this work provide fully integrated configurations and not discrete systems. An OTA was designed in order to be used in the implementation of the above shapers. The filter configurations are analytically compared in relation to power consumption, total harmonic distortion, dynamic range and noise performance. The OTA LC Ladder architecture is proved to be the optimum in low energy radiation detection applications, according mainly to its output noise and power consumption. Consequently, the LC Ladder method is the most suitable in order to design the shaper of an integrated readout front end system in which a current mode preamplifier is used.

REFERENCES

[1] B. Kriger, I. Kipnis, B. A. Ludewigt, "XPS: A multi-channel preamplifier-shaper IC for X-ray spectroscopy", *IEEE Transactions on Nuclear Science*, vol. 45, no. 3, pp. 732-734, June 1998.

[2] M. Pedrali-Noy et al., "PETRIC – a positron emission tomography readout integrated circuit", *IEEE Transactions on Nuclear Science*, vol. 48, no. 3, pp. 479-483, June 2001.

[3] B. Krieger et al., "An 8x8 pixel IC for X-ray spectroscopy", *IEEE Transactions on Nuclear Science*, vol. 48, no. 3, pp. 493-497, June 2001.

[4] Erik H. M. Heijne, Pirre Jarron, "A low noise CMOS integrated signal processor for multi-element particle detectors", *Digest of Papers ESSCIRC*, pp. 68-69, 1988.

[5] F. S. Goulding, D. A. Landis, "Signal processing for semiconductor detectors", *IEEE Transactions on Nuclear Sci*ence, vol. NS-29, pp. 1125-1141, June 1982.

[6] W. Sansen, Z. Y. Chang, "Limits of low noise performance of detector readout front ends in CMOS technology*", IEEE Transactions on Circuits and Syst*ems, vol. 37, no. 11, pp. 1375-1382, November 1990.

[7] Z. Y. Chang, W. Sansen, "Effect of 1/*f* noise on the resolution of CMOS analog readout systems for microstrip and pixel detectors", *Nuclear Instruments and Methods in Physics Research,* vol. 305, pp. 553-560, 1991.

[8] T. Noulis, S. Siskos, G. Sarrabayrouse, "Analysis of input and feedback capacitances effect on low noise preamplifier performance for X rays silicon stripe detectors", *19th Conference on Design of integrated circuits and systems*, Bordeaux, France, November 24-26, 2004.

[9] T. Noulis, S. Siskos, G. Sarrabayrouse, "Effect of technology on the input transistor selection criteria of a low noise preamplifier", *IEEE Mediterranean Electrotechnical Conference*, Dubrovnik, Croatia, May 12-15, 2004.

[10] J. Wulleman, "Current mode charge pulse amplifier in CMOS technology for use with particle detectors", *Electronics Letters*, vol. 32, No. 6, pp. 515-516, March 1996.

[11] Fei Yuan, "Low voltage CMOS current-mode preamplifier: Analysis and design", *IEEE Transactions on Circuits and Syst*ems, vol. 53, no. 1, pp. 26-39, January 2006.

[12] T. L. Deliyannis, Yichuang Sun, Kel Fidler, *Continuous-time Active Filter Design*, Florida, CRC Press LLC, 1999.

[13] R. Schaumann, Mac.E van Valkenburg, *Design of Analog Filters*, New York, Oxford, Oxford University Press, 2001.

[14] R. Senani, "Novel lossless synthetic floating inductor employing a ground capacitor", *Electronics Letters*, vol. 18, no. 10, pp. 413-414, May 1982.

[15] T. Noulis, C.Deradonis, S. Siskos, G. Sarrabayrouse, "Novel fully integrated OTA based front end analog processor for X – rays silicon strip detectors", accepted for publication to *13th IEEE Mediterranean Electrotechnical Conference*, Malaga, Spain, May 16-19, 2006.

A Space Application Current Mode CMOS Low Noise Preamplifier for X-rays Detection System

T. Noulis[1] and S. Siskos[2]

Abstract — **A systematic design guideline is presented for the noise optimization of a current mode CMOS preamplifier suitable for space applications X-rays silicon strip detectors. A novel current architecture preamplifier is implemented using a third generation current conveyor. The preamplifier implement-tation is designed in AMS 0.35 μm process. Analysis is supported by simulation results, which confirm its great performance mainly in terms of the total output noise and the power consumption. This current structure appears to be preferable to the typical voltage one due to its lower noise performance and its noise independence on the detector capacitance variations.**

Keywords — **current mode, low noise preamplifier, X-rays detection system.**

I. INTRODUCTION

Radiation detection has been increasingly developed in various fields of radioactivity control, medical imaging and space science. In X-rays detection front-end systems CMOS technologies are widely used since they ensure the required integration density and seem very promising in terms of radiation hardness. A block diagram of such a detection system is shown in Fig. 1. A reversely biased diode (Si or Ge) detects radiation events by generating electron–hole pairs proportional to the absorbed energies. A low noise charge sensitive preamplifier (CSA) is widely used at the front end due to its low noise configuration and insensitivity of the gain to the detector capacitance variations. The generated charge Q is integrated onto a feedback capacitance. This is fed to a bandpass filter (shaper) where pulse shaping is performed to optimize the signal to noise (S/N) system ratio. The resulting output signal is a narrow pulse suitable for further processing [1], [2].

The noise performance of the amplification stage determines the overall system noise and should be therefore optimized. A voltage mode folded cascode architecture is commonly used because of its low input capacitance [1]-[5]. However, a current mode structure could be an attractive alternative to the more typical voltage mode ones, since the signal is processed in the current domain, avoiding charging and discharging of the parasitic capacitance to high voltage levels and keeping the internal nodes of the circuit at low impedance values.

In this work, a current mode preamplifier configuration is proposed. A CMOS third generation current conveyor (high gain current conveyor CCII∞) is designed and noise minimization techniques are presented in order to achieve optimum noise performance.

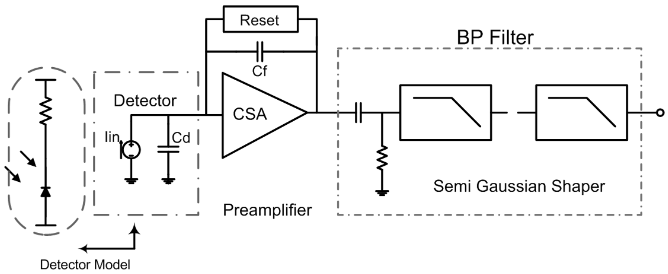


Fig. 1. Block diagram of a front-end readout system.

II. NOISE MINIMIZATION TECHNIQUES

A current mode preamplifier can be implemented using a current conveyor and particularly a high gain one (CCII∞). The input stage of almost all current mode feedback amplifiers is either a positive or a negative second generation current conveyor [6], [7]. Therefore, a conveyor noise model (Fig. 2) can be used to simplify noise calculations. The noise sources in a dual output MOS CCII∞ are shown in Fig. 3, where the only MOSFETs not included in the conveyor noise model are the output stage transistors (M1, M2, MSS1 and MSS2).

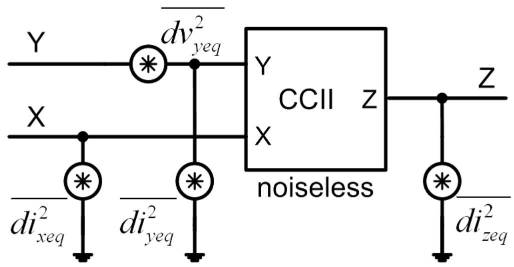


Fig. 2. Equivalent noise sources of a second generation current conveyor.

[1]T. Noulis is with Electronics Laboratory of Physics Department, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece, e-mail: tnoul@physics.auth.gr.
[2]S. Siskos is with Electronics Laboratory of Physics Department, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece, e-mail: siskos@physics.auth.gr.

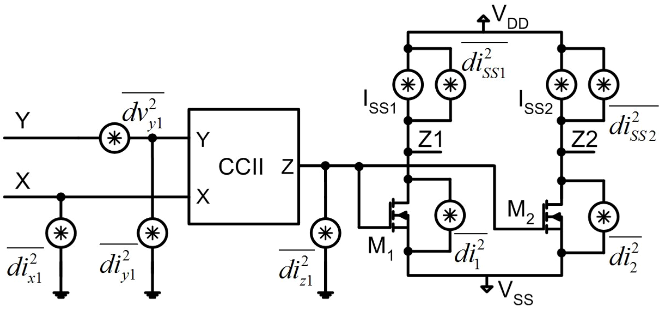


Fig. 3. Noise sources in a dual output MOS high gain current conveyor.

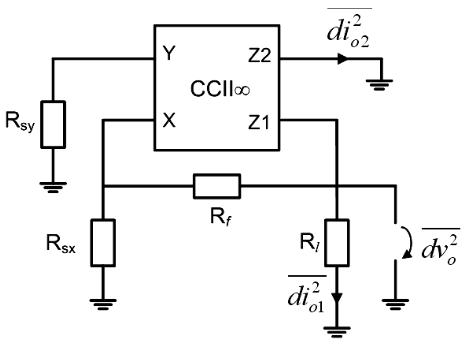


Fig. 4. Noise test set-up for a dual output high gain current conveyor.

By using the circuit configuration of Fig. 4, the noise of most high gain current conveyor applications can be evaluated. Since the closed loop input impedance at the X-terminal is in most cases significantly lower than R_{sx}, the output noise voltage $\overline{dv_o^2}$ at the Z1-output can be approximated at low frequencies as:

$$\overline{dv_o^2} \approx \left(1+\frac{R_f}{R_s}\right)^2 \left(\overline{dv_{y1}^2} + R_{sy}^2 \overline{di_{y1}^2} + 4kTR_{sy}df\right) + R_f^2\left(\overline{di_{x1}^2} + \overline{di_{z1}^2} + \frac{4kT}{R_{sx}\|R_f}df\right) + \frac{R_f^2}{\left(1+G_i+\frac{R_f}{R_l}\right)^2}\left(\overline{di_1^2} + \overline{di_{SS1}^2} + \frac{4kT}{R_f\|R_l}df\right) \quad (1)$$

where k is the Boltzmann constant and T is the temperature.

In high gain current conveyors the open-loop current gain G_i is high and therefore noise contribution of output transistors (M1 and MSS1) can be neglected unless very high frequencies are considered. If the output current noise is required, it can be obtained by letting:

$$\overline{di_{o1}^2} = \frac{\overline{dv_o^2}}{R_l^2} \quad (2)$$

Contrary, the replica output is outside the feedback loop and transistors M2 and MSS2 noise contribution can not be neglected, and thus the output current noise is:

$$\overline{di_{o2}^2} = \frac{\overline{dv_o^2}}{\left(R_f\|R_l\right)^2} + \overline{di_2^2} + \overline{di_{SS2}^2} \quad (3)$$

Because the replica output is typically used in the closed loop current conveyor configuration without resistive feedback, the output noise current is given by:

$$\overline{di_{o2}^2} = G_{icl}^2\left(\overline{di_{x1}^2} + \overline{di_{z1}^2} + \overline{di_1^2} + \overline{di_{SS1}^2}\right) + \overline{di_2^2} + \overline{di_{SS2}^2} \quad (4)$$

where G_{icl} is the closed loop current gain depending on the aspect ratios of the output transistors. As a result, the output transistors (M2 and MSS2) noise contribution is more significant whereas the conveyor input voltage noise is omitted. This situation differs from the normal one with high gain voltage mode opamps where normally only input transistors noise needs to be considered. Consequently, the noise model for a high gain current conveyor should also include output current noise sources, as depicted in Fig. 5 [7]. In this model, the equivalent noise sources at Y-terminal are identical to the noise sources of the input CCII-. However, the remaining noise sources are divided between the X-terminal and the two Z-outputs accordiong to the following equations:

$$\overline{di_{xeq}^2} = \overline{di_{x1}^2} + \overline{di_{z1}^2} \quad (5)$$

$$\overline{di_{zeq1}^2} = \overline{di_1^2} + \overline{di_{SS1}^2} \quad (6)$$

$$\overline{di_{zeq2}^2} = \overline{di_2^2} + \overline{di_{SS2}^2} \quad (7)$$

A simple optimization technique in order to design a low noise current conveyor implies the use of conveyors with a current gain from node X to node Z, since this reduces the ouput transistors noise contribution and can be controlled in current mode opamps. It would also be essential to design low g_m current mirror structures, high g_m structures for the X input stage and the Y to X level shifter stage.

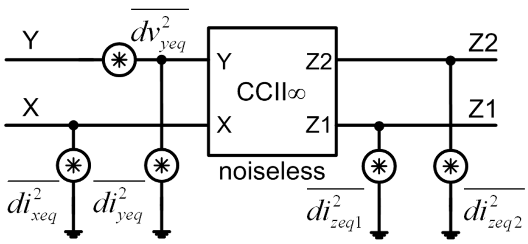


Fig. 5. The equivalent noise sources of a multi-input high gain current conveyor.

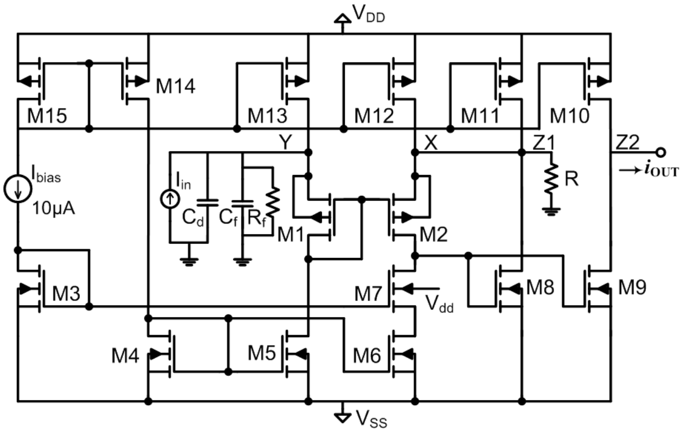


Fig. 6. Preamplifier current mode implementation with a high gain current conveyor.

III. Design of Low Noise Preamplifier

The preamplifier integrates the input signal from a silicon strip detector. The circuit has to be designed to work with current (dc) couple detectors and should be able to supply a current in the range from a few picoamperes to a few nanoamperes through the feedback loop in order to match the detector leakage current.

The current mode preamplifier using a CCII∞ implementation with an n-well CMOS process is shown in Fig.6. This conveyor uses an input voltage follower structure typical to class-A second generation positive current conveyors [8], [9]. The input voltage follower is implemented with PMOS transistors M1, M2 and the input voltage swing can be maximized by using floating n-wells for these transistors. Since the offset voltage between Y- and X-terminals remains minimal, this high gain conveyor can be used as a drop-in replacement for voltage mode operational amplifier [7]. In certain applications, the input impedance may be too low. Its performance can be improved by using cascode current sources rather than the transistors M5, M13 and M12. This high-gain conveyor has two current outputs Z1 and Z2. A CCII+ structure is implemented from this amplifier by connecting the Z1 output to X-terminal. When only one current output is needed, the output current swing can be doubled by joining the two outputs together. In order the circuit to operate as a non-inverting loss integrator a feedback capacitor C_f is connected in parallel with a resistor R_f in node Y and a resistor R is connected between node X and the ground. This circuit design was based on the respective noise optimization criteria described above.

TABLE I

Design Specifications

Detector Diode PIN (Si) – Preamplifier	
Detector capacitance	C_d =2 - 5 pF
Leakage current	I_{leak} = 10 pA
Q collected per event	$Q_{typical}$=28000e$^-$ Q_{max}=300000e$^-$
Time needed for 90% of total Q	300 nsec
Temperature	-40^0 C
Power consumption per channel	< 8 mW

The above preamplifier circuits have been designed for a low energy X-rays strip detector for space applications. The design specifications are listed in Table I.

IV. Simulation Results

The preamplifier configuration was simulated using HSPICE (BSIM3V3.2 Level 49) and designed in 0.35 μm process by Austria Micro Systeme (AMS). Voltages V_{DD} and V_{SS} are 2.5 V and -2.5 V respectively, feedback capacitance C_f was 10 pF, R_f and R were 5 kΩ and 25 Ω. The bias current I_{bias} was selected to be 10 μA in order to achieve lower noise performance. The current output signal of the CCII∞ preamplifier implementation is shown in Fig. 7. The fact that the current CSA output signal is not inverted does not complicate the signal processing, since the use of a respective CCII current mode shaper structure (with inverting Z- and non-inverting Z+ outputs) would provide a more flexible system design. Table II summarizes the performance of the circuit.

TABLE II

Performance Characteristics

	Current Preamplifier
Charge time	200 ns
Discharge time	0.46 μs
Power consumption	2.5 mW
Gain	45 dB
fp	1.92 MHz
enc_{total}/τ_s =10μs C_d = 2pF	28 e$^-$
Charge time	200 ns

As it seems from Fig. 7 the preamplifier output increases to a maximum value when the charge of the feedback capacitance ends. Charge time (collection of input signal by CSA) is very low, which proves that the circuit is fast enough to perceive signals of 300 ns. The discharge time is also relatively very low since in this application, two successive events have a time distance of about 1 ms. The requested power consumption limits are also satisfied since the current mode preamplifier consumes 2.5 mW, providing the shaper the capability to consume 5.5 mW. The CCII∞ opamp, appears to have lower gain performance in relation to typical voltage configurations but in a much more wider bandwidth. Moreover the current mode CSA has very low *enc*. Fig. 8 shows the noise spectral density of the current configuration.

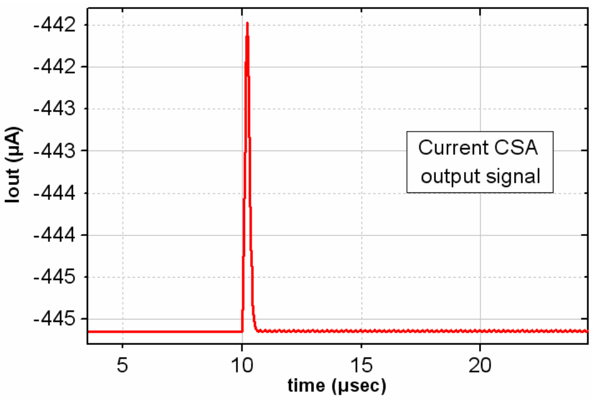


Fig. 7. Current mode preamplifier output signal.

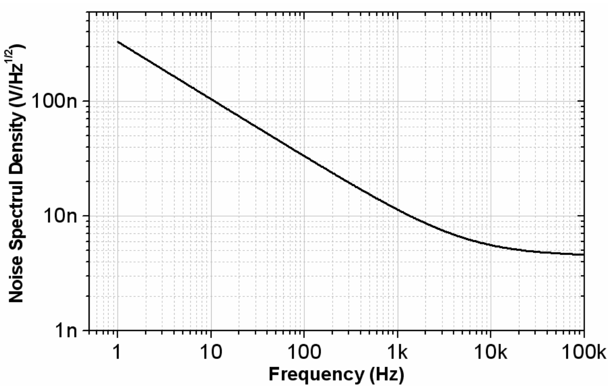


Fig. 8. Preamplifier noise spectral density.

Furthermore, a study was made as to examine and contrast the dependence of noise performance on the detector capacitance. It was calculated that in a current mode CSA the total *enc* (detector readout system noise performance is expressed as the equivalent noise charge (*enc*) [1], [2]) is constant with C_d. Specifically, this corresponds to a constant equivalent noise charge of 28 e^- at a detector with a capacitance of 2 pF.

V. Conclusion

In this paper, a noise optimization technique for a current mode preamplifier was presented. In particular, low noise design criteria were applied for the first time in a current mode structure suitable for X-rays detection systems. It is confirmed by simulations that with a current mode preamplifier very low *enc* levels can be provided. Moreover, a great advantage of the current mode configuration is the opportunity to implement analog systems with high speed and wider frequency bandwidth. It is also noticeable that the detector capacitance increase does not result to higher noise performance, as it happens with the typical voltage mode architecture. However, it should not be neglected that the amplification gain performance appears to be worse in relation to the typical folded cascode architecture. The above factor renders the current mode preamplifier optimum especially for large capacitance detection applications.

References

[1] Z. Y. Chang, W. Sansen, “Effect of 1/f noise on the resolution of CMOS analog readout systems for microstrip and pixel detectors,” *Nuclear Instruments and Methods in Physics Research*, vol. 305, pp.553-560, 1991.

[2] W. Sansen, Z. Y. Chang, “Limits of low noise performance of detector readout front ends in CMOS technology”, *IEEE Transactions on Circuits and Systems*, vol. 37, no. 11, pp.1375-1382, November 1990.

[3] T. Noulis, S. Siskos, G. Sarrabayrouse, “Analysis of input and feedback capacitances effect on low noise preamplifier performance for X rays silicon stripe detectors”, *19th Conference on Design of Integrated Circuits and Systems*, Bordeaux, France, November 24-26, 2004.

[4] E. F. Tsakas, A. N. Birbas, “Noise optimisation for the design of a reliable high speed X-rays readout integrated circuit”, *Microelectronics Reliability*, vol. 40, pp. 1937-1942, February 2000.

[5] T. Noulis, S. Siskos, G. Sarrabayrouse, “Effect of technology on the input transistor selection criteria of a low noise preamplifier”, *IEEE Mediterranean Electrotechnical Conference*, Dubrovnik, Croatia, May 12-15, 2004.

[6] E. Bruun, “Analysis of the noise characteristics of CMOS current conveyors”, *Analog Integrated Circuits and Signal Processing*, vol.12, pp. 71-78, 1997.

[7] K. Koli, *CMOS Current Amplifiers: Speed Versus Nonlinearity*, Finland, Helsinki University of Technology, 2000.

[8] D. Johns, K. Martins, *Analog Integrated Circuit Design*, New York, Wiley Interscience Publications, 1997.

[9] K. Koli, K. Halonen, “Inverting transimpendence amplifier for current- and voltage-mode applications”, *Proceedings of the 11th European Conference on Circuit Theory and Design*, Davos, August 1993.

Numerical modelling of the two-state lasing in 1.55µm (113)B InAs/InP quantum dot lasers for optical telecommunications

K. Veselinov[1], F. Grillot[1], Al. Bekiarski[2], J. Even[1], S. Loualiche[1]

Abstract — **Two rate equation models (REM) based on the intraband energy competition and efficient carrier relaxation are presented. The comparison between these two theoretical approaches leads to qualitative understanding of the origin of the two-state lasing in 1.55 µm InAs/InP quantum dot lasers.**

Index Terms — **Quantum dots, semiconductor lasers, rate equa- tions, optical telecommunications.**

I. Introduction

While optics has proven to be the most practical response to the high traffic rate demand for long-haul transmission, its extension to the metropolitan networks down to the home remains an open challenge. The implementation of optics at transmission rate where other technical solutions exist requires cost reduction. As a consequence, semiconductor lasers based on low dimensional heterostructures such as quantum dots (QDs) laser are very promising. Indeed, QDs structures have attracted a lot of attention in the last decade since they exhibit many interesting and useful properties such as low threshold current, temperature insensitivity, chirpless behavior and optical feedback resistance. As a result, thanks to QDs lasers, several steps toward cost reduction can be reached such as: improving the laser resistance to temperature fluctuation in order to remove temperature control elements, or designing feedback resistant laser for isolator-free and optics-free module. Most investigations reported in the literature deal with In(Ga)As QDs grown on GaAs substrates [1,2]. Also numerous theories about carrier dynamics in these structures have been introduced [3,4]. It is however important to stress that In(Ga)As / GaAs QDs devices do not allow a laser emission above 1.35 µm which is detrimental for optical transmission. In order to reach the standards of long-haul transmissions, 1.55 µm InAs QD lasers on InP substrate have been developed. In this paper, based on a rate equations model, carrier dynamics is at first investigated. Numerical results reporting the double laser emission are then shown exhibiting two different behaviors. All these recent results will be reviewed in the following.

[1] K. Veselinov, F. Grillot, J. Even, and S. Loualiche are with the National Institute of Applied Sciences, UMR CNRS FOTON, INSA, 20 Avenue des Buttes de Coesmes, CS 14315, 35043 RENNES Cedex, FRANCE (e-mail: kiril.veselinov@ens.insa-rennes.fr; frederic.grillot@insa- rennes.fr; jacky.even@insa-rennes.fr; slimane.loualiche@insa-rennes.fr)

[2] A. Bekiarski is with the Technical University of Sofia, 8, Kliment Ohridski, St., Sofia-1000, BULGARIA (e-mail: aabbv@tu-sofia.bg)

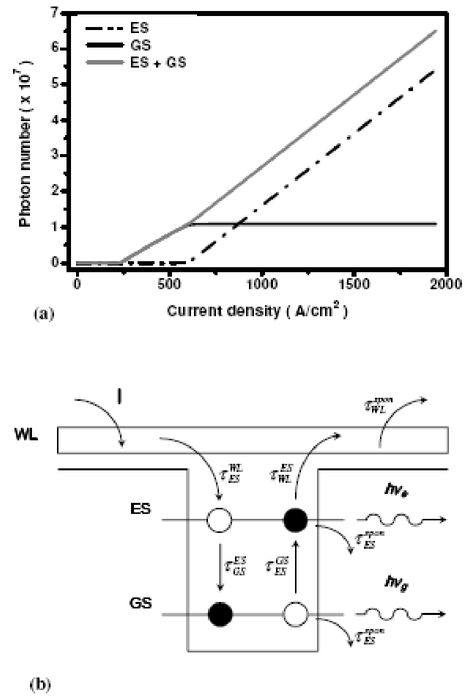


Fig. 1. Calculated photon number in the active region as a function of the injected current density (a). Schematic representation of the intraband energy competition model (b).

II. Two-state lasing competition

A numerical model based on a set of rate equations is used to study carrier dynamics in a QD laser. Its active region consists of a QD ensemble, where different dots are inter-connected by a wetting layer (WL) as shown in Fig. 1(b). In this ensemble, two energy levels are assumed: the ground state (GS) and the excited state (ES) [7]. It is also assumed that there is only one QD ensemble, i.e., all dots have the same size meaning the inhomogeneous broadening is neglected. Electrons and holes are considered to be captured and emitted in pairs. Carriers are supposed to be injected directly from the contacts into the WL levels, so the barrier dynamics is not taken into account in the model. Calculated photon number in the cavity is reported in Fig. 1(a) as a function of the injection current density. Once the threshold of the ES lasing is reached, the emission of the GS saturates and the ES emission increases linearly. This behaviour has been already reported for the InAs/GaAs system [4] as a two-state lasing competition between the GS and ES due to the finite

GS relaxation time. However, this approach do not match with recent experimental results obtained on the InAs/InP system which have shown no saturation of the GS lasing [3].

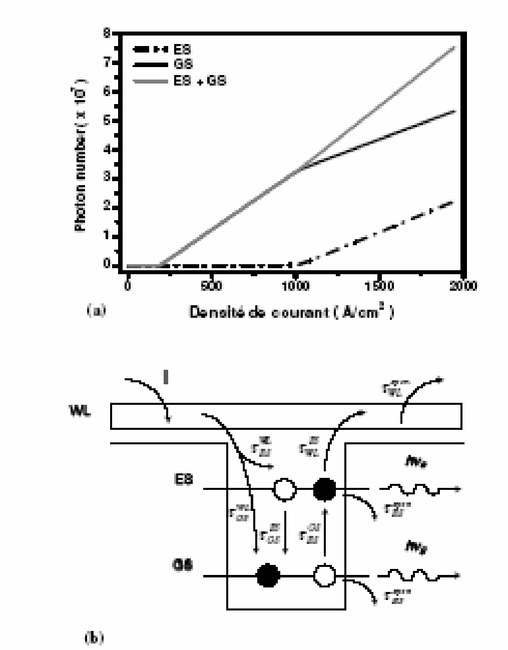


Fig. 2. Calculated photon number in the active region as a function of the injected current density (a). Schematic representation of the efficient carrier relaxation model (b).

III. EFFICIENT CARRIER RELAXATION

Using the same simulation parameters, carriers are this time captured into the ES or directly into the GS within the same time $\tau_{WL} = \tau_{WL}$ (Fig. 2(b)) whereas they can also relax from the ES to the GS. It is assumed that at low injection rate, the relaxation is phonon-assisted while the Auger effect dominates when the injection gets larger [6]. In order to investigate the properties of the InAs/InP(113)B QD device under electrical pumping, steady-state solutions of the REM have been determined. Simulation of the lasing performance is shown in Fig. 2(a) where the the calculated ES and GS photon number is reported as function of the injection current density showing two thresholds corresponding to the two laser emissions. A single longitudinal mode is assumed and the gain nonlinearities are not taken into account. When the ES stimulated emission appears, only a slight decrease of the GS slope efficiency is predicted. At the same time, the global slope efficiency increases. This behaviour is different from the one associated to the InAs/GaAs QD laser. Here, the double laser emission seems to result from the efficient carrier relaxation into the GS [7]. Although the competition between GS and ES transitions of different QDs is not taken into account, these numerical results are in good agreement with experimental ones recently reported for an optical pumped InAs/InP diode laser [5].

IV. CONCLUSION

In summary, calculated photon number has been predicted versus the injected current density. Using a rate equation model based on the intraband energy competition and on an efficient carrier relaxation, two theoretical approaches have been investigated. On one hand, numerical results describing carrier dynamic behaviour have been presented confirming previously observed phenomena for InAs/GaAs system. On the other hand, it has been shown that the direct relaxation channel included in the model matches very well the different experimental results already published and leads to qualitative understanding of InAs/InP QD lasers. These results show that laser characteristics such as threshold current or external efficiency can be predicted. It is of first importance for the realization of QD semiconductor devices for optical telecommunications.

ACKNOWLEDGMENT

This work was supported by ePIXnet (European Network of Excellence on Photonic Integrated Components and Circuits), SANDiE network of excellence (Self-Assembled semiconduc- tor Nanostructures for new Devices in photonics and Electron- ics) and CREFID (Centre Regional Francophone d'Ingenierie pour le Developpement).

REFERENCES

[1] K. Mukai, Y. Nakata, K. Otsubo, M. Sugawara, N. Yokoyama and H. Ishikawa, 1.3-µm CW Lasing of InGaAs-GaAs Quantum Dots at Room Temperature with a Threshold Current of 8 mA, IEEE Photonics Technology Letters, 11 1205, 1999.

[2] K. T. Tan, C. Marinelli, M. Thompson, A. Wonfor, M. Silver, R. Sellin, R. Penty, I. White, M. Lammlin, N. Ledentsov, D. Bimberg, A. Zhukov, V. Ustinov and A. Kovsh, High Bit Rate and Elevated Temperature Data Transmission Using InGaAs Quantum -Dot Lasers, IEEE Photonics Technology Letters, 16 1415, 2004.

[3] M. Sugawara, N. Hatori, H. Ebe, Y. Arakawa, T. Akiyama, K. Otsubo, and Y. Nakata, Modelling room-temperature lasing spectra of 1.3 µm self- assembled InAs/GaAs quantum-dot lasers: Homogeneous broadening of optical gain under current injection, Journal of Applied Physics, vol. 97, 043523, 2005.

[4] A. Markus, J. X. Chen, C. Paranthoen, A. Fiore, C. Platz, and O. Gauthier- Lafaye, Simultaneous two-state lasing in quantum-dot lasers, Ap. Phys. Let., vol 82, no. 12, pp. 1818-1820

[5] C. Platz, C. Paranthoen, P. Caroff, N. Bertru, C. Labbe, J. Even, O. De- haese, H. Folliot, A. Le Corre, S. Loualiche, G. Moreau, J. C. Simon, and A. Ramdane, Comparison of InAs quantum dot lasers emitting at 1.55 µm under optical and electrical injection, Semicond. Sci. Technol., vol. 20, pp. 459-463, 2005.

[6] P. Miska, C. Paranthoen, J. Even, O. Dehaese, H. Folliot, N. Bertru, S. Loualiche, M. Senes and X. Marie, Optical spectroscopy and mod- elling of double-cap grown InAs/InP quantum dots with long wavelength emission, Semicond. Sci. Technol., vol. 17, pp. L63-L67, 2002.

[7] P. Miska, K. Veselinov, F. Grillot, J. Even, C. Platz, P. Caroff, C. Cornet, C. Paranthoen, N. Bertru, C. Labbe, O. Dehaese, H. Folliot, A. Le Corre, S. Loualiche, G. Moreau, J. C. Simon, X. Marie, and A. Ramdane, Carrier dynamics and saturation effect in (311)B InAs/InP quantum dot lasers, in International Workshop on PHysics and Applications of SEmiconductor LASERs, March 2005, Metz (France).

Application of Electro-Thermal Analogy for Complex Simulation of Hybrid Power Controllers

Grzegorz Błąd [1], Dariusz Klepacki [1], Jerzy Potencki [1] and Anna Andonova [2]

Abstract – **Temperature plays a very important role in proper operation of microelectronic circuits. It determines exploitation parameters and – in the most cases – reliability of circuit or whole system. From this reason, taking into consideration the thermal problems is strongly recommended on the beginning of design process. The paper presents some aspects of describing of thermal model of the thick-film microcircuits with equivalent models (based on RC elements) as well as example of the complex analysis of hybrid power controller of household equipment in PSPICE program.**

Keywords – **temperature field, PSPICE simulation**

I. Introduction

Temperature plays a very important role in proper operation of microelectronic circuits. It determines exploitation parameters and – in the most cases – reliability of circuit or whole system [2]. From this reason, taking into consideration the thermal problems is strongly recommended on the beginning of design process.

The thick-film microcircuits are very complicated objects to their formal description from point of view of heat transfer. The complexity of heat exchange mechanism makes the mathematical analysis only approximation of the real-world conditions. In reality, only numerical calculations can solve the systems of differential equations with very complicated boundary and initial conditions.

Using the mathematical analogy of equations (in stationary and dynamic states) which describe the temperature field and electric potential field, it is possible to create the electrical equivalent circuit of thermal system which can be simulated in PSPICE program. The proposed method of temperature simulation using equivalent RC models is based on well – known Beuken theory and is much more easier and "intuitive" for designers of the electronic equipment. [1-7].

II. RC modelling procedure

Based on Fourier theory of heat conduction, Beuken in 1934 proved that using electrical model of the analyzed structure is possible to obtain the information about its temperature. In such idea (based on combined connections of the RC elements) voltage is the analogue of temperature, and resistance and capacitance are analogous of thermal resistance and capacitance, respectively.

The division of real-world circuit on differential elements in the modeling procedure is required. It is nearly the same to finite differences method.

The next step is concentration of thermal capacity in the node and thermal resistances between particular nodes. As results the resistors' network is created with terminated capacitors.

Thermal resistance is determined by thermal conductivity in the selected area of analyzed object and it can be expressed by formula:

$$R_{T\lambda}(\lambda(T)) = \frac{l}{\lambda(T)\cdot F} \tag{1}$$

where: $\lambda(T)$ is the material thermal conductivity dependent on temperature, l – material length and F means material area excited by heat flux.

For boundary elements in RC thermal model of the microcircuit the calculation of thermal resistance which characterized the heat exchange with environment is required. In general case it can be calculated from equation:

$$R_T(\alpha(T)) = \frac{1}{\alpha(T)\cdot F} \tag{2}$$

where: $\alpha(T)$ is the total coefficient of heat exchange with environment (by convection and radiation).

The value of α coefficient is very hard to determination because it is strongly dependent on many physical parameters. The most often it is determined from experimental measurements or calculated from criterion numbers (Nusselt and Rayleigh).

Thermal capacitance is determined by specific heat of the selected area of analyzed object as well as density of material. In the general case it can be calculated from expression:

$$C_T(T) = c_p(T)\cdot\rho\cdot V \tag{3}$$

where: c_p is specific heat, ρ – material density and V means volume.

Such defined elementary RC elements are the basis for creation RC network which represents analyzed thermal object. The main idea of the multilayer microcircuits modeling is presented in Fig. 1.

[1] Department of Electronic and Telecommunications Systems, Rzeszow University of Technology, W. Pola 2, 35-959 Rzeszow, Poland, E-mail: gblad@prz.rzeszow.pl

[2] Department of Microelectronics, University of Sofia, FETT Sofia 1797, Bulgaria, E-mail: ava@ecad.tu-sofia.bg

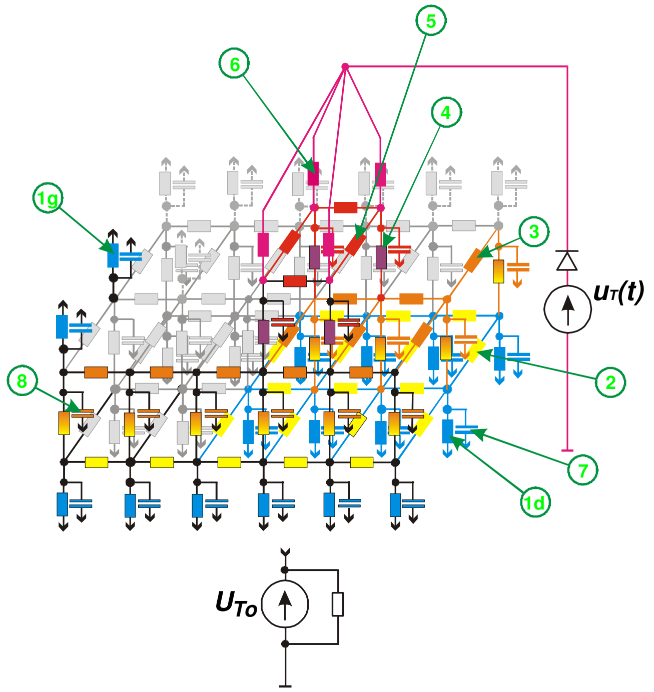


Fig. 1. RC model of multilayer structure: 1d and 1g - thermal resistances of heat exchange by lower and upper surface, respectively, 2,3 and 5 - thermal resistances of substrate and layers, respectively, 4 – thermal contact resistance (between layers), 6 – resistance which connects heat source with layer node, 7 – thermal capacitance of substrate, 8 – thermal capacitance of layer, UTo – voltage source which represents ambient temperature, uT(t) – voltage source which represents heat source.

The majority of physical parameters (important from the thermal point of view) of particular thick-film components are strongly dependent on temperature. For proper model creation the very good knowledge about such temperature relations is required.

For simulation of temperature field distribution using equivalent RC models the PSPICE program was applied. The non-linear thermal resistances and capacitances (with taking into consideration their dependence on temperature) were modeled in PSPICE program as special "subcircuits" – so-called controlled resistances and capacitances.

III. Example of Simulations

As example of complex electro-thermal simulation the hybrid power controller HR093 was chosen (see the picture in Fig.1). It is manufactured by HYBRES factory (Rzeszow, Poland) for controlling of power in household equipment (especially in vacuum cleaners).

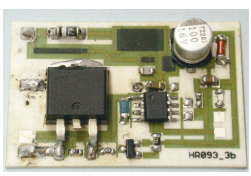


Fig. 2. Picture of controller.

The block and schematic diagrams (with topology) of the above-mentioned controller are presented in Fig. 3 and 4, respectively.

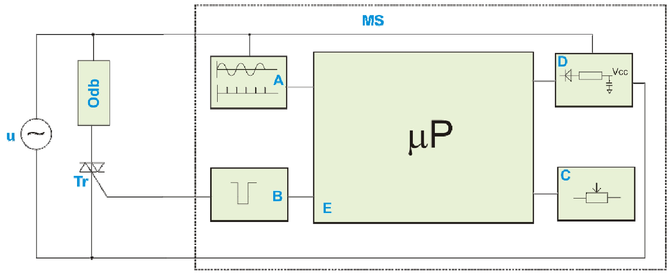


Fig. 3. Block diagram of power controller: A-zero detection module, B-pulse forming circuit, C-control of supplied power, D-supplier, E-microprocessor control unit, Tr-triac T1635, Odb-power receiver, u-voltage source.

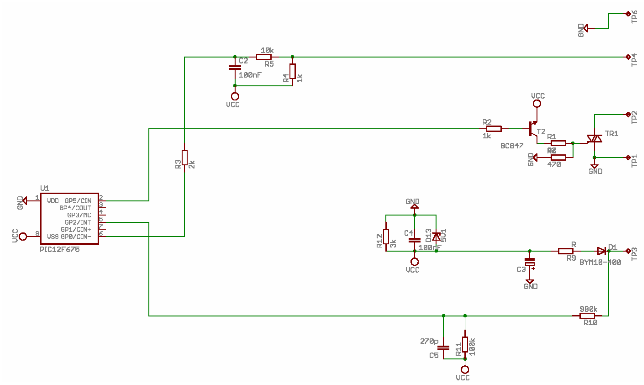


Fig. 4. Schematic diagram of power controller.

Such device was simulated in PSPICE program as typical electrical circuit, as RC model of thermal object (for temperature field determination) and again electrically simulated with taking into account the temperature changes in particular points of analyzed circuit.

A. *Electrical simulation*

Transient analysis was conducted for different power levels with taking into consideration the resistance and inductance of windings. The results of simulations are presented in Fig. 5. and 6.

Those results were basis for determination of average value of power for particular components of microcircuit (necessary to the thermal analysis – estimation of heat fluxes density).

B. *Thermal simulation*

The main heat sources in analyzed controller are supplying resistor R and triac Tr. The simulations of temperature distribution were made using EQ Term program. It allow to create RC elements network (on the basis of geometrical data and grid size) in the acceptable form for PSPICE program. After simulation it is possible to view the all temperature in each layer (in grid points) of microcircuit in analyzed time steps.

The heat fluxes density (as results of electrical simulation) were equal 6016 W/m^2 (for triac) and 5000 W/m^2 (for resistor). The results are shown in Fig. 7.

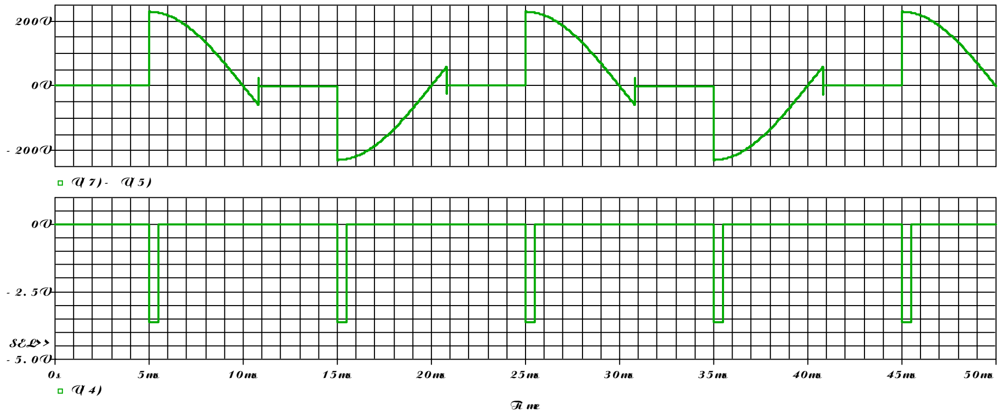


Fig. 5. Voltage courses on load (upper) and triac gate (lower) for average motor power.

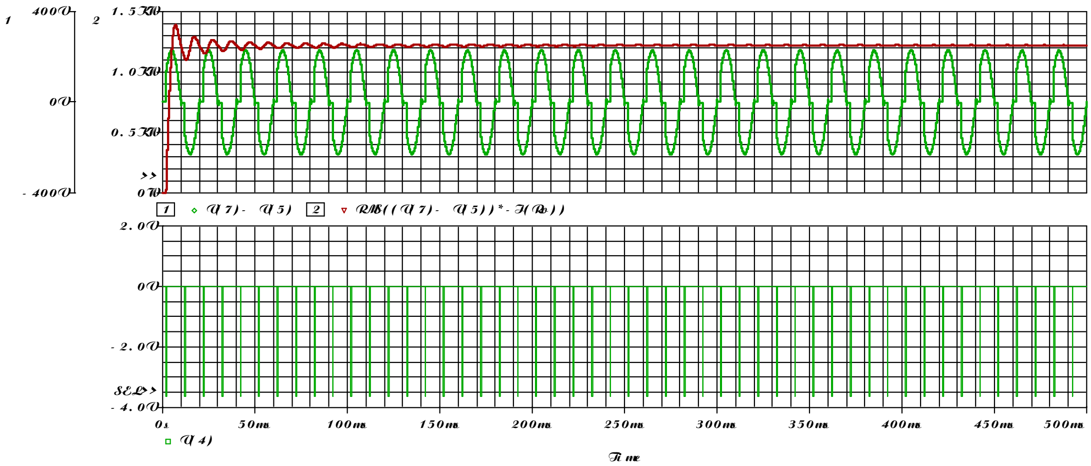


Fig. 6. Voltage courses and average power on load (upper) and triac gate (lower) for maximal motor power.

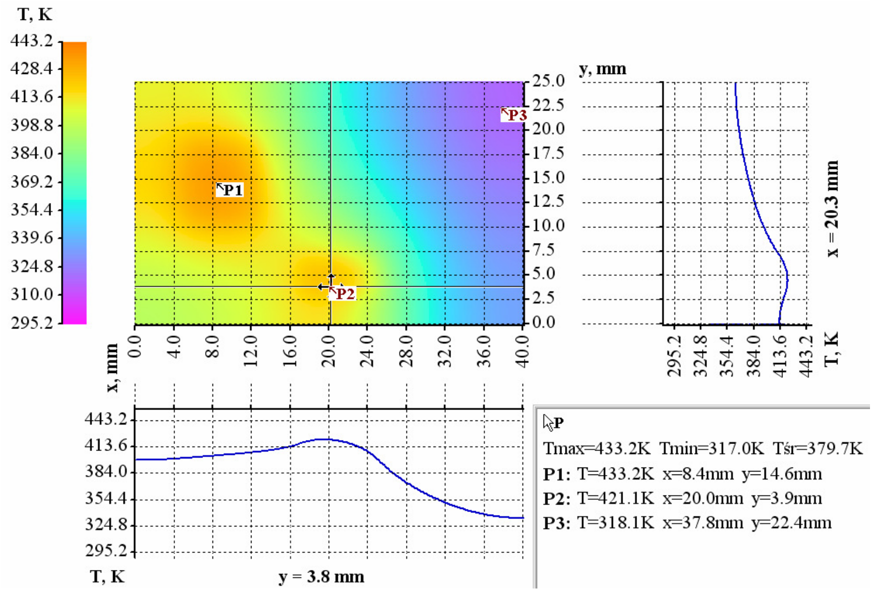


Fig. 7. Temperature field distribution in controller substrate for maximal power of motor.

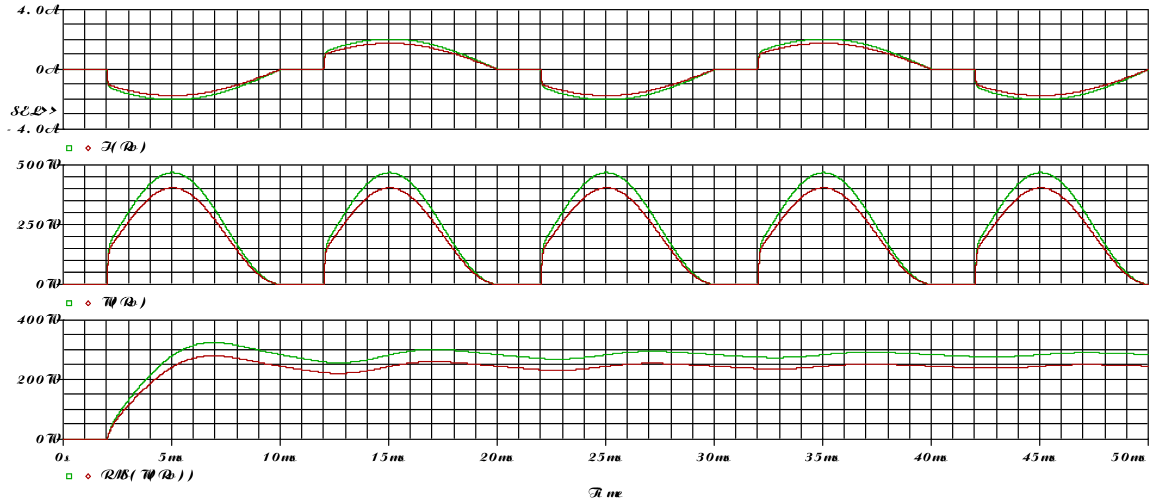


Fig. 8. . Results of transient analysis with (red color) and without (green color) taking into account the local values of temperature (courses on load): current (upper), pulse power (middle), average power (lower).

C. *Electrical simulation with temperature influences*

The newest versions of PSPICE program allow to introduce to the simulated circuit temperature gradient (by definition of different local temperature on each element important from the thermal point of view).

In the case of analyzed controller the local temperature was given for supplying resistor and triac only (values obtained from thermal simulations using equivalent RC network). The results of calculations (transient analysis) are presented in Fig.8.

IV. EXPERIMENTAL VERIFICATION

The measurements were made using thermovision camera V–20ER005-25 (Vigo Warszawa). They are carried out for open system with natural convection and cooling by one area. The series commutator motor was used as load. Thermogram for maximal motor power is presented in Fig. 9.

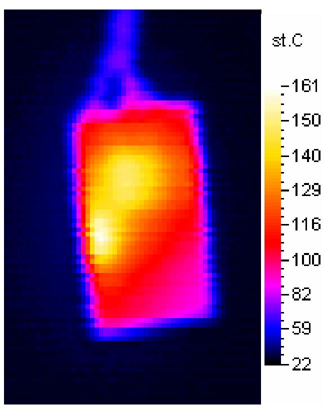


Fig. 9. Temperature distribution on controller substrate (maximal motor power).

V. CONCLUSIONS

The presented method of complex simulation using electro-thermal analogy (equivalent RC models based on Beuken theory) is very useful (and more easier) for designers of the electronic equipment. The big advantage of such method is possibility of its application not only for thick-film microcircuits. The main disadvantage is necessity of well knowing of the complete thermal characteristic of the analyzed object.

REFERENCES

[1] Beuken D.L. *Wärmeverluste bei periodisch betriebenen Öfen*, Dissertation, 1936 Freiburg, Germany.

[2] Andonova A., Ph. Philippov, N. Atanasova, Methodology of Estimate Reliability of Highly Reliable Components by Monitoring Performance Degradation, *Proceedings of 24th ISSE,* May 5-9, Bulgaria, 2004, pp. 125-128.

[3] Błąd G., D. Klepacki , J. Potencki , Simulation of Temperature Distribution in Thick-Film Sensors, *Proceedings of 8th International Symposium For Design And Technology In Electronic Modules SIITME'2002*, Cluj-Napoca, September 19-22, 2002, pp.34-37.

[4] Błąd G., D. Klepacki , Modelowanie pól temperatury w grubowarstwo-wych strukturach sensorowych; *ELEKTRONIZACJA* No 7-8, 2002, pp. 39-42 (in Polish).

[5] Błąd G., Z. Hotra ., D. Klepacki., J. Potencki , Modelling of Temperature Field Distribution in Thick-Film Sensor Structures, *Proceedings of XXVI Conference of IMAPS Poland Chapter, Warszawa,* 25-27 September 2002, pp. 109-112.

[6] Potencki J., A. Pietrikova, Possibilities of Temperature Fields Shaping in Multilayer Thick-Film Heating Systems, *SPIE*, vol. 3730, 1999, pp. 130-138.

[7] Błąd G., D. Klepacki, J. Potencki, Nonlinear Discrete Elements in Electro-Thermal Model of Thick-Film Structure, *28th International Spring Seminar on Electronic Technology ISSE 2005*, Wiener Neustadt, Austria, May 18-22, pp.231-236, 2005.

R-2R Digital-to-Analog Converter: Analysis and Practical Design Considerations

Dimitar P. Dimitrov[1]

Abstract – **The following article is an attempt to describe the rudiments of R-2R structures in reverse connection and their application in D/A converters.**
Analysis of R-2R ladders is made simple and intuitive by using Thevenin equivalent circuits. Formulae expressing all the basic properties of R-2R structures are derived using this approach. Practical design considerations are also discussed in great detail. To prove the theoretical analysis, test structures are designed and fabricated in 1.0 µm and 0.6µm double-poly, double-metal CMOS processes. Experimental results are then analyzed.

Keywords – **ADC, R-2R, mixed-signal, CMOS**

I. INTRODUCTION

R-2R structures of digital-to-analog converters (DAC) are very popular for their simplicity. For a DAC with a resolution of N bits 2N+1 resistors and 2N switches are necessary. In addition, only two resistor values are needed: R and 2R. Moreover, since 2R=R+R (and vice versa: R=2R//2R), only one resistor value is actually required. Thus the entire R-2R ladder is implemented as an array of equal resistors.

There are two basic types of R-2R DACs [1] [2] [3]:
- current mode (current steering).
- voltage mode (R-2R ladder in inverse connection).

The current mode DAC has been considered as the traditional approach. In this approach, the R-2R ladder is used to produce a set of binary-weighted currents whose sum is then converted to voltage (Fig.1)

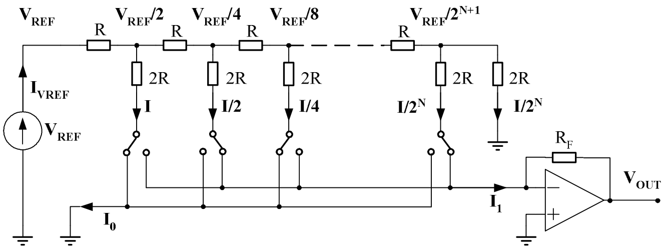


Fig. 1. Current mode R-2R DAC

The voltage at the output of the current-mode DAC is Eq. 1:

$$V_{OUT} = -R_F \times \sum_{k=0}^{k=2^{n-1}} I_k = -V_{REF} \times \frac{R_F}{2R} \times D \qquad (1)$$

Where D is the digital word applied to the converter:

$$D = (b_{N-1}x2^{N-1};\ b_{N-2}x2^{N-2}; \ldots b_1x2^{1};\ b_0x2^{0})$$

An important disadvantage of the current mode R-2R structures is the need for an operational amplifier that performs the current-to-voltage conversion.
The opamp itself introduces errors such as offset voltage, limited slew rate, limited output swing, etc.
The output of the voltage-mode R-2R DAC is voltage, so no opamp is needed provided the load impedance is high enough, which is usually the case in CMOS circuits.
In this article, the long-neglected voltage-mode R-2R DAC is discussed in some more detail. Expressions are derived for all the basic properties of the R-2R ladders by means of Thevenin equivalent circuits. The emphasis is on the practical design considerations and design methodology.
The rudiments of R-2R ladder in voltage mode are discussed in Section II.
Errors and error sources are dealt with in Section III Experimental results are given in Section IV.

II. THE VOLTAGE-MODE R-2R DAC

Fig. 2 shows a network of N cascaded R-2R links, numbered from 0 to N-1.

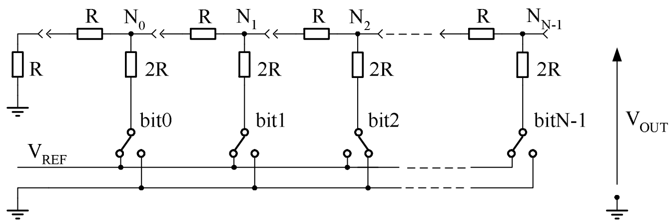


Fig. 2. The voltage mode R-2R DAC

Note that in contrast to the current-steering ladder, the basic R-2R links are connected in reverse order.
A terminating resistor of value R is connected to the leftmost link (numbered 0), so that the equivalent impedance seen to the left each link is exactly R and the equivalent impedance seen at the output node V_{N-1} is also R.
The Thevenin equivalent circuit for an arbitrary R-2R link is shown in Fig. 3

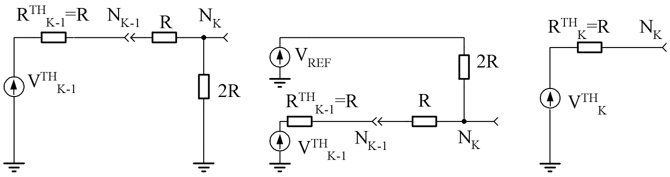


Fig. 3. Thevenin equivalent circuit of the voltage mode R-2R DAC

[1] Dimitar P. Dimitrov works at Melexis-Bulgaria Ltd.
84 Ami Boue Str., 1612 Sofia, Bulgaria
e-mail: ddi@melexis-bg.com

For an arbitrary link K, the following expression holds:

$$V_K^{th} = \frac{V_{N-1}^{th}}{2} + bit_K \times \frac{V_{REF}}{2} \rightarrow for...bit_K \subset (0,1) \qquad (2)$$

The index "th" stands for "Thevenin" equivalent circuit. Applying this recurrent formula to the entire chain of N basic links, numbered from 0 to N-1, yields the voltage at the last node (the output of the DAC) Eq. 3

$$V_{OUT} = V_{N-1} = b_{N-1} \times \frac{V_{REF}}{2^1} + b_{N-2} \times \frac{V_{REF}}{2^2} + ... + b_0 \times \frac{V_{REF}}{2^N} = \frac{V_{REF}}{2^N} \times D \qquad (3)$$

where D is the digital code applied:

$D = (b_{N-1}x2^{N-1}; b_{N-2}x2^{N-2};\dots b_1x2^1; b_0x2^0)$.

In summary, the entire R-2R ladder can be replaced with a Thevenin equivalent circuit with equivalent open-circuit voltage V_{OUT} and output impedance Z_{OUT}, Eq (4):

$$V_{OUT} = \frac{V_{REF}}{2^N} \times D \qquad Z_{OUT} = R \qquad (4)$$

Power consumption

A point worth mentioning is that in contrast to the current-modeR-2R ladder, the power consumption of the voltage-mode R-2R ladder is not constant but varies with the code applied.

For all bits = "0" power consumption is also zero (all 2R resistors are connected to GND). When only one bit is "1" the consumption is $V_{REF}/3R$. Maximum consumption occurs at codes 010101…01 and 101010…11 (Fig. 4)

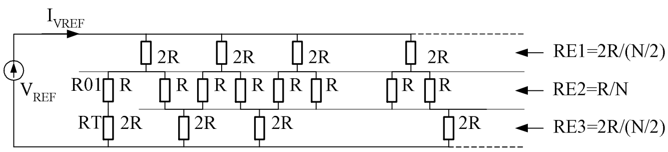


Fig. 4. The load applied to the reference source

If the number of R-2R links (i.e. number of bits) is large enough, the problem can be simplified by assuming that the contribution of the resistors R01 and RT is negligible in comparison to all the other resistors. Then all nodes across lines A-A and B-B would be at potentials V_A and V_B respectively. Thus the impedance seen by the reference is:

$$R_E = \frac{2R \times 2}{N} + \frac{R}{N} + \frac{2R \times 2}{N} = \frac{9R}{N} \qquad (5)$$

And the current flowing out of the reference is:

$$I_{VREF} = \frac{V_{REF}}{R_E} = V_{REF} \times \frac{N}{9R} \qquad (6)$$

This formula is accurate if the number of bits is large enough. For N more than or equal to 8 bits the error is less than 10%.

The power consumed from the digital supply is actually zero, since no current flows through the switch gates.

III. ERROR SOURCES IN VOLTAGE-MODE R-2R DAC

Effects of tolerances and device mismatch

Device mismatch is the major source of error in any DAC. The stochastic matching between two identically designed resistors is defined as the standard deviation of the normal distribution for the relative difference δ_R.

$$\delta_R = \frac{\Delta R}{R} \qquad (7)$$

The matching of two identically designed resistors with size W x L is described by the following model [4]:

$$\delta_R = \frac{A_R}{\sqrt{W \times L}} \qquad (8)$$

where A_R is a process-dependent matching parameter.
The influence of device mismatch of each R-2R link increases as its rank in the ladder increases. The Differential Nonlinearity (DNL) and the Integral Nonlinearity (INL) are likely to reach their maximum around the major carry points that involve the most significant bit (i.e transitions like 0111…111 -> 1000…000). The worst case occurs when all resistors are at their minimum (maximum) values and only the 2R resistor of the most significant bit (b_{N-1}) is at its maximum (minimum) value.

$$R = R - \Delta R = R(1-\delta_R) \qquad 2R_{BITN-1} = 2(R+\Delta R) = 2R(1+\delta_R) \qquad (9)$$

The equivalent Thevenin circuit is shown in Fig. 5

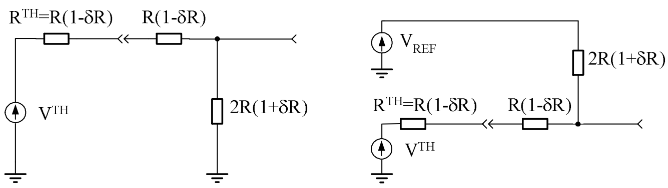


Fig. 5. The effect of device mismatch

Using this circuit the absolute values of DNL and INL can be expressed in terms of least-significant bits:

$$DNL = 2^N \times \delta_R \qquad (10)$$

$$INL = 2^{N-1} \times \delta_R \qquad (11)$$

Switch imperfection

a) On-resistance of closed switches. The on-resistance of a closed switch adds to the resistance of the corresponding 2R resistor. That is, its actual value becomes $Ra=2R+R_{ON}$. Since the value of R resistors is not affected by switches, the actual division ratio of the basic R-2R link changes, resulting in integral and differential nonlinearity.
The equivalent Thevenin circuit with R_{ON} included is shown in Fig. 6.

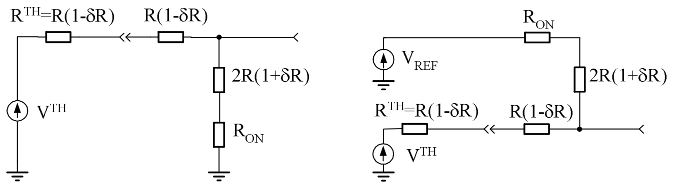


Fig. 6 – The influence of the switches

With R_{ON} included, the expressions for DNL and INL are (in terms of LSB):

$$DNL = 2^N \times \left(-\delta_R - \frac{R_{ON}}{4R} \right) \tag{11}$$

$$INL = 2^N \times \left(\frac{4R \times \delta_R - R_{ON}}{8R + 2R_{ON}} \right) \tag{12}$$

Note that since δ_R can be either positive or negative, the first term in the brackets in Eq. (11) and (12) can also be either positive or negative. However, R_{ON} is always positive.

b) Switch leakage. Since every 2R resistor is connected either to V_{REF} or to GND, which are both low impedance nodes, the leakage currents flow into V_{REF} or into GND rather than into the resistor array. That is why switch leakage currents do not affect DAC performance provided the reference source is low ohmic. Tests of real devices show that the effect of switch leakage is negligible.

Poor layout

Poor layout can significantly spoil device matching, thus leading to INL and DNL. Straight-forward layout (devices in a row) is the simplest and most compact way to place a set of resistors. However, it is prone to process deviations along the wafer. A layout with one axis of symmetry can significantly compensate for process deviations along this axis. A layout which is symmetrical in respect to two axes of symmetry (common-centroid layout) can effectively compensate for the process deviations along both the horizontal and the vertical axes at the price of an increased die area and an elaborate interconnection scheme.

Influence of the reference source

As mentioned above, the current flowing out of the reference source varies with the digital code applied. The changes in the current consumed cause changes in the voltage drop across the output impedance R_I of the reference. These code-dependent voltage drops result in integral nonlinearity. The error caused by this parasitic voltage drop is:

$$\delta_{VREF} = \frac{\Delta V_{REF}}{V_{REF}} = \frac{R_I (N-3)}{9R} \tag{13}$$

A fact which is often neglected at the layout phase is that the wires connecting the converter to the reference source and to ground also have parasitic resistance, which adds to the output impedance of the reference source.

Influence of the load impedance

As shown in Fig. 3 and Eq. 3, the voltage mode R-2R DAC can be represented by its equivalent Thevenin circuit, which has an open-circuit ideal voltage source and output impedance Z_O. When a finite load impedance Z_L is applied, the actual output voltage is:

$$V_{OUT} = V_{REF} \times D \times \frac{Z_L}{Z_L + Z_O} \tag{14}$$

All codes are attenuated by the same factor. That is, the finite load impedance causes an additional gain error expressed in Eq. 12

IV. EXPERIMENTAL RESULTS

The voltage-mode DAC shown in Fig. 2 was fabricated in 1.0 µm and 0.6µm double-poly, double-metal CMOS processes. Several R-2R structures were prepared for a resolution of 12, 10 and 8 bits. To check the influence of the layout, three different layout schemes were employed for the 8-bit DAC. Every effort was made at the layout phase to ensure good device matching and to keep the lengths of the interconnecting wires equal. Tests were performed by means of a special test kit and a KEITHLY 2000 digital voltmeter. The test results are given in Fig. 7 – Fig. 10. The minimum measured values are in blue, the maximum measured values are in red. The increase of both INL and DNL at the major carry point is easily seen. Device mismatch is the major source of error in voltage-mode R-2R DAC. The errors of the converters are actually determined by the mismatch of the two most significant bits. As predicted by Eq. 10 and Eq. 11, for the same device mismatch, the DNL is (in terms of LSB) twice as big as INL. If DNL is higher than 1 LSB the conversion monotonicity is not guaranteed. The improvement of conversion linearity (both INL and DNL) for the symmetrical layouts (one axis of symmetry and common-centroid) as compared to the straightforward layout is partially due to the fact that in symmetrical layout schemes devices are split into several unit devices whose total area is larger than the area of the straight-forward layout. Test results are summarized in Table I:

TABLE I
MEASURED R-2R DAC NONLINEAREITY

DAC type: process and layout scheme	INL [LSB]	DNL [LSB]
12-bit; 1µm CMOS common-centroid layout	1,5	2,6
10-bit; 1µm CMOS one axis of symmetry	0,6	1,1
8-bit; 1 µm CMOS one axis of symmetry	0,18	0,32
8-bit; 0,6µm CMOS straightforward layout	0,11	0,22
8-bit; 0,6µm CMOS one axis of symmetry	0,09	0,17
8-bit; 0,6µm CMOS common-centroid layout	0,05	0,09

V. Conclusions

The voltage-mode R-2R DAC (R-2R in reverse connection) has been discussed in this article as an effective way to circumvent the problems associated with the opamp in current-mode R-2R DACs. Theoretical analysis using Thevenin equivalent circuits has been proposed that gives insight into the voltage-mode R-2R DAC operation. The derived expressions have been proved by means of design experiment. Experimental results agree with the formulae derived in Section II. Ways to reduce the INL and DNL by means of proper layout techniques have also been discussed in brief.

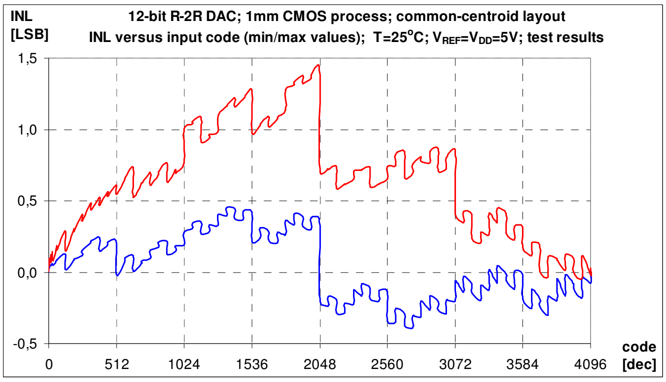


Fig. 7. INL of 12-bit R-2R DAC

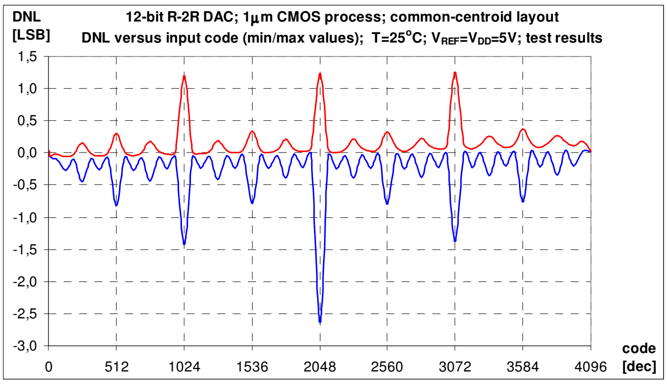


Fig. 8. DNL of 12-bit R-2R DAC

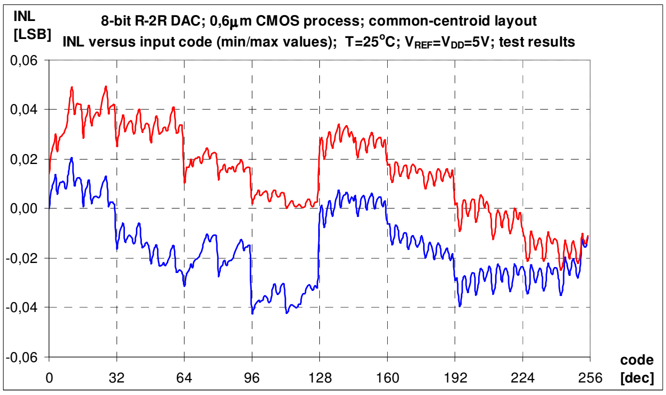


Fig. 9. INL of 8-bit R-2R DAC; Common-centroid layout

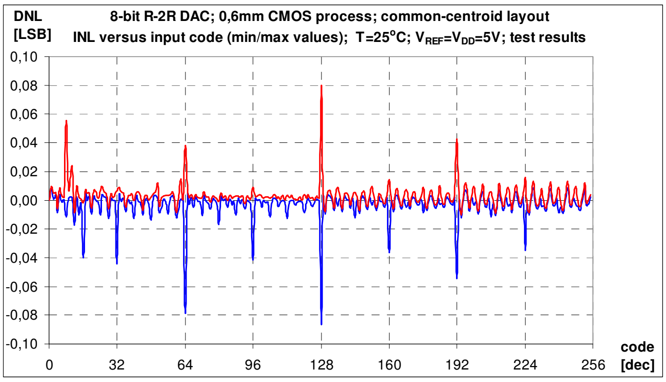


Fig. 10. INL of 8-bit R-2R DAC; Common-centroid layout

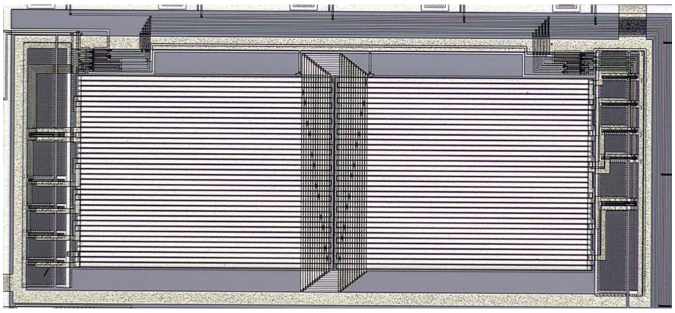


Fig. 11. Die photograph of 12-Bit DAC, 1μm CMOS process

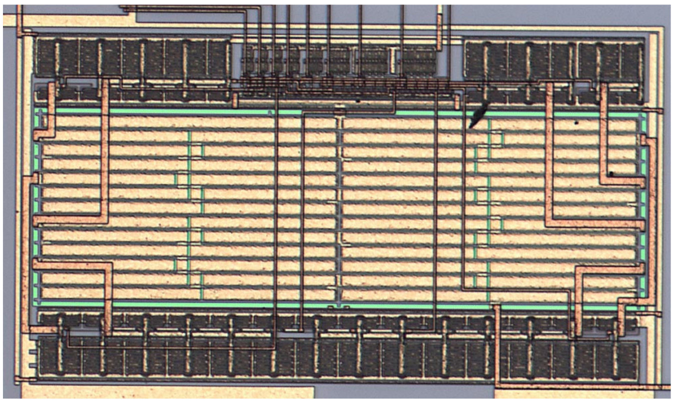


Fig. 12. Die photograph of 8-Bit DAC, 0.6μm CMOS process; Straight-forward layout

References

[1] Mikael Gustavsson, J. Jacob Wikner, Nianxiong Nick Tan, "CMOS Data Converters for Communications", Kluwer Academic Publishers, 2002, pp. 64-79.

[2] R. Jacob Baker "CMOS, Mixed-Signal Circuit Design", Wiley-IEEE Press, July 2002

[3] David A. Johnes, Ken Martin, "Analog Integrated Circuit Design", John Wiley & Sons, 1997, pp. 507-11.

[4] Marcel J. M. Pelgrom, Aad C. J. Duinmaijer, Anton P. G. Welber,s "Matching propertries of MOS transistors", IEEE JSSC, vol. SC-24, October 1989, pp. 1433-39

[5] John A. Schoeff, "An inherently Monotonic 12 Bit DAC", IEEE JSSC vol. SC-14, December 1979

Noise Characteristics for Amplifier Model with a Thevenin Source

Pesha D. Petrova[1]

Abstract - **Accurate noise analysis is a prerequisite for low-noise circuit design. In this paper, noise analysis for V_n–I_n amplifier model with a Thevenin source is presented. Noise characteristics, such as signal – to – noise ratio, noise factor, noise figure and noise temperature, as well as, the optimum source impedance, which minimizes the noise factor, are obtained. MATLAB simulation results are presented.**

Keywords – **amplifier noise model, noise factor, noise figure, noise temperature, signal – to – noise ratio.**

I. INTRODUCTION

Noise is one of the most important factors affecting the operations of communications circuits. That's why the main objective of noise analysis is to design low-noise circuits.

When designing an amplifier for a specific application, there are many characteristics to be met and decisions to be made. They include gain, bandwidth, impedance levels, feedback, stability, dc power, cost and signal - to - noise ratio (SNR) requirements. Usually requirements low - noise design contradict to the other important circuit parameters such as bandwidth, gain, input/output characteristics, etc. The low-nose design is different for each specific case, but amplifier designers can elect one of two paths [1]:

- The wrong approach is to worry about the gain and bandwidth first and later in the design power they check for the noise.
- Alternatively, one can design the system with initial emphasis on noise performance

Although there are many low-noise devices available they do not perform equally for all signal sources.

To obtain the optimum noise performance, it is necessary to select the proper amplifying device (BJT, FET, or IC) and operating point for the specific input source.

In order to predict the noise behavior of an amplifier correctly, accurate noise models are required. Without them, the design and optimization of an amplifier's noise characteristics cannot be successful.

If it is necessary to study a circuit's noise in terms of the physical phenomena, the Van der Ziel models [2] which are well accepted models and have a solid physical basic, can be used. But for amplifier design purposes, both amplifier noise models, with a Thevenin or with a Norton input source, are more appropriate.

[1]Pesha D. Petrova is with the Department of Communication Techniques and Technologies, Technical University of Gabrovo, 5300 Gabrovo, 4, Hadji Dimitar St., Bulgaria, E-mail: daneva@tugab.bg

II. NOISE CHARACTERISTICS FOR AMPLIFIER MODEL WITH A THEVENIN SOURCE

The amplifier model with a Thevenin noise input source, described in [3], is shown in Fig.1, where the control voltage Vi is the voltage across Zi.

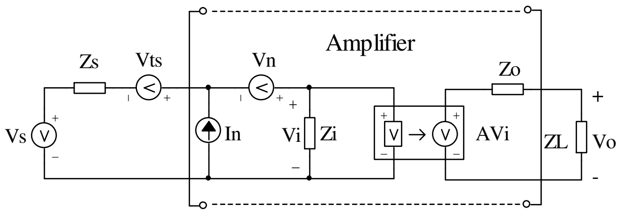


Fig. 1. Vn – In amplifier model with Thevenin source.

A. Signal – to – Noise Ratio

If the amplifier is noiseless, the signal – to - noise ratio is given by $SNR = v_s^2 / v_{ts}^2$, where v_s^2 is the mean-square source voltage and v_{ts}^2 is the mean-square thermal noise voltage generated by the source impedance. When the amplifier noise is included, the signal – to - noise ratio is $SNR = v_s^2 / v_{ni}^2$. The total mean-square noise voltage v_{ni}^2 at the input of the amplifier can be written as

$$v_{ni}^2 = 4kTR_S\Delta f + v_n^2 + 2v_n i_n \operatorname{Re}(cZ_S^*) + i_n^2 |Z_S|^2 \quad (1)$$

where $Z_S = R_S + jX_S$ is the source impedance and $c = c_r + jc_i$ is the correlation coefficient between noise sources V_n and I_n.

Thus, the signal – to – noise ratio can be determined as

$$SNR = \frac{v_s^2}{v_{ni}^2} =$$

$$= \frac{v_s^2}{4kTR_S\Delta f + v_n^2 + 2v_n i_n \operatorname{Re}(cZ_S^*) + i_n^2 |Z_S|^2}. \quad (2)$$

The signal – to – noise ratio is maximized by minimizing v_{ni}^2. The source impedance which minimizes v_{ni}^2 can be obtained by setting $\partial v_{ni}^2 / \partial R_S = 0$ and $\partial v_{ni}^2 / \partial X_S = 0$ and solving for R_S and X_S. The solution for R_S is

$$Rs = -\frac{2kT\Delta f + v_n i_n}{i_n^2} \quad . \tag{3}$$

It is negative. Because this is no realizable, $R_S = 0$ is the realizable solution for the least noise. The solution for X_S is

$$X_S = -c_i \frac{v_n}{i_n} \quad . \tag{4}$$

Therefore, the source impedance which minimizes the total noise input voltage v_{ni}^2, is given as

$$Z_S = R_S + jX_S = 0 - jc_i \frac{v_n}{i_n} \quad . \tag{5}$$

Because minimum noise occurs for $R_S = 0$, it can be concluded that a resistor should never be connected in series with a source at the amplifier input if noise performance is a design criterion. If a series resistor is required, e.g. for stability, it should be much smaller than R_S. Although the output impedance of a source is usually fixed, the signal – to – noise ratio can be improved by adding a reactance in series with the source which makes the total series reactance equal to the imaginary part of Z_S in Eq. (5). When this is the case, after substituting Eq. (5) into Eq. (1), the equivalent noise input voltage can be determined by

$$v_{ni}^2 = 4kTR_S\Delta f + v_n^2\left(1-c_i^2\right) + 2c_r R_S v_n i_n + i_n^2 R_S^2 \quad . \tag{6}$$

B. Noise Factor and Noise Figure

Dividing the noiseless amplifier signal – to – noise ratio by noise amplifier signal – to – noise ratio gives the noise factor as

$$F = \frac{v_s^2 / v_{ts}^2}{v_s^2 / v_{ni}^2} = \frac{v_{ni}^2}{v_{ts}^2} =$$

$$= 1 + \frac{v_n^2 + 2v_n i_n \operatorname{Re}(cZ_S^*) + i_n^2 \left|Z_S\right|^2}{4kT_0 R_S \Delta f} \tag{7}$$

where T_0 is the standard temperature.

It follows from this expression that a noiseless amplifier has the noise factor $F = 1$. A common way of presenting the noise factor is the so-called noise figure N.

$$N = 10\log(F) \, . \tag{8}$$

Eqs. (7) and (8) give a first hint of why a low noise figure is considered important, and why a high gain is often associated with a low noise figure. Noise figure is, however, often overstated in its importance in a low noise system design. It is the correct balance between noise figure and gain that is important rather than a low noise figure itself. It may not be obvious, but an amplifier with a noise figure (in dB) numerically greater than its gain (in dB) is not useful. Indeed, it can be proven that there exists a combination of passive components which would perform better than an amplifier with lower gain than noise factor.

Often, it is convenient to express F in terms of the amplifier noise resistance R_n, noise conductance G_n and the correlation impedance Z_c. These are related to v_n^2, i_n^2, and c by [4], [5]:

$$R_n = \frac{v_n^2}{4kT_0\Delta f} \, , \tag{9}$$

$$G_n = \frac{i_n^2}{4kT_0\Delta f} \quad , \tag{10}$$

$$Z_c = R_c + jX_c = c\frac{v_n}{i_n} = \left(c_r + jc_i\right)\frac{v_n}{i_n} \, . \tag{11}$$

R_n, and G_n, respectively, represent normalized values of v_n^2 and i_n^2, where the normalization factor is $4kT_0\Delta f$.

If the amplifier noise parameters are expressed in terms of R_n, G_n and c, the noise factor is given by

$$F = \frac{v_{ni}^2}{v_{ts}^2} = 1 + \frac{v_n^2 + 2v_n i_n \operatorname{Re}\left(cZ_S^*\right) + i_n^2\left|Z_S\right|^2}{4kT_0 R_S \Delta f} =$$

$$= 1 + \frac{R_n + G_n\left[2\left(R_S R_c + X_S X_c\right) + \left(R_S^2 + X_S^2\right)\right]}{R_S} \, . \tag{12}$$

The optimum source impedance which minimizes the noise factor F is obtained by setting $\partial F / \partial R_S = 0$ and $\partial F / \partial X_S = 0$ and solving for R_S and X_S. It follows from $\partial F / \partial R_S = 0$ that

$$G_n R_S^2 - 2G_n X_S X_c - G_n X_S^2 - R_n = 0 \, . \tag{13}$$

Solving Eq. (13) for R_S gives

$$R_{Sopt} = \sqrt{1 - c_i^2}\,\frac{v_n}{i_n} \, . \tag{14}$$

From $\partial F / \partial X_S = 0$ it follows that

$$2G_n X_c + 2G_n X_S = 0 \text{, i.e.,}$$

$$X_{Sopt} = -X_c \, . \tag{15}$$

Therefore, the optimum source impedance is obtained as

$$Z_{opt} = R_{opt} + jX_{opt} = \left[\sqrt{1 - c_i^2} - jc_i\right]\frac{v_n}{i_n} =$$

$$= \sqrt{\frac{R_n}{G_n} - X_c^2} - jX_c \, . \tag{16}$$

It follows from Eq. (16) that the imaginary part of Z_{opt} is equal to the imaginary part of Z_S in Eq. (5), which minimizes the signal – to – noise ratio. The corresponding minimum value of the noise factor can be obtained by replacing R_S and X_S in Eq. (12) with R_{Sopt} and X_{Sopt}, respectively. In this case the optimum noise factor is given by

$$F_{\min} = 1 + \frac{v_n i_n}{2kT_0 \Delta f}\left(c_r + \sqrt{1 - c_i^2}\right) =$$

$$= 1 + 2G_n\left(R_c + R_{opt}\right). \tag{17}$$

The noise factor can also be expressed in terms of $F_{\min}$ and Z_{opt}. It follows from Eqs. (12) and (17) that the difference $F - F_{\min}$ is given by

$$F - F_{\min} = \frac{R_n + 2G_n\left(R_c R_S + X_c X_S\right) + G_n\left(R_S^2 + X_S^2\right)}{R_S}$$

$$- 2G_n\left(R_c + R_{opt}\right) =$$

$$= \frac{R_n - 2G_n\left(R_{opt} R_S + X_{opt} X_S\right) + G_n \left|Z_S\right|^2}{R_S} \tag{18}$$

where $X_c = -X_{opt}$ has been used. Adding and subtracting the term $G_n\left(R_{opt}^2 + X_{opt}^2\right) = G_n\left|Z_{opt}\right|^2$ can complete the square in the nominator of this expression. This leads to the equation

$$F_{\min} = \frac{R_n + G_n\left[\left(R_S - R_{opt}\right)^2 + \left(X_S - X_{opt}\right)^2\right]}{R_S} -$$

$$- \frac{G_n\left|Z_{opt}\right|^2}{R_S}. \tag{19}$$

By substituting for $\left|Z_{opt}\right|^2$ as

$$\left|Z_{opt}\right|^2 = \frac{R_n}{G_n} \tag{20}$$

in Eq. (19), the noise factor is obtained as

$$F = F_{\min} + \frac{G_n}{R_S}\left[\left(R_S - R_{opt}\right)^2 + \left(X_S - X_{opt}\right)^2\right] =$$

$$= F_{\min} + \frac{G_n}{R_S}\left|Z_S - Z_{opt}\right|^2. \tag{21}$$

It is clear from Eq. (21) that the amplifier's noise behaviour is completely characterised by the optimum source impedance Z_{opt} which results to a minimum noise factor $F_{\min}$, and the noise conductance G_n describing how rapidly the noise factor increases for a deviation from the optimum source admittance.

C. The Noise Factor Fallacy

The noise factor can be misleading characteristic. If an attempt is made to minimize F by adding resistor in series with the source at the input of an amplifier, the signal – to – noise ratio is always decreased. This is referred to as the noise factor fallacy or the noise figure fallacy. Potential confusion can be avoided if low - noise amplifiers are designed to maximize the signal – to – noise ratio. This is accomplished by minimizing the equivalent noise input voltage.

D. Noise Temperature

The internal noise generated by an amplifier can be expressed as an equivalent input – termination noise temperature. When the source is represented by a Thevenin equivalent circuit, as is shown in Fig. 1, the noise temperature T_n is the temperature of the source resistance that generates a thermal noise voltage equal to the internal noise generated in the amplifier when referred to its input. For the Thevenin source, the noise temperature is defined by

$$4kT_n R_S \Delta f = v_n^2 + 2v_n i_n \operatorname{Re}\left(cZ_S^*\right) + i_n^2\left|Z_S\right|^2 \tag{22}$$

where $R_S = \operatorname{Re}\left(Z_S\right)$. It follows that the noise temperature is

$$T_n = \frac{v_n^2 + 2v_n i_n \operatorname{Re}\left(cZ_S^2\right) + i_n^2\left|Z_S\right|^2}{4kR_S \Delta f}. \tag{23}$$

The noise temperature is related to the noise factor by

$$T_n = \left(F - 1\right)T_0. \tag{24}$$

III. EXAMPLE

A few simulations for the amplifier noise model, shown in Fig. 1, have been performed in MATLAB. Some of noise characteristics, obtained with $v_n = 2nV$, $i_n = 10pA$, and $c = 0{,}1$, are presented in Figs. 2 and 3.

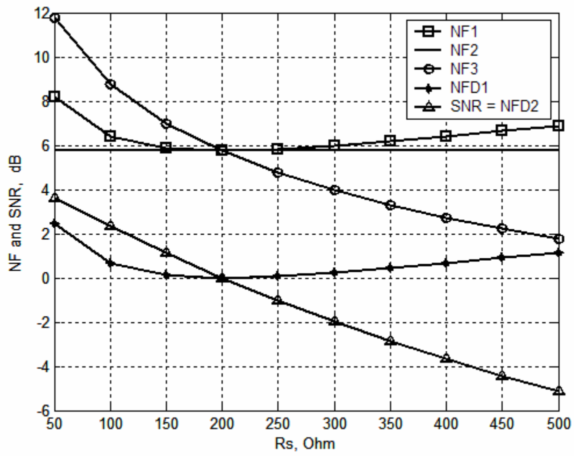


Fig. 2. Variations in SNR and NF over the source resistance range from $R_S = 50\Omega$ to $R_S = 500\Omega$

In Fig. 2 the noise figure $NF1$ for a given source resistance R_S is $NF1 = 10\log(v_{ni}^2(R_S)/v_{ts}^2(R_S))$. If a resistance is added in series with the source impedance to minimize the noise factor, the new noise figure can be expressed as $NF2 = 10\log(v_{ni}^2(R_{Sopt})/v_{ts}^2(R_{Sopt}))$ with R_{Sopt} given by Eq. (14) with $c_i = 0$. The decrease in the noise figure can be found as $NFD1 = NF1 - NF2$, and the dB decrease in SNR as $SNR = 10\log(v_{ni}^2(R_{Sopt})/v_{ni}^2(R_S))$.

The simulation results in Fig. 2 illustrate how the noise figure appears to be decreased by adding resistance in series with an amplifier input. However, the SNR is lowered. The fallacy comes from treating the added resistance as part of the source rather than part of the amplifier. In reality, the added resistor increases the amplifier noise, but the source noise remains constant. Because of that, the correct way to determine the noise figure $NF3$ with the added resistor is $NF3 = 10\log(v_{ni}^2(R_{Sopt})/v_{ts}(R_S))$. In this case the noise figure decreases by $NFD2 = NF3 - NF1$. Thus, this is the same decrease as the dB decrease in the SNR.

In Fig. 3 the noise temperature waveforms, that follow the noise factors $F1$, $F2$ and $F3$ changes, are presented.

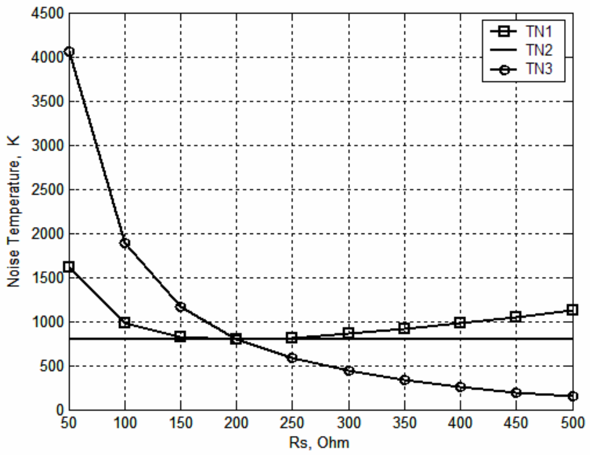


Fig. 3. Variations in noise temperature over the source resistance range from $R_S = 50\Omega$ to $R_S = 500\Omega$

The results in Figs. 2 and 3 show that the minimum noise factor (noise figure) do not correspond to the maximum signal – to – noise ratio. It can also be concluded that if a series resistor must be included at the amplifier input, its value should be much smaller than the source impedance.

IV. Conclusion

A simple but effective approach to determine the noise characteristics for amplifier model with a Thevenin source is developed. Expressions for signal – to– noise ratio, for noise factor (noise figure) and noise temperature are derived. The possibilities to improve the noise characteristics are studied. The effect of series resistor, added at the amplifier input, on the noise characteristics is shown. The noise factor (noise figure) fallacy due to the added resistor is commented. The correct way to determine the noise characteristics for this case is proposed. The approach for analysis can be used for low - noise amplifiers design.

References

[1] David E. Johnson, Johnny R. Johnson, and John L. Hilburn, *Electric Circuit Analysis.* Prentice-Hall, Inc., 1989.

[2] A. Van der Ziel, *Noise in Solid State Devices and Circuits,* John Wiley & Sons, Inc., New York, 1986.

[3] P. Petrova, “Amplifier Noise Model with Thevenin Input Sourse”, *XL International Scientific Conference on Information, Communication and Energy Systems and Technologies, Conference Proceedings*, vol. 1, pp.105-108, Niš , Serbia and Montenegro, 2005.

[4] J. Xu, Y. Dai, and Y. Li, “The study of the relation between R_n-G_n noise model and E_n-I_n noise model of an amplifier,” *IEEE Trans. Circuits Syst.I*, vol. 45, pp. 154-156, Feb. 1998.

[5] Jieh-Tsorng Wu, *Noise Analysis and Modeling*, National Chiao-Tung University, 2001.

Model building and testing procedure – An analogue multiplexer SPICE macromodel improved with temperature effects

Ivailo M. Pandiev[1]

Abstract – **A systematic procedure to design a behaviour SPICE macromodels is described in this paper. The proposed modelling procedure can be split into three basic steps: 1) structuring the model; 2) build the model; 3) validate the model. Using this method, an improved SPICE macromodel of a CMOS analogue multiplexer is presented in which the temperature effect of a switch on-resistance is modelled. Also, for creating the simulation model, techniques known from modelling operational amplifiers have been adapted. The proposed macromodel allows the simulation of circuits with respect to the behaviour in both the time and frequency domains, including error parameters and the temperature dependence of switch on-resistance. Model parameters for the integrated circuit ADG408 from Analog Devices are extracted as an example. The simulation results are compared with manufacturer's data and good correspondence is reported.**

Keywords – **Analogue circuits, Temperature effects, CMOS analogue multiplexers, SPICE, Modelling.**

I. INTRODUCTION

The CMOS analogue multiplexers have become an essential component in the design of electronic systems which require the ability to control and select a specified transition path for the analogue signal. These analogue devices are widely used for implementing multi-channel amplifiers, sample-and-hold circuits and PGAs. [1-3].

The switch on-resistance (R_{ON}) is an important consideration in applying analogue multiplexers. When one of the switches is closed, DC-performance is affected mainly by on-resistance and leakage current. Also R_{ON} changes as a function of the temperature and from switch-to-switch. However, temperature variation of the switch on-resistance is not presented in the available behaviour SPICE macromodels of analogue multiplexers [4, 5]. The majority of published SPICE macromodels only attempt to model R_{ON} , usually by a fixed on-resistance model parameter, defined into the ideal model with voltage-controlled switches. They therefore cannot be used to simulate the dynamic variations in R_{ON} due to changing temperature. To solve this problem, here is proposed a design procedure for modelling temperature dependence of the on-resistance in the CMOS analogue multiplexer macromodels. Also, techniques described in [6-8] have been adapted for macromodelling the temperature effects.

[1]Ivailo M. Pandiev is with the Faculty of Electronics from Technical University of Sofia, Kliment Ohridski 8, 1000 Sofia, Bulgaria. E-mail: ipandiev@tu-sofia.bg

II. MACROMODEL BUILDING PROCEDURE

The macromodel building and testing procedure presented in this section is based on a Top-Down analysis approach and by applying simplification and build-up technique, known from modelling operational amplifiers [9-12].

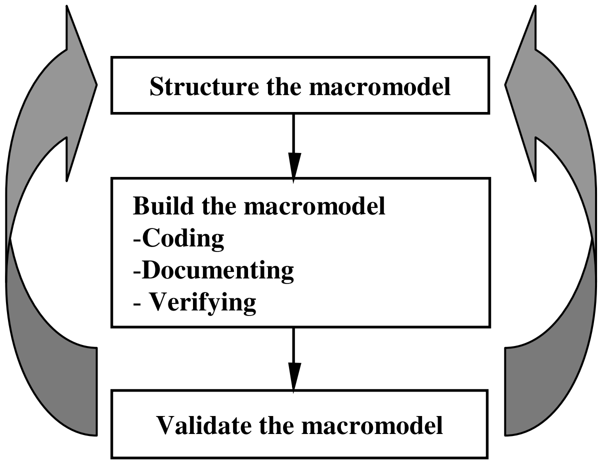


Fig. 1. Model building and testing procedure.

The process of model building and testing, as shown in Fig. 1, can be broken down into three steps: 1) structure the model; 2) build the model; 3) validate the model. Despite the fact that the process of model building is shown in a linear fashion, moving between structuring, building and validation of the model is a process that has iterative nature.

Each step of the procedure can be split into smaller steps which follow a similar iterative pattern. An outline of these steps can be resumed as follows.

1) The first step is to choose an internal structure for the model. The structure can be different from the actual structure of the IC. Also, the structure should consist of the elements that need to be defined and the data and logic, required to drive the model. It effectively is a paper version of the computer model.

2) The second step is the building of the model with the simulation software. This process consists of three distinct activities: coding, entering the model into the computer; documenting, explaining the model structure using software facilities and other techniques; verifying, ensuring that the code is correct. Each of these activities is performed in small steps, iteratively building and improving the model.

3) The last step is the validation of the macromodel. Validation is the process of determining whether a simulation model

is an accurate representation of the real system, for the particular objectives of the study. If a model is “valid”, then it can be used to make decisions about the system similar to those that would be made if it were feasible and cost-effective to experiment with the system itself.

Validation is achieved by comparing the model predictions with actual experimental results or with data sheet parameters. For the goals of validation it is necessary to collect data that represents the average behavior of the real IC. It is important to check not only the average levels of those data, but also to compare their spread, applying the methods of statistical analysis and calculus probability.

III. Building the Macromodels Sensing Temperature Effects of the Switch On-Resistance

The design procedure described in the previous section can be used to model second-order effects of the analogue multiplexers.

The analysis of switch on-resistance R_{ON} versus temperature at specified input signal level for different analogue multiplexers has shown that in most cases this electrical characteristic could be approximated by quadratic or exponential transfer function. Normally, semiconductor data books provided by IC-manufacturers contain electrical diagrams regarding the temperature dependence of the switch on-resistance that could be used in modelling the real devices.

The schematic diagram of the proposed macromodel is shown in Fig. 2. It’s based on a previous analogue multiplexer macromodel [4] from standard SPICE library, with the additional capability to model the temperature effects of the switch on-resistance. The input switch section, decoder section and supply current section are implemented as behavioural models using ABM (Analogue Behaviour Modelling) technique and are represented in Fig. 2 as black-boxes marked ‘demuxsw’, ‘scl’ and ‘pscc’ respectively. The temperature dependence of the R_{ON} has been modelled by adding new elements into existing behaviour SPICE macromodels for analogue multiplexers as follows.

As it is shown in Fig. 2, a linear two-port voltage-controlled current source (VCCS) G_T is connected to the output section of the existing SPICE macromodel, creating a new node 88. The current through G_T will have to follow the change in the temperature. The temperature-dependent controlled voltage of the VCCS G_T comes from a separate stage in Fig. 2. This additionally defined temperature stage consists of an ideal current source I_T and a SPICE temperature-dependent resistor R_T, which is controlled with equation [13]:

$$R_T(T) = R_T(T_0)\left[1 + TC1(T - T_0) + TC2(T - T_0)^2\right] \qquad (1)$$

where $R_T(T_0)$ is the value of the resistor at $T_0 = 27°C$ (SPICE-Option TNOM), T is the temperature in $°C$, $TC1$ is the linear temperature coefficient and $TC2$ is the quadratic temperature coefficient. The equation (1) will fit a quadratic curve through three points in a temperature graph by solving three equations with three arguments.

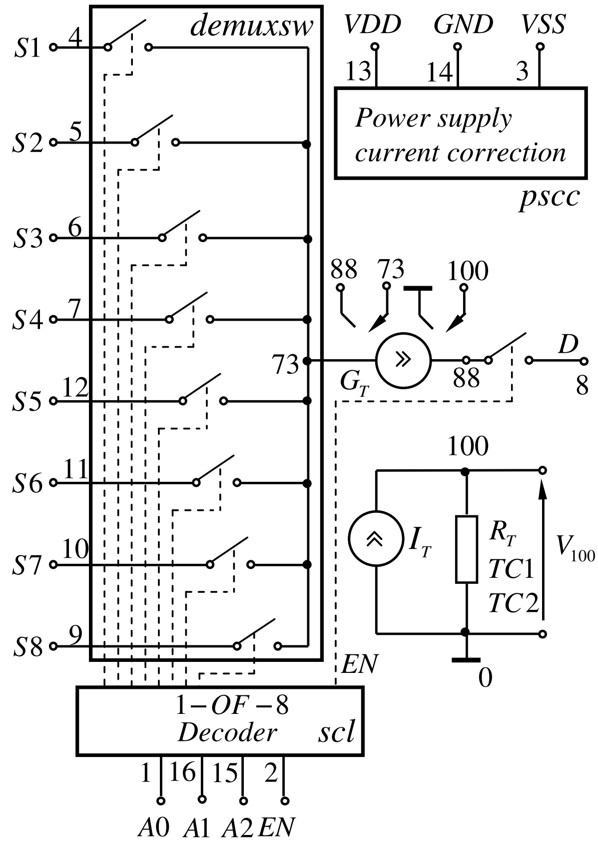


Fig. 2. Conceptual schematic of the proposed analogue multiplexer macromodel improved with temperature

The current I_T will flow through the resistor R_T, towards the internal ground. In such a way the voltage $V_{100,0}$ will depend upon the temperature.

$$V_{100,0} = I_T R_T(T_0)\left[1 + TC1(T - T_0) + TC2(T - T_0)^2\right] \qquad (2)$$

The current source I_T is fixed at $1A$ such that to simplify the calculations for the other components in the macromodel.

The controlling voltage $V_{100,0}$ is a quadratic temperature-dependent input value for the VCCS G_T in the model. The current generated by the G_T will have the following form:

$$I(G_T) = k_{G_T} V_{73,88} V_{100,0} = k_{G_T} V_{73,88} I_T R_T(T) \qquad (3)$$

where k_{G_T} $[S/V]$ is a linear coefficient. For convenience, it is chosen to be equal to one.

Since the current source G_T is controlled by two voltages $V_{73,88}$ and $V_{100,0}$, the relation between the on-resistance value at a certain temperature $R_{ON}(T)$ and the value at nominal temperature $R_T(T_0)$ is

$$R_{ON}(T) = R_{ON}(T_0) + \frac{V_{73,88}}{I(G_T)} = R_{ON}(T_0) + \frac{1}{k_{G_T} I_T R_T(T)} \qquad (4)$$

where $R_{ON}(T_0)$ is the switch on-resistance at a temperature $T_0 = 27°C$.

The temperature coefficients $TC1$ and $TC2$, and the resistance $R_T(T_0)$ from Eq. (2) are calculated by employing the least squares sense technique.

The main advantage of this approach is that parameter extraction can be performed only from manufacturer's data, even for integrated circuits whose internal structure is unknown.

Another temperature stage with an exponential behaviour is shown in Fig. 3. Its consists of the voltage source V_{TE}, the diode D_{TE} and the correction current source I_{TE}. The exponential temperature dependence of saturation current $I_S(T)$ in the SPICE diode model [13] is determined by

$$I_S(T) = I_{So} e^{(T/T_0 - 1)E_G/(NV_T)} (T/T_0)^{XTI/N} \tag{5}$$

where I_{So} is the saturation current at T_o, T is the temperature in K (Kelvin), XTI is the saturation current temperature exponent, N is the emission coefficient, V_T is the thermal voltage, and E_G is the energy gap. The parameter extraction can be done for the desired temperature behaviour. The voltage source V_{TE} is fixed at -10V to provide a reverse voltage drop across D_{TE}. The resistor R_{TE} is set to 1Ω in order to minimize the generated thermal noise and to simplify the calculations for the other components in the stage.

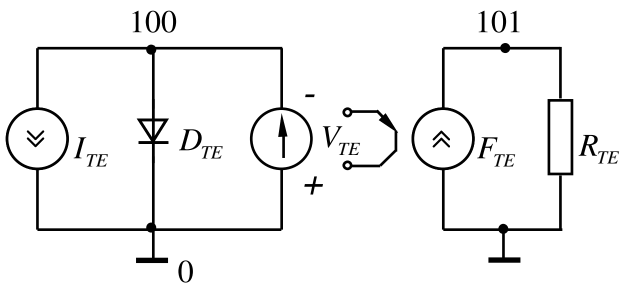


Fig. 3. Stage for generating exponential temperature dependence of the R_{ON}.

The current $I_S(T)$, generated from D_{TE}, will flow through the voltage source V_{TE} towards the ground. The current $I_S(T)$ is an exponential temperature-controlled input value for current-controlled current source (CCCS) F_{TE}, defined in the macromodel. The current generated by the F_{TE} will have the following form:

$$I(F_{TE}) = k_{1,TE} I_{VTE}(T) \tag{6}$$

The leakage current thus generated, $I(F_{TE})$ is routed through the resistor R_{TE}. The linear coefficient $k_{1,TE}$ of the CCCS is selected to be equal to one. The voltage V_{101} works as a linear temperature-controlled input value for VCCS G_T in the model.

III. MACROMODEL PERFORMANCE

The performance of the macromodel has been compared with data sheet parameters of the integrated analogue multiplexer ADG408 [14]. The existing macromodel for the ADG408 was used, with additional elements modelling a quadratic temperature function of the switch on-resistance (Fig. 2), inserted between the node 73 and the output terminal.

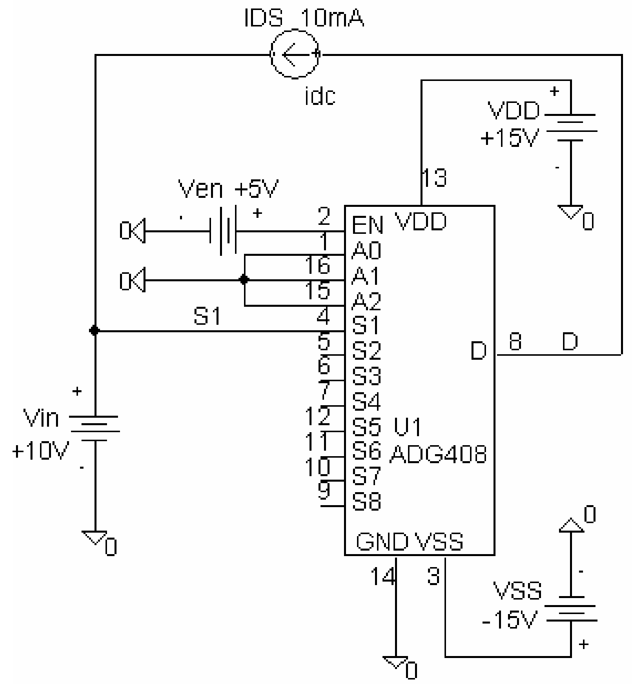


Fig. 4. Test circuit for simulation.

The test circuit for simulation shown in Fig. 4 is created following the test conditions, given in the semiconductor data book of the corresponding IC. The power supply voltages of the circuit are chosen ±15V and logic supply voltage is set to +5V. The DC input signal level V_{in}, applied to node S1 (pin 4), is set to +10V. The ideal current sources $I_{DS} = 10mA$ is connected between nodes S1 and D. For the simulation testing a parametric DC analysis is performed with the sweep parameters within the range from V_{in}=-15V to V_{in}=+15V. The enabling input voltage source $V_{EN} = +5V$ is applied to node EN (pin 2). The computer simulations are implemented for three values of the temperature, namely $25°C$, $85°C$ and $125°C$. Figure 5 and 6 shows the on-resistance versus input voltage for different temperature of the real device ADG408 and the simulation output, using the values in the netlist of Table 1. In Table 1 only additional elements to the existing model [4] are represented. Notice that the simulated response compares quite closely with the actual response. The maximum error between simulation results and manufacturer's data are not higher than 10%, which guarantees the sufficient degree of accuracy.

Table 1. SPICE netlist of the output section for ADG408 model.

```
.SUBCKT ADG408/AD-X 4 5 6 7 12 11 10 9 15 16 1 2 8 13 14 3
*  Section for control line A0, A1 and A2
*  Main Series Switch combination
*  Output and enable switch section *** see Reference [4] ***
GT 73  88 POLY(2) (73 88) (100 0) 0 0 0 0 1
S_EN_1     88  8  2  14    Sdemux
C_EN_1     2  8          4.2E-12 ; SETS CHARGE INJECTION
.MODEL  Sdemux  VSWITCH  (RON=1  ROFF=1E12  VON=2.0
VOFF=1.4)
* Temperature stage
IT 0 100 1
RT 0 100 0.95692800467
+ TC -0.02224408931100 1.306292978368239e-004
.ENDS
```

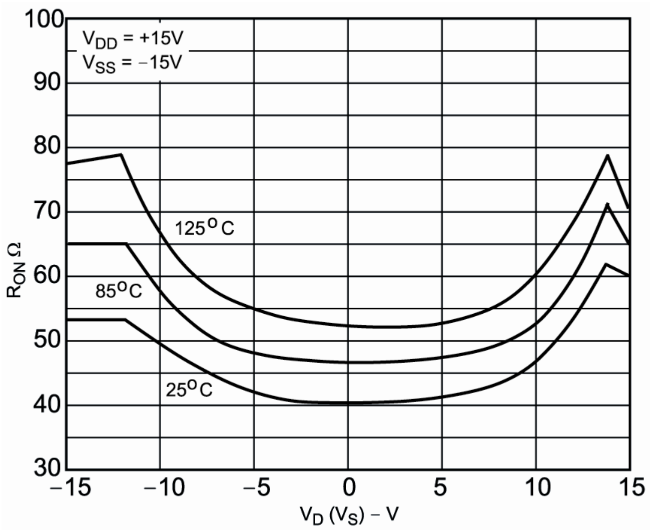


Fig. 5. ADG408 temperature effects.

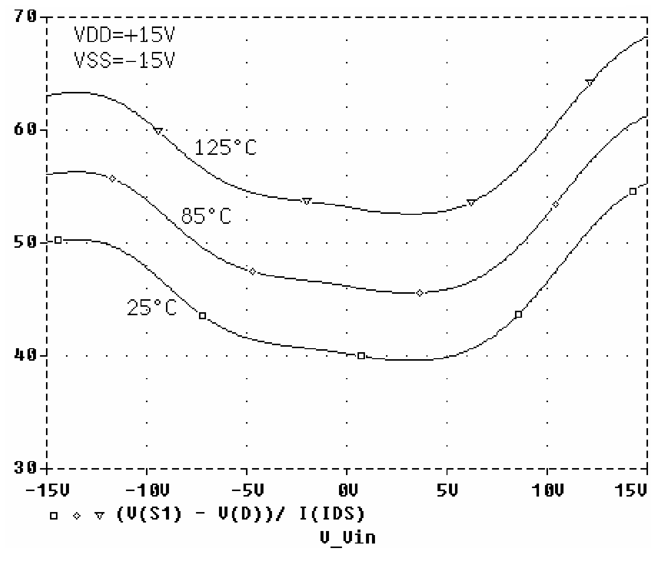


Fig. 6. ADG408 simulated temperature effects.

IV. CONCLUSIONS

This paper presents the development of the analogue multiplexer SPICE macromodel, aimed at improving the modelling of switch on-resistance as a function of temperature. Modelling of the temperature effects of the circuit is independent from the actual technical realizations and the model parameters and can only be obtained from manufacturer's data of the real IC. The macromodel can be used to simulate different kinds of analogue circuits in respect to temperature variations. The detailed comparison with semiconductor data book demonstrates a good correspondence with selected diagrams of the analogue multiplexer ADG408 from Analog Devices. The maximum error between simulation results and manufacturer's data is not higher than 10%. However, other analogue multiplexers, especially those using different topologies or technologies, may well exhibit different functions of the on-resistance R_{ON} versus temperature. These matters are the subjects of further research.

REFERENCES

[1] W. Kester, "*Switches and Multiplexers. Useful Design Techniques and Tips for Switches and Multiplexers*", Analog Devices, Norwood, MA, USA, pp. 7.62-7.66, 2005.

[2] R. L. Geiger, P. E. Allen and N. R. Strader, *VLSI Design Techniques for Analog and Digital Circuits*, pp. 289-302, McGraw-Hill, 1996.

[3] *Op Amp Applications*, Analog Devices, Norwood, MA, USA, pp. 6.128-6.129 and pp. 3.26-3.27, 2002.

[4] Switches/Multiplexers - SPICE Models, Analog Devices 2005, http://www. analog.com/.

[5] *Analog Switches and Multiplexers Macromodels*, Maxim Integrated Products 2005, http://www.maxim-ic.com/tools/spice.

[6] H. Chenmin, D. Leach and Sh. Chan, "An improved macromodel for operational amplifiers," *International Journal of Circuit Theory Applications*, vol. 18, pp. 190-203, 1990.

[7] U. Jörges, G. Jummel and G. Müller, "A macromodel of sample-and-hold circuits," *International Journal of Circuit Theory Applications*, vol. 25, pp. 483-502, 1997.

[8] "*Development of an extensive SPICE macromodel for current-feedback amplifiers*," National Semiconductor, Application Note AN 840, 1992.

[9] Cousineau, M., D. Standarovski, M. Lescure, *Analog circuit modeling method – Sample-and-Hold model development*, MIXDES 2003, Poland, June 2003, pp. 395-398.

[10] G.R. Boyle et al., "Macromodeling of integrated circuit operational amplifiers," IEEE J. of Solid-State Circuits, vol. SC-9, No 6, pp. 359-363, Dec. 1974.

[11] A.I. Kayssi, "Macromodel construction and verification," IEEE J. of Circuit&Devices, pp. 34-39, 1998.

[12] A.M. Law, "How to conduct a successful simulation study," Proceeding of the 2003 Winter Simulation Conference, pp. 66-70, 2003.

[13] *MicroSim PSpice A/D Reference Manual*, MicroSim, Irvine, California USA, Chapter 2, 1996.

[14] CMOS analogue multiplexer ADG408 - data sheet, Analog Devices, Norwood, MA, USA, 2003.

A logarithmic amplifier SPICE macromodel improved with DC offset voltages and input bias currents

Ivailo M. Pandiev[1]

Abstract – **An improved SPICE macromodel of a logarithmic amplifier is presented here, in which the offset voltage and input bias current are modelled. The simulation model is developed through modifying the existing macromodels employing the mechanism of controlled sources and sub-circuits. The proposed macromodel allows simulating analogue circuits with respect to the behaviour in both the time and frequency domains, including error parameters and temperature dependence of the input (output) offset voltage and the input bias current. As an example, model parameters for the logarithmic amplifier LOG101 (Texas Instruments) are extracted. Simulated results and selected diagrams are compared with the manufacturer's data.**

Keywords – **Analogue circuits, Offset voltage and current, Input bias current, Logarithmic amplifier, SPICE, Modelling.**

I. INTRODUCTION

Logarithmic amplifiers (log amps) are widely used for compressing signals and computation. Although digital ICs have mostly replaced the log amps in applications that require computing, engineers continue to use log amps to compress signals. Therefore, the log amp remains an essential component in many video, medical, test and measurement systems. [1-3].

Accuracy considerations for log amps are somewhat more complicated than the ones for the other operational amplifiers. This is because the transfer function is nonlinear and log amps have two inputs, each varying in a wide dynamic range. The accuracy for any combination of the input voltages is determined from the total error. This is the deviation of the actual output voltage $V_{out(actual)}$ from the ideal output voltage $V_{out(ideal)}$. Thus,

$$V_{out(actual)} = V_{out(ideal)} \pm V_{error} \tag{1}$$

The $\pm V_{error}$ represents the sum of all the individual components of error normally associated with the logarithmic amplifier when operating in current input mode. As with any transfer function, the error generated by the function itself may be referred to the output or to the inputs. The individual error components of log amps are: input/output offset voltage, input bias and offset currents, as well as their temperature drifts. For most of the integrated log amps in voltage input mode of operation, the $V_{out(actual)}$ as a function of the major components of error can be found by

$$V_{out(actual)} = (1V)(1 \pm \Delta K)\, log \frac{\dfrac{V_1}{R_1} - I_{B1} \pm \dfrac{V_{OS1}}{R_1}}{\dfrac{V_2}{R_2} - I_{B2} \pm \dfrac{V_{OS2}}{R_2}} \pm 2Nn \pm V_{OSO} \tag{2}$$

where ΔK is gain error of the output voltage, I_{B1} and I_{B2} are the bias currents of input 1 and 2, V_{OS1} and V_{OS2} – input offset voltages, N – log conformity error (usually from 0,01% to 0,06%), n – number of decades and V_{OSO} – output offset voltage.

The offset voltage and bias currents of the log amps are important parameters, taken into consideration during the design of photodiode signal compression amplifiers, analogue signal compression before ADCs, etc. An accurate model of the log amp input bias current and input/output offset voltages must therefore take into account not only the voltage and current for *T*=25°C, normally quoted on data sheet, but also the temperature drifts in the operating range.

The majority of published SPICE log amp macromodels only attempt to model input bias current and offset voltage at *T*=25°C, usually by ideal DC sources, connected to the input terminals [4, 5]. They therefore cannot be used to simulate dynamic variations in input bias current and offset voltage in function with the change of temperature. In response to this problem, here is presented a log amp SPICE macromodel, with input/output offset voltages and input bias currents improved versus temperature. The proposed simulation macromodel is developed following the design procedure given in [7].

II. BUILDING THE MACROMODELS SENSING INPUT AND OUTPUT OFFSETS

The schematic diagram of the proposed macromodel is shown in Fig. 1. It is based on a previous SPICE macromodel [4], with the additional capability to simulate the input/output offset voltage and the input bias/offset current versus temperature. In Fig. 1, there are two operational amplifier models represented as black-boxes, described in full detail in [6].

Various parameters have been modelled as follows.

A. Modelling the input and output offset voltages

These parameters are modelled by adding new elements and stages to the equivalent circuit of the existing macromodel of the log amp. As it is shown in Fig. 1, two linear voltage-controlled voltages sources (VCVSs) E_{OS1} and E_{OS2} are

[1]Ivailo M. Pandiev is with the Faculty of Electronics from Technical University of Sofia, Kliment Ohridski 8, 1000 Sofia, Bulgaria. E-mail: ipandiev@tu-sofia.bg

connected to the non-inverting inputs of the amplifiers, creating the new nodes 161 and 162. The state of those nodes will have to follow the change in the input offset voltage. Node 16 through $R_{PI} = 10\Omega$ is connected to node 6, which is the reference node of the macromodel.

The current I_{TVOS} (and I_{TVOSO}) will flow through the resistor R_{TVOS} (and R_{TVOSO}), towards the internal ground. In such a way the voltage V_{100} (and V_{101}) will depend upon the temperature.

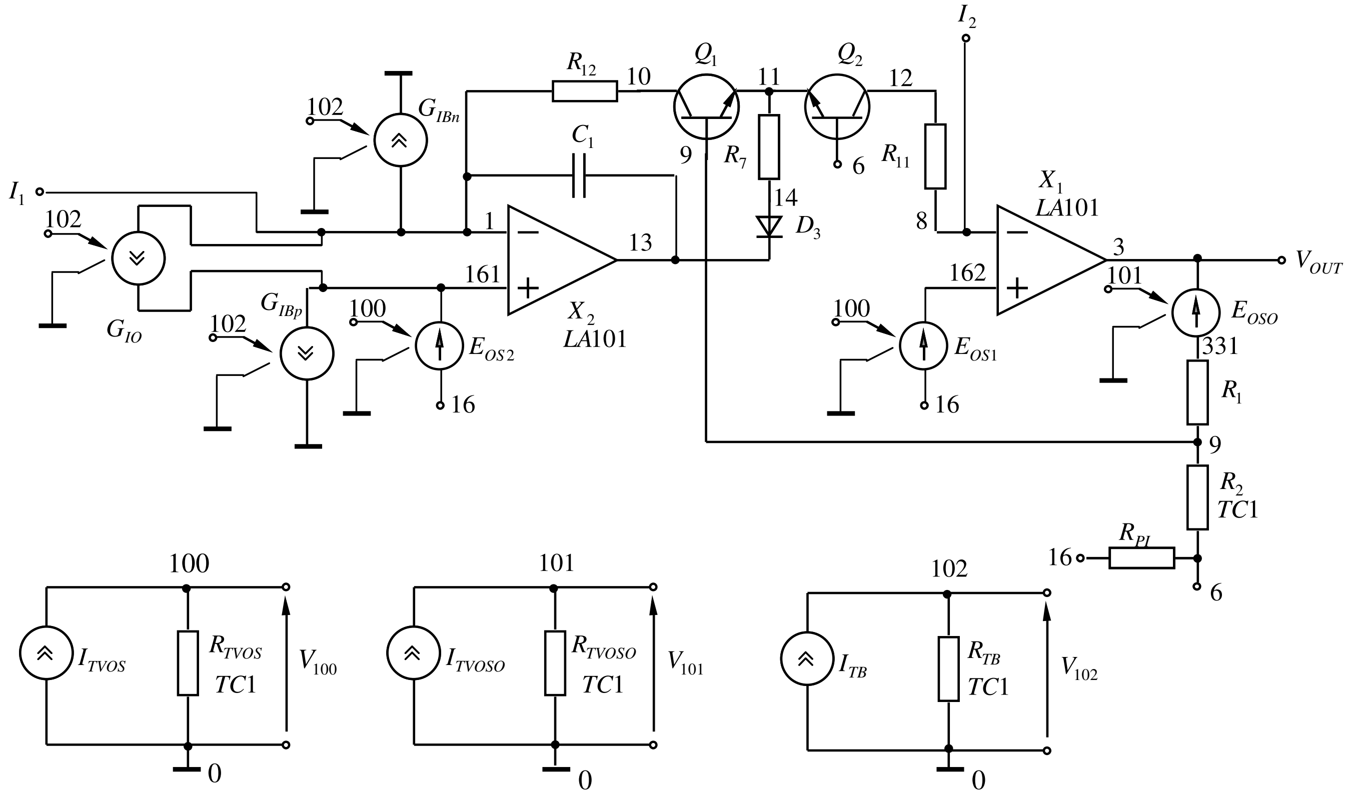


Fig. 1. Conceptual circuit of the proposed log amp SPICE macromodel.

The output offset voltage is modelled with one VCVS E_{OSO}, connected between nodes 3 (output terminal) and 331. The temperature-dependent controlling voltages of the E_{OS1}, E_{OS2} and E_{OSO} come from separate temperature stages in Fig. 1. These additionally defined stages consist of ideal current sources (I_{TVOS} and I_{TVOSO}) and SPICE temperature-dependent resistors (R_{TVOS} and R_{TVOSO}), controlled with the equation [8]:

$$R\ (T) = R(T_0)\left[1 + TC1(T - T_0) + TC2(T - T_0)^2\right] \quad (3)$$

where $R(T_0)$ is the value of the resistor at $T_0 = 27°C$ (SPICE-Option TNOM), T is the temperature in $°C$, $TC1$ is the linear temperature coefficient and $TC2$ is the quadratic temperature coefficient. A sufficient degree of accuracy for the purpose of modelling offset voltage temperature drift has been provided by choosing $R(T)$ as a linear temperature-dependent resistor ($TC2 = 0$), having the following characteristic equation:

$$R(T) = R(T_0)[1 + TC1(T - T_0)] \quad (4)$$

$$V_{100,0} = I_{TVOS} R_{TVOS}(T_0)[1 + TC1(T - T_0)] \quad (5a)$$

$$V_{101,0} = I_{TVOSO} R_{TVOSO}(T_0)[1 + TC1(T - T_0)] \quad (5b)$$

The parameters of the current I_{TVOS} (and I_{TVOSO}) and the resistor R_{TVOS} (and R_{TVOSO}) are calculated so that the temperature stages provide a voltage $V_{100} = V_{101} = 1mV$ at the temperature $T_0 = 27°C$. Signals, generated at nodes 100 and 101, are used for forming the characteristic equations of E_{OS1}, E_{OS2} and E_{OSO} as follows:

$$V(E_{OS1}) = k_{o,EOS1} + k_{1,EOS1} V_{100}, \quad (6a)$$

$$V(E_{OS2}) = k_{o,EOS2} + k_{1,EOS2} V_{100}, \quad (6b)$$

$$V(E_{OSO}) = k_{o,EOSO} + k_{1,EOSO} V_{102}. \quad (6c)$$

The linear temperature coefficient $TC1$ from Eq. (4), and characteristic coefficients from Eq. (6a), (6b) and (6c) are calculated by employing the least squares sense technique.

B. Modelling the input bias current and the input offset current

The analysis of data sheets for the log amps has shown that in most cases the input characteristics could be approximated by nonlinear functions, in which input bias currents double with every 10°C temperature rise. Normally, semiconductor data books, provided by IC-manufacturers, contain information regarding the temperature dependence of the input currents that could be used in modelling the real log amps.

The input bias/offset currents are implemented by non-linear voltage-controlled current sources (VCCSs) G_{IBn}, G_{IBp} and G_{IO}, connected to the inverting and non-inverting inputs of the amplifiers and an additionally defined temperature stage consisting of an ideal current source I_{TB} and a temperature-dependent resistor R_{TB}, as shown in Fig. 1.

The temperature stage provides a voltage V_{102} for modelling linear temperature dependences using the SPICE model (4) of the linear temperature-dependent resistor ($TC2 = 0$). For convenience, the parameters of the current I_{TB} and the resistor R_{TB} are extracted so that the temperature stage provides a voltage $V_{102} = 1V$ at the temperature $T_0 = 27°C$. Since SPICE-simulators will give an error message if the resistor R_{TB} goes negative at any temperature, the value of the linear coefficient $TC1$ can be obtained by

$$TC1 \leq 10^{-1} / (T_0 - T_{\min}) \tag{7}$$

where T_{min} is the minimum temperature in the operating range of the real logarithmic amplifier.

The controlling voltage V_{102} is a linear temperature-controlled input value for the VCCSs G_{IBn}, G_{IBp} and G_{IO} in the model. The approximation of the non-linear behaviour is realized using appropriate polynomial coefficients of the controlled sources. A sufficient degree of accuracy for the purpose of modelling temperature dependence has been provided by choosing the polynomial sources realized by sixth-order functions. The VCCS functions are given by

$$I(G_{IBn}) = \sum_{i=0}^{n=6} k_{i,IBn} V_{IBn}^{i} , \tag{8a}$$

$$I(G_{IBp}) = \sum_{i=0}^{n=6} k_{i,IBp} V_{IBp}^{i} , \tag{8b}$$

$$I(G_{IO}) = \sum_{i=0}^{n=6} k_{i,IO} V_{IO}^{i} . \tag{8c}$$

The polynomial coefficients ($k_{i,IBn}, k_{i,IBp}$ and $k_{i,IO}$ $i = \overline{0,n}$) are calculated using MATLAB. In particular, the function *polyfit.m* is implemented to define the unknown coefficients of the controlled sources. In general, the function *polyfit* (*x,y,n*) for a given data in a vector $x\,(T,°C)$ finds a *n*-th order polynomial function *p* such that $p(x)$ fits the data in a vector *y* (positive and negative input bias current) by using the least squares sense. The main advantage of this approach is that parameter extraction can be performed only from data sheet parameters, even for circuits whose internal structure is unknown.

III. Macromodel performance

The verification of the SPICE macromodel (Fig. 1) with temperature-dependent input and output offset voltages and currents is carried out by comparing simulation results with data sheet parameters of integrated log amp LOG101 [9]. The sub-circuit listing, given in Table 1, contains elements (marked in bold style) for simulating the DC voltage and current error of the LOG101. The test circuit for simulation is created following the test conditions given in the semiconductor data book of the corresponding IC. The power supply voltages of the circuit are chosen to be ±15V. The simulation modelling is implemented within EDA OrCAD. During the process of simulation a DC sweep analysis is specified and performed with the temperature ranging from -5°C to 75°C. In the Fig. 2, 3, 4 and 5 are presented the simulation output ($x_{,M}$), the data sheet parameters (x_{DS}) for the input/output offset voltages, input bias/offset currents and error ($\delta = [(x_M - x_{DS}) / x_M].100\%$) as a function of the temperature. It can be observed, that the accuracy of the macromodel is quite good ($\delta < 5\%$).

Table 1. SPICE netlist for LOG101 macromodel

```
* PINOUT ORDER I1 VOUT V+ V- GND I2
* PINOUT ORDER  1   3   4  5  6   8
.SUBCKT LOG101 1 3 4 5 6 8
R1 9 331 3340
R2 6 9 212 TC1=3.47M
Q1 10 9 11 QL101
Q2 12 6 11 QL101
X1 162 8 4 5 3 LA101
X2 161 1 4 5 13 LA101
* Input and output offset voltages
EOSO 3 331 101 0 3.00386666666667
EOS1 162 16 100 16 0.2663
EOS2 161 16 100 16 0.2663
* Input bias and offset currents
GIBn 1 0 POLY(1) (102 0) 0.007518015195925
+-0.04450102859145  0.109768095815289  -0.144422090643820  +0.10689777201416
-0.042204603929841 0.006943840145432
GIBp 161 0 POLY(1)   (102 0) 0.007518015195925
+-0.04450102859145 0.109768095815289 -0.144422090643820 +0.106897772014167
-0.042204603929841 0.006943840145432
GIO 1 161 POLY(1) (102 0) 0.0007518015195925
+-0.0044501028591453 0.0109768095815287
+-0.0144422090643818 0.0106897772014165
+ -0.0042204603929840 0.0006943840145432
* Temperature stage for input offset voltage
ITVOS 0 100 1m
RTVOS 100 0 1 TC1=0.00813516
* Temperature stage for output offset voltage
ITVOSO 0 101 1m
RTVOSO 101 0 1 TC1=6.436148963546978e-004
* Temperature stage for input bias and offset current
ITB 0 102 1
RTB 102 0 1 TC1=0.001
RPI 16 6 10
R5 0 8 1E12
R6 0 1 1E12
R7 14 11 670
C1 1 13 154E-12
D3 14 13 DS
R8 14 13 1E8
R9 11 1 1E12
R10 8 11 1E12
R11 8 12 82
R12 10 1 82
.ENDS
.SUBCKT LA101 3 2 7 4 6
* SPICE netlist for LA101 ***see Reference [6] ***
.ENDS
* END MODEL LOG101
```

For the macromodel, the simulation time is observed to be comparable to the existing models from the standard SPICE libraries. At present, the advantage of such *complex* macromodels seems to be the study of the effects of specific component parameters of the complete electronic sub-systems.

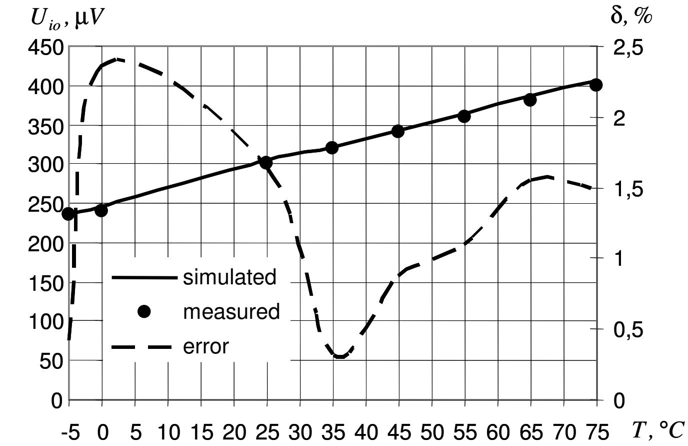


Fig. 2. Input offset voltage as a function of temperature for LOG101

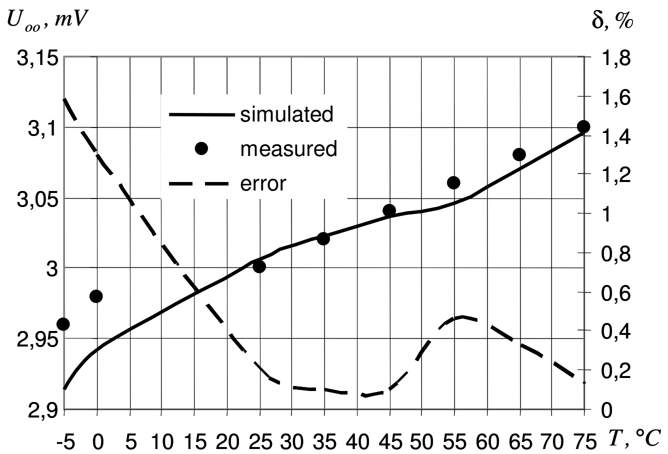


Fig. 3. Output offset voltage as a function of temperature for LOG101

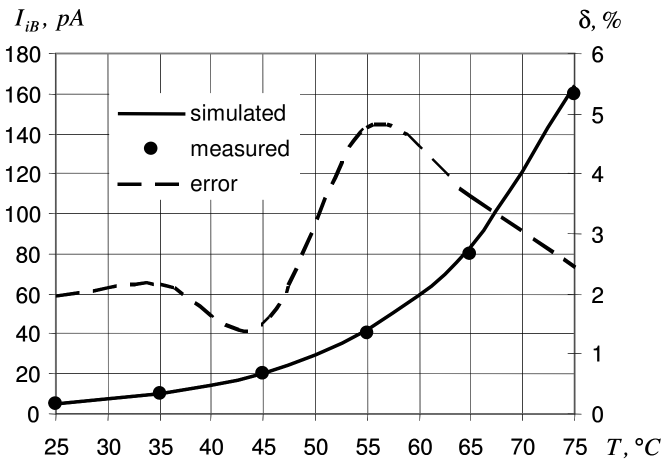


Fig. 4. Input bias current as a function of temperature for LOG101

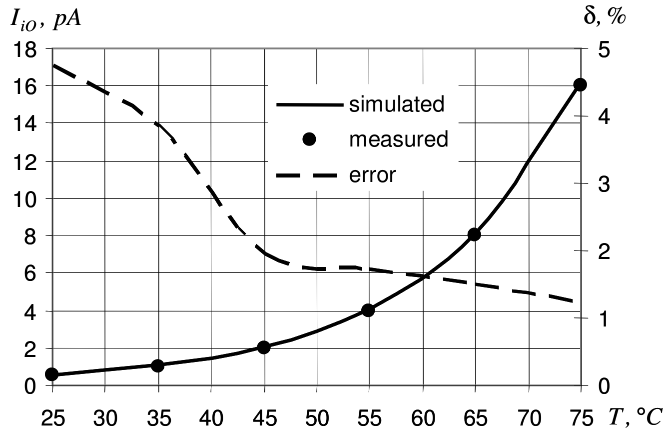


Fig. 5. Input offset current as a function of temperature for LOG101

IV. Acknowledgment

The author wish to thank the reviewers who offered useful and constructive comments.

V. Conclusions

This paper presents the new developments of log amp simulation macromodel, aimed at improving the modelling of the offset voltage and input bias current versus temperature. The model is independent from actual technical realizations and is based on compromises regarding the exact circuit structure in the model. The efficiency of the model was proved by comparison of simulation results and data sheet parameters of the integrated logarithmic amplifier LOG101 from Texas Instruments.

References

[1] R. L. Geiger, P. E. Allen and N. R. Strader, *VLSI Design Techniques for Analog and Digital Circuits*, pp. 289-302, McGraw-Hill, 1996.

[2] Tietze, U., Ch. Schenk. *Halbleiter-Schaltungstechnik*. 11.1 Auflage. Springer-Verlag, pp. 785-800, 2001.

[3] *Integrated dc logarithmic amplifiers*, AN 3611, Maxim-Dallas Semiconductor, 2005, http://maxim-ic.com.

[4] *Simulation models – logarithmic amplifiers*, Texas Instruments, 2005, http://www.ti.com.

[5] S. R. Chowdhury, C. Pramanik, H. Saha, "ASIC Design of the Linearisation Circuit of a PTC Thermistor", *18th Int. Conf. on VLSI Design'05*, pp. 866-869, 2005.

[6] Boyle G., B. Cohn. D. Pederson, J. Solomon. "Macromodeling of Integrated Circuit Operational Amplifiers". *IEEE Journal of Solid State Circuits*, Vol. SC-9, No.6, 1974, pp. 353- 363.

[7] I.M. Pandiev, "Model building and testing procedure - An analogue multiplexer SPICE macromodel improved with temperature effects," ICEST 2006 (submitted).

[8] *MicroSim PSpice A/D Reference Manual*, MicroSim, Irvine, California USA, Chapter 2, 1996.

[9] *Precision logarithmic amplifier LOG101* - data sheet, Texas Instruments, USA, 2002.

A Survey of Three System-on-Chip Buses: AMBA, CoreConnect and Wishbone

Milica Mitić and Mile Stojčev

Abstract – **This article gives an overview of three popular bus organized computer architectures (CAs), called AMBA, CoreConnect and Wishbone. It starts with a brief introduction to on-chip CA, then looks at bus organizations, and concludes with a discussion related to a comparative performance analysis of all three CAs.**

Keywords – **SoC buses, on-Chip bus, AMBA, CoreConnect, Wishbone**

I. INTRODUCTION

Shrinking process technologies and increasing design sizes have led to highly complex billion-transistor integrated circuits (ICs). As a consequence, manufacturers are integrating increasing numbers of components on a chip. A heterogeneous system-on-a-chip (SoC) might include one or more programmable components such as general purpose processsors cores, digital signal processor cores, or application-specific intellectual property (IP) cores, as well as an analog front end, on-chip memory, I/O devices, and other application specific circuits [1].

On-chip bus organized CA is among the top challenges in CMOS SoC technology due to rapidly increasing operation frequencies and growing chip size. Usually, IP cores, as constituents of SoCs, are designed with many different interfaces and communication protocols. Integrating such cores in a SoC often requires insertion of suboptimal glue logic. Standards of on-chip bus structures were developed to avoid this problem. Currently there are a few publicly available bus architectures from leading manufacturers, such as CoreConnect from IBM [2], AMBA from ARM [3], SiliconBackplane from Sonics [4], and others. This paper focuses on SoC CAs providing a survey of three popular bus organized CAs, called AMBA, CoreConnect and Wishbone from an industrial and research viewpoint.

II. ON-CHIP COMMUNICATION ARCHITECTURES

A. Background

The design of on-chip CAs addresses the following three issues [5]:

1. **Definition of CA topology** - defines the physical structure of the CA. Numerous topologies exist, ranging from single shared bus to more complex architectures such as bus hierarchies, token ring, crossbar, or custom networks.
2. **Selection and configuration of the communication protocols** – for each channel/bus in the CA, communication protocols specify the exact manner in which communication transaction occur. These protocols include arbitration mechanisms (e.g. round robin access, priority-based selection [2], [3], time division multiplexed access [4], which are implemented in centralized or distributed bus arbiters.
3. **Communication mapping** – refers to the process of associating abstract system-level communications with physical communication paths in the CA topology [5].

B. Topologies

In respect to topology on-chip communication architectures can be classified as:

Shared bus: The system bus is the simplest example of a shared communication architecture topology and is commonly found in many commercial SoCs [6]. Several masters and slaves can be connected to a shared bus. A block, bus arbiter periodically examines accumulated requests from the multiple master interfaces, and grants access to a master using arbitration mechanisms specified by the bus protocol.

Hierarchical bus: this architecture consists of several shared busses interconnected by bridges to form a hierarchy. SoC components are placed at the appropriate level in the hierarchy according to the performance level they require. Low-performance SoC components are placed on lower performance buses, which are bridged to the higher performance buses so as to not burden the higher performance SoC components. Commercial examples of such architectures include the AMBA bus [3], CoreConnect [2]. Transactions across the bridge involve additional overhead, and, during the transfer, both buses remain inaccessible to other SoC components. Hierarchical buses offer large throughput improvements over the shared busses due to: (1) decreased load per bus; (2) the potential for transactions to proceed in parallel on different buses; and (3) multiple ward communications can be preceded across the bridge in a pipelined manner [5].

Ring: in numerous applications, ring based applications are widely used, such as network processors, ATM switches [2], [5]. In a ring, each node component (master/slave) communicates using a ring interface, usually implemented by a token-pass protocol.

C. On-Chip communication protocols

Communication protocols deal with different types of resource management algorithms used for determining access right to shared communication channels. From this point of view, in the rest of this section, we will give a brief comment related to the main features of the existing communication protocols, which are:

Milica Mitić and Mile Stojcev are with the Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Nis, Serbia and Montenegro,
E-mail: [milicam, stojcev]@elfak.ni.ac.yu

Static-priority: employs an arbitration technique. This protocol is used in shared-bus communication architectures. A centralized arbiter examines accumulated requests from each master and grants access to the requesting master that is of highest priority. Transactions may be of non-preemptive or preemptive type. AMBA and CoreConnect use this protocol [3], [2].

Time Division Multiple Access (TDMA): the arbitration mechanism is based on a timing wheel with each slot statically reserved for unique master. Special techniques are used to alleviate the problem of wasted slots. Sonics uses this protocol [4].

Lottery: a centralized lottery manager accumulates request for ownership of shared communication resources from one ore more masters, each of which is, statically or dynamically, assigned a number of "lottery tickets" [7].

Token passing: this protocol is used in ring based architectures. A special data word, called token, circulates on the ring. An interface that receives a token is allowed to initiate a transaction. When the transaction completes, the interface releases the token and sends it to the neighboring interface. For example, VCI uses this protocol [8]

Code Division Multiple Access (CDMA): this protocol has been proposed for sharing on-chip communication channel. In a sharing medium, it provides better resilience to noise/ interference and has an ability to support simultaneously transfer of data streams. But this protocol requires implementation of complex special direct sequence spread spectrum coding schemes, and energy/battery inefficient systems such as pseudo-random code generators, modulation and demodulation circuits at the component bus interfaces, and differential signaling [9].

III. SoC BUSES OVERVIEW

In the sequel an overview of the more relevant SoC CAs (AMBA, CoreConnect and Wishbone) will be given. Due to space limitation the discussion will be focused on describing the more distinctive features of each of them.

A. AMBA

AMBA (*Advanced Microcontroller Bus Architecture*) [3], [10], is a bus standard devised by ARM with aim to support efficient on-chip communications among ARM processor cores. Nowadays, AMBA is one of the leading on-chip busing systems used in high performance SoC design. AMBA (see Fig. 1) is hierarchically organized into two bus segments, system- and peripheral-bus, mutually connected via bridge that buffers data and operations between them. Standard bus protocols for connecting on-chip components generalized for different SoC structures, independent of the processor type, are defined by AMBA specifications. AMBA does not define method of arbitration. Instead it allows the arbiter to be designed to best suit the applications needs. The three distinct buses specified within the AMBA bus are:

- **ASB (*Advanced System Bus*) –** first generation of AMBA system bus used for simple cost-effective designs that support burst transfer, pipelined transfer operation, and multiple bus masters.
- **AHB (*Advanced High-performance Bus*)** – as a later gene–ration of AMBA bus is intended for high performance high-clock synthesizable designs. It provides high-bandwidth communication channel between embedded processor (ARM, MIPS, AVR, DSP 320xx, 8051, etc.) and high performance peripherals/ hardware accelerators (ASICs MPEG, color LCD, etc), on-chip SRAM, on-chip external memory interface, and APB bridge. AHB supports multiple bus masters operation, peripheral and burst transfer, split transactions, wide data bus configurations, and non tristate implementations. Constituents of AHB are: AHB-master, slave-, arbiter-, and –decoder.
- **APB (*Advanced Peripheral Bus*** – is used to connect general purpose low-speed low-power peripheral devices. The bridge is peripheral bus master, while all buses devices (Timer, UART, PIA, etc) are slaves. APB is static bus that provides a simple addressing with latched addresses and control signals for easy interfacing.

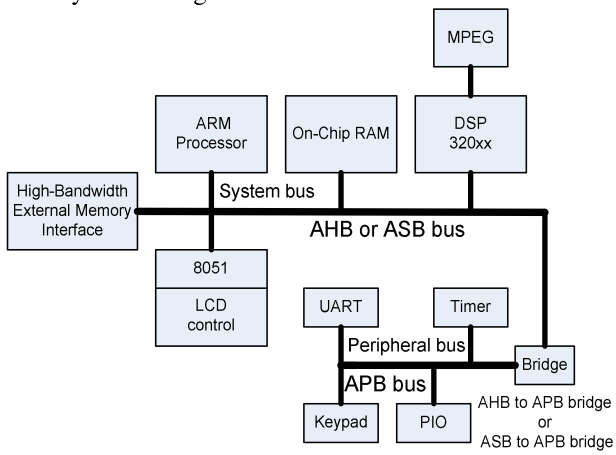


Fig. 1. AMBA based system architecture

Recently, two new specifications for AMBA bus, Multi-Layer AHB and AMBA AXI, are defined. [11], [12]. Multi-layer AHB provides more flexible interconnect architecture (matrix which enables parallel access paths between multiple masters and slaves) with respect to AMBA AHB, and keeps the AHB protocol unchanged. AMBA AXI is based on the concept point-to-point connection.

Goog overview papers related to AMBA specifications are references [11], [12] and [13].

B. CoreConnect

CoreConnect [2] is an IBM-developed on-chip bus. By reusing of processor, subsystem and peripheral cores, supplied from different sources, it enables their integration into a single VLSI design. CoreConnect is hierarchically organized architecture. It is comprised of three buses that provide an efficient interconnection of cores, library macros, and custom logic within a SoC (see Fig. 2).

- **PLB (*Processor Local Bus*) –** is the main system bus. It is synchronous, multi-master, central arbitrated bus that allows achieving high-performance and low-latency on-chip communication. Separated address, and data buses support concurrent read and write transfers. PLB macro, as glue logic, is used to interconnect various master and slave macros. Each PLB master is attached to the PLB through separate addresses, read-data and write-data buses, and other control signals. PLB slaves are attached to PLB through shared, but decoupled, address, read data, and write data buses. Up to 16 masters can be supported by the arbitration unit, while there are no restrictions in the number of slave devices [10].

- **OPB (*On-chip Peripheral Bus*) -** is optimized to connect lower speed, low throughput peripherals, such as serial and parallel port, UART, etc. Crucial features of OPB are: fully synchronous operation, dynamic bus sizing, separate address and data buses, multiple OPB bus masters, single cycle transfer of data between bus masters, single cycle transfer of data between OPB bus master and OPB slaves, etc. OPB is implemented as multi-master, arbitrated buses. Instead of tristate drivers OPB uses distributed multiplexer. PLB masters gain access to the peripherals on the OPB bus through the OPB bridge macro. The OPB bridge acts as a slave device on the PLB and a master on the OPB.
- **DCR bus (*Device Control Register bus*)** – is a single master bus mainly used as an alternative relatively low speed datapath to the system for: (a) passing status and setting configuration information into the individual device-control-registers between the Processor Core and others SoC constituents such as Auxilliary Processors, On-Chip Memory, System Cores, Perhipheral Cores, etc; and (b) design for testability purposes. DCR is synchronous bus based on a ring topology implemented as distributed multiplexer across the chip. It consists of 10-bit address bus and 32-bit data bus. CoreConnect implements arbitration based on a static priority, with programmable priority fairness.

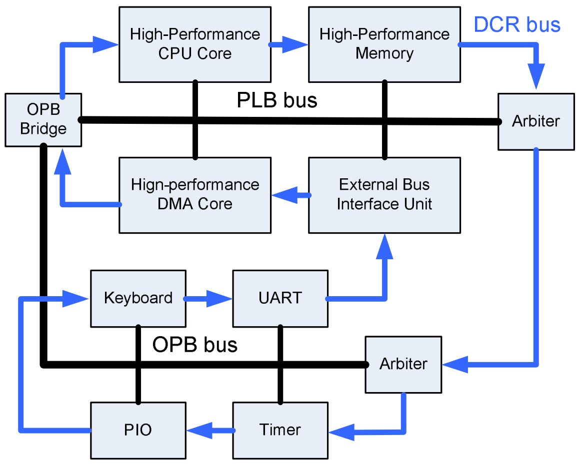


Fig. 2. CoreConnect bus based system

C. Wishbone

Wishbone [14] bus architecture was developed by Silicore Corporation. In August 2002, OpenCores (organization that promotes open IP cores development) put it into the public domain. This means that Wishbone is not copyrighted and can be freely copied and distributed.

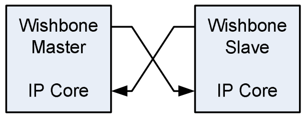


Fig. 3. Point to point interconnection

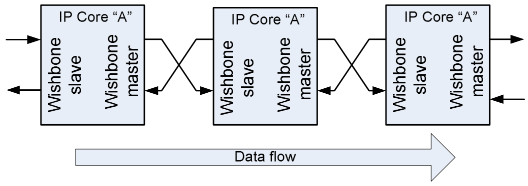


Fig. 4. Dataflow interconnection

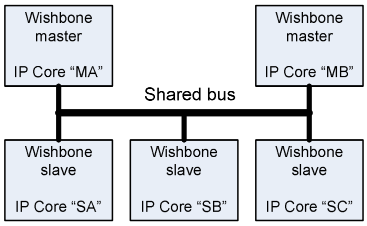


Fig. 5. Shared bus

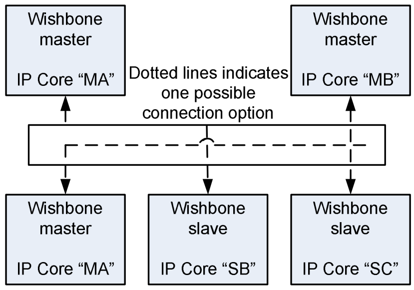


Fig. 6. Crossbar switch interconnection

The Wishbone defines two types of interfaces, called master and slave. Master interfaces are IPs which are capable of initiating bus cycles, while slave interfaces are capable of accepting bus cycles [10]. The hardware implementations support various types of interconnection topologies such as:
point-to-point connection (Fig. 3) – used for direct connection of two participants that transfer data according to some handshake protocol

a) dataflow interconnection (Fig. 4) – used in linear systolic array architectures for implementation of DSP algorithms
b) shared bus (Fig. 5)- typical for MPSoCs organized around single system bus
c) crossbar switch interconnection (Fig. 6) - usually used in MPSoCs when more than one masters can simultaneously access several different slaves. The master requests a channel on the switch, once this is established, data is transferred in a point-to-point manner.

The Wishbone supports different types of bus transactions, such as read/write, implementing blocking/unblocking access. A Read-Modify-Write transfer is also supported.
Wishbone doesn't define hierarchical buses. In applications where two buses should exist, one slow and one fast, two separated Wishbone interfaces could be created.
Designer can also choose arbitration mechanism and implements it to best fit the application needs.

TABLE 1: SOC BUSES FEATURES OVERVIEW

	Bus topology					**Arbitration**						**Bus width[bit]**		**Transfers**				
Name	Point-to-point	Ring	Shared	Hierarchical	Int. network	Static priority	TDMA	Lottery	Round-robin	Token passing	CDMA	Data bus	Address bus	Handshaking	Split	Pipelined	Burst	Op. freq.
AMBA	-	-	-	x	-	2*	2*	2*	2*	2*	2*	5*	32	x	x	x	x	8*
Core Connect	-	1*	-	1*	-	3*	-	-	-	-	-	6*	7*	x	x	x	x	9*
Wishbone	x	x	x	-	x	4*	4*	4*	4*	4*	4*	8, 16, 32, 64	1-64	x	n/a	-	x	8*

Exceptions for Table 1: 1* Data lines shared, control lines point-to-point ring; 2* Application specific except for APB which requires no arbitration; 3* Programmable priority fairness; 4* Application specific, arbiter can be designed regarding to the application requirements; 5* For AHB and ASB bus width is 32, 64, 128 or 256 byte, for APB 8, 16 or 32 byte; 6* For PLB bus width is 32, 64, 128 or 256 byte, for OPB 8, 16 or 32 byte and for DCR 32 byte; 7* For PLB and OPB bus width is 32 byte, and for DCR 10 byte; 8* User defined operating frequency; 9* Operating frequency depending on PLB width

IV. COMPARISON OF SOC BUS

In Table 1 are given some common features for presented SoC buses, such as topology, arbitration, transfers, and bus width. All presented buses are synchronous.

AMBA and CoreConnect are hierarchical buses. Wishbone does not defines hierarchical bus interconnection, but allows various other possible interconnections, such as point-to-point, ring, unilevel shared bus, crossbar switch interconnection, etc.

Arbitration method for AMBA and Wishbone is application specific, which means that arbiter can be designed regarding to the application requirements. CoreConnect defines static priority.

Presented SoC buses support various transfer types. All support handshaking, split transfer and burst transfer, while pipelined transfer support AMBA and CoreConnect, but not Wishbone.

Address and data bus width are configurable. For AMBA and CoreConnect data bus width depends on type of the bus (for AHB and ASB bus width is 32, 64, 128 or 256 byte, for APB 8, 16 or 32 byte and for PLB bus width is 32, 64, 128 or 256 byte, for OPB 8, 16 or 32 byte and for DCR 32 byte).

Operating frequency is for all buses user defined. CoreConnect defines maximum frequency depending on the PLB width (for 32 b PLB width maximal frequency is 256 MB/s, for 64 b PLB width 800 MB/s and for 128 b PLB width, 2.9 GB/s)

V. CONCLUSION

Complex VLSI IC design is being revolutionized by the widespread adoption of the SoC paradigm. The benefits of the SoC approaches are numerous, including improvements in system performance, cost, size, power dissipation, and design turn-around time. In order to exploit these advantages to the fullest, system design methodology must optimize CA requirements. During this, we have defined the on-chip CA as a fabric that integrates the various SoC components that provides them with a mechanism for the exchange data. This paper gives an overview of three popular on chip CAs, called AMBA, CoreConnect and Wishbone. At the start a background material concerning typical topologies and communication protocols is presented. In the central part an overview of most widely used on-chip CAs is provided. Finally, a short analysis related to the possibilities of all three buses is given.

REFERENCES

1. Bennini L., DeMicheli G., *Networks on Chips: A New SoC Paradigm*, IEEE Computer, Vol. 35, No. 1, January 2002, pp. 70- 78.
2. http://www.ibm.com/chips/products/coreconnect
3. http://www.arm.com
4. http://www.sonicsinc.com
5. Lahiri K., Raghunathan A., Dey S., *Design Space Exploration for Optimizing On-Chip Communication Architectures*, IEEE Trans. on CAD of ICs and Systems, Vol. 23, No. 6, June 2004, pp. 952 – 961.
6. OMI 324 PI Bus, Rev. 0.3d, OMI standards Draft, 1994 Siemens.
7. Dally W. J., Towel B., *Principles and Practices of Interconnection Networks*, Elsevier, Amsterdam, 2004.
8. Virtual Component Interface Standard version (OCB 2.2.0) VCI Alliance, April 2001, http://www.vsi.org, available at January 2006.
9. Shandhag N.R., Reliable *and Efficient System-on-Chip Design*, IEEE Computer, Vol. 37, No. 3, pp. 42-50, March 2004.
10. Ayala J., Lopez-Vellejo M., Bertozzi D., Benini L., *State-of-the-Art SoC Communication Architectures*, pp. 20.1- 20.22, in Embedded Systems Handbook, ed by Zurawski R., CRC Press, Boca Raton, 2006.
11. ARM.AMBA Specifications v2.0, 1999.
12. ARM.AMBA Multi-Layer AHB Overview, 2001.
13. ARM.AMBA AXI Protocol Specifications, 2003.
14. http://www.opencores.org/projects.cgi/web/wishbone/wishbone

29 June – 1 July 2006, Sofia, Bulgaria

Delay Locked Loop with Double Edge Synchronization

Goran S. Jovanović[1] and Mile K. Stojčev[2]

Abstract – **In CMOS multistage clock buffer design, the duty-cycle of clock is liable to be changed when the clock passes through several buffer stages. The pulse-width may be changed due to unbalance of the *p* and *n* MOS transistors in the long buffer. This paper describes a delay locked loop with double edge synchronization for use in a clock alignment process. Results of its SPICE simulation, that relate to 1.2μm CMOS technology, shown that the duty-cycle of the multistage output pulses can be precisely adjusted to (50±1)% within the operating frequency range, from 55MHz up to 166MHz.**

Keywords – **DLL, Duty Cycle Corrector, Delay Line.**

I. Introduction

Almost all contemporary digital VLSI systems and other digital systems rely on clock pulses to control the movement of data. To reach the highest circuit speed in CMOS applications, the clock distribution system must be carefully designed. A great deal of attention has been paid to clock recovery, clock regeneration, timing, and distribution [1].

Automatic control techniques, such as Phase-Locked Loop (PLL) and Delay-Locked Loop (DLL) have been widely used in high-speed clock alignment applications such as double-data rate (DDR) SDRAMs, pipelined microprocessors, network processors, etc. [2].

In a PLL implementation the chip has its own reference clock oscillator (VCO) that is phase-locked to an external reference clock. In general, a PLL clock aligner is superior in applications where noise on the reference clock dominates, and self-induced jitter within the VCO is negligible. On the other hand, a DLL provides superior performance when a clean reference clock is available. A DLL is commonly used to lock the phase of the buffered clock to that of the input data. Typically, we meet this in applications where no clock synthesis is required, such as often the situation for multi-chip digital systems with well-designed system clock distribution network [2].

In high-speed design a multistage clock buffer implemented with a long inverter chain is often needed to drive a heavy capacitive load. For these designs, as well as for applications in which the timing of both edges of the clock is critical [10], it is difficult to keep the clock duty cycle at its ideal value 50 %, primarily due to various asymmetries in signal paths and unbalances of the *p* and *n* transistors in the long buffer. As a consequence the clock duty cycle will deteriorate from 50 %, and in the worst case, the clock pulse may disappear inside the clock buffer, as the pulse width becomes too narrow or too wide [6].

Duty-cycle distortion is usually addressed in PLLs by simply running the PLL's VCO at twice the system frequency and using a post divider triggered on one edge of the VCO output to produce the output clock of the PLL. This ensures good 50 % duty cycle. In a DLL, however no frequency multiplication is possible. Therefore, the duty-cycle of the output signal must be corrected to 50 %. A conventional solution is to attach duty-cycle correction circuit to the clock output driver with the price of added area [4].

In this paper, we describe a new structure of a DLL circuit with clock alignment capability of both leading and trailing output pulse edges. This circuit can be used to obtain correct the duty-cycle factor (50 %) in a multistage clock buffer.

[1]Goran S. Jovanović is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedova 14, 18000 Niš, Serbia and Montenegro, Email: joga@elfak.ni.ac.yu

[2]Mile K. Stojčev is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedova 14, 18000 Niš, Serbia and Montenegro, Email: stojcev@elfak.ni.ac.yu

II. Clock Aligners

The goal of clock distribution network is to organize clocks so that the delays from the source point of each clock or clock phase to its destination points are identical. In reality, however, no matter how each clock path is constructed, due to variations in wire delays and driver delays, no uniform and possibly time-varying clock load, and negative effects of supply and substrate noise, any two paths within the VLSI IC will always have a delay difference. A clock aligner's task is to phase-align a chip internal clock with a reference clock, effectively removing the variable buffer delay and reducing uncertainty in clock phase between communicating VLSI IC constituents. Clock aligners (see Fig. 1) can be built using either PLLs or DLLs.

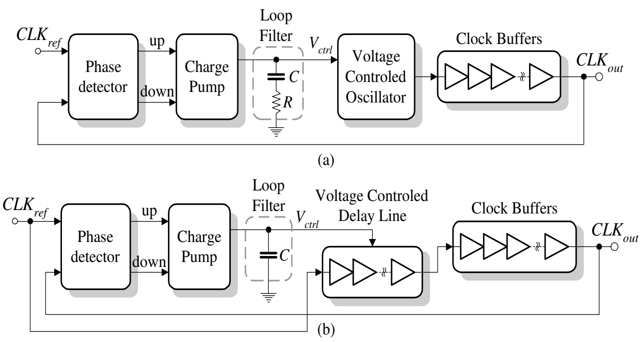


Fig. 1. Clock aligner implementations: (a) PLL clock aligner, (b) DLL clock aligner

In a PLL implementation (Fig. 1(a)) the circuit has its own oscillator (VCO) that is phase-locked to a reference clock. The phase shift introduced by the buffer delay, T_B, is assumed to be changing as a consequence of wiring delay, temperature and voltage variations, etc. The buffer's delay is eliminated by inclusion in the control loop.

In DLL implementation (Fig. 1(b)) a variable delay line (VCDL) is inserted between the reference clock, CLK_{ref}, and the output clock, CLK_{out}. The delay is regulated so that $T_D+T_B=N*T_{ref}$ (usually N=1). The buffer's delay is placed within a control loop and is eliminated.

PLL and DLL implementations have complementary advantages and disadvantages. Table I summarizes the main features of PLL and DLL clock aligners.

TABLE I COMPARISONS OF PLL AND DLL

PLL	DLL
– jitter accumulation	– no jitter accumulation
– higher-order system	– 1st-order system
– can be unstable	– always stable
– hard to design	– easier to design
– costly to integrate LF	– easier to integrate LF
– less referent signal dependent	– referent signal dependent
– easy frequency multiplication	– difficult frequency multiplication
	– limited locking range

Many factors control the speed of CMOS ICs. There are device dimensions, clocking strategy, architecture, clock distribution, etc. Here we focus our attention on clocking strategy and clock distribution problems.

To meet the demand for high-speed operation today, many systems adopt a double data rate (DDR) technologies, such as DDR SDRAM, double sampling ADC, clock and data recovery circuits, microprocessor circuits, etc. In these systems, both rising and falling edges of the clock are used to sample the input data, requiring that the duty-cycle of the clock be precisely maintained at 50%. Therefore, how to generate a clock with precise 50% duty-cycle for high-speed operation is an important issue. Namely, in high-frequency operations, clock outputs with a short cycle time can be severely distorted as clock passes through many delay cells. Even if the duty cycle of CLK_{ref} is 50% at the entrance that of CLK_{out} may deviate significantly from 50%. As a consequence, it can cause the output to have phase error, which could be fatal, especially in high-speed communication applications.

A conventional solution is to attach duty-cycle correction circuits to all output drivers with the price of added area, increased jitter, and further phase mismatch due to enlarged path [4]. There are several different methods for implementing 50% duty-cycle correctors, both in PLL and DLL control loops, intended to adjust the output duty-cycle of the multistage driver [6-9]. Each of these methods has its advantages and drawbacks. In all these circuits, the variable delay element is one of the key building blocks. Its precision directly affects the overall performance of the circuit.

III. DLL WITH DOUBLE EDGE SYNCHRONIZATION

The structure of the proposed Delay Locked Loop with Double Edge Synchronization (DLL–DES) clock alignment circuit is pictured in Fig. 2. The click aligner is composed of a voltage controlled delay line, VCDL, two phase detectors, PD1 and PD2, two charge-pumps, CP1 and CP2, two first order low-pass filters, LP1 and LP2, and a multistage clock buffer, CB. The negative feedback in the loop adjusts the delay through the VCDL by integrating the phase shift errors that result between the periodic reference input, CLK_{in}, and the multistage output, CLK_{out}.

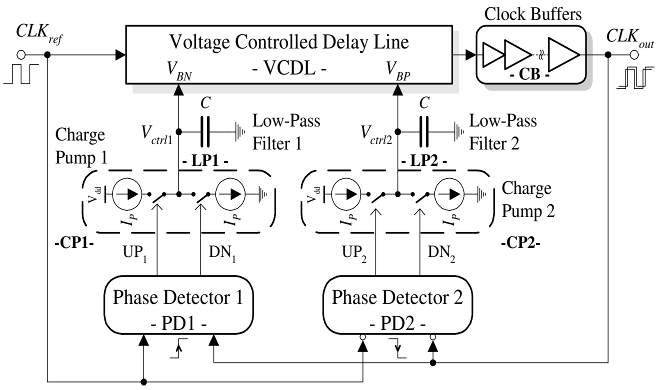


Fig. 2. DLL's architecture with double edges synchronization

The underlying idea for this approach is to provide delay regulation for both a rising and trailing edge of the output clock pulse CLK_{out}. For implementation of variable delay regulation the building block VCDL is used. The control voltage V_{BN} (V_{BP}) defines delay regulation of a rising (trailing) clock pulse edge. The phase detector PD1 (PD2) compares phase shifts of rising (trailing) edges between the input, CLK_{in}, and output, CLK_{out}, clock pulses. UP1 (UP2) pulses cause I_p to add charge to loop filter capacitor C, whereas DN1 (DN2) pulses remove charge. The LP1's (LP2's) output, V_{ctrl1} (V_{ctrl2}), is connected to the VCDL control input at node V_{BN} (V_{BP}). When the system, from Fig. 2, enters in stable state both edges of CLK_{out} are synchronized and phase shifted in respect to the referent clock CLK_{in}. An important feature of this architecture is that the duty-cycle of CLK_{out} is maintained at value of 50 %.

IV. CIRCUITS IMPLEMENTATION

In the sequel we will describe, in more details, the structure and principle of operation of each constituent of a DLL-DES based clock aligner.

A. Voltage controlled delay line

The actual implementation of a VCDL is comprised of a number of cascaded variable delay buffers. Each delay buffer (adjustable timing element) is of identical structure. Current starved delay element (CSDE) was chosen as a convenient candidate for realization of the delay buffer. The main design decision for such a choice was the following: CSDE provides independent delay regulation of both rising and falling clock pulse edges. Independent delay regulation can be achieved by varying the current of *p* and *n* MOS transistors.

In conventional CSDE (see Fig. 3(a)) a single control voltage V_{ctrl}, generated by a bias circuit, modulates the on resistance of pull-down M_3, and through a current mirror, pull-up M_4 [5]. The variable resistances control the current available to charge or discharge the parasitic load capacitance.

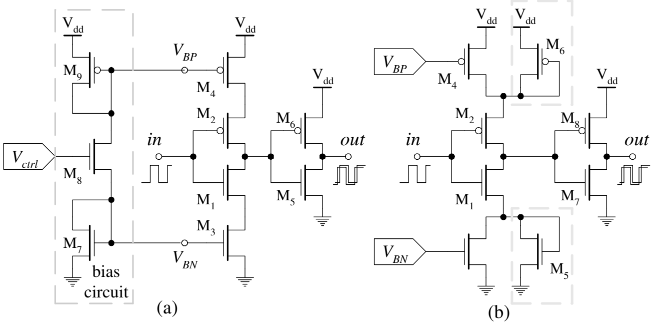


Fig. 3. Modified current starved delay element

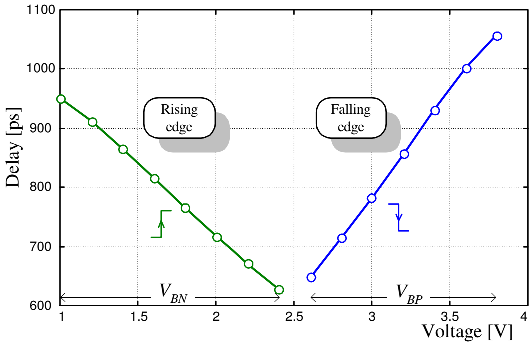


Fig. 4. Rising and Falling edge delay in term of control voltages

In order to achieve independent, instead of single, variable resistances control, we propose here a modified version of CSDE, as one given in Fig. 3(b). In our approach, both control voltages, V_{BN} and V_{BP}, directly drive gates of M_3 and M_4 MOS transistors, respectively. Transistors M_5 and M_6 act as symmetric loads and are used for two purposes: a) to linearize a voltage-to-delay transfer function of the CSDE; and b) provides correct initial condition for DLL operation even in a case when both control voltages V_{BN} and V_{BP} are out-of-regulation limits (for example, M_4 and M_3 are switched off). The modified CSDE was designed for 1.2μm CMOS techno-logy, for 5V power supply voltage. A SPICE simulation results that correspond to delay functions of both rising and falling pulse edges are given in Fig. 4. The obtained results in Fig. 4 show that linear regulation of voltage-versus-delay can be achieved. In general, CSDE offers good delay line stability in respect to temperature and supply voltage variations. Its main disadvantage is relatively limited range of delay regulation, i.e. low-sensitivity.

B. Phase detector

The phase detector measures the phase difference between the time reference and the delay chain. High precision dynamic phase detection circuit based on true single phase logic [3] is adopted in our design. The main advantages of this circuit are simple hardware structure, high-speed of operation, and small dead zone [5]. The UP*x* and DN*x* (*x* refers to 1 or 2) are used to control the charge-pump circuit CP*x*. The PD1 (PD2) is sensitive to rising (falling) clock pulse edge. A modification, in respect to standard solution [5], is performed by substituting MOS transistors P_{12}, N_{12}, P_{22}, and N_{22} (see Fig. 5 (a)) with complementary ones N_{11}, P_{13}, N_{21}, and P_{23} (see Fig. 5(b), respectively.

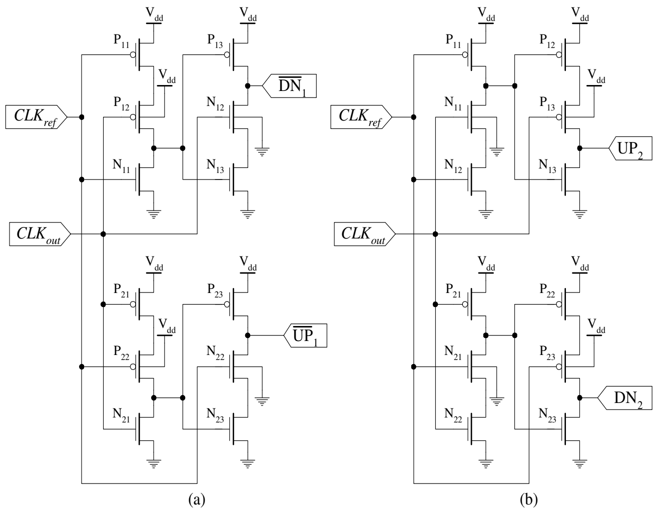


Fig. 5. Implementation of phase detectors for (a) raising and (b) falling edges

Operational principles of PD1 and PD2 are shown in Fig. 6. The widths of UP and DN signals are proportional to the phase difference of the input signals. Fig. 6a (6b) shows the operation of PD1 (PD2). Waveforms on the left side of Fig. 6a (6b) correspond to a case when the signal CLK_{out} (see Fig. 2) leads in respect to the signal CLK_{ref}. Otherwise, timing diagrams on the right side are valid (CLK_{ref} leads the CLK_{out} signal).

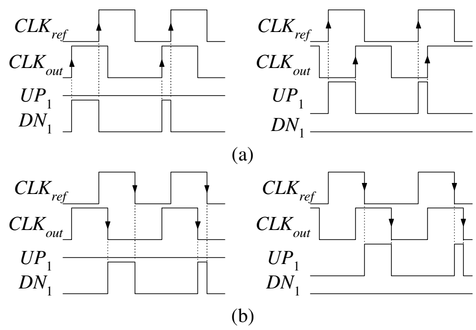


Fig. 6. Waveforms of input and output signals for (a) phase detector 1 and (b) phase detector 2

C. Charge pump and loop filter

The charge-pump and loop filter structure is presented in Fig. 7. Transistors P_1 and N_1 act as switching elements driven by pulses UP and DN, while transistors P_2 and N_2 are employed as current sink and source, respectively. The charge-pump charges or discharges the filter capacitor, C. The voltage on this capacitor, V_{ctrl} (V_{BP} or V_{BN} in Fig. 2), sets the VCDL stage propagation delay. The charge-pump the realization of an integrator transfer function with no additional active amplifier, resulting in a zero-phase error in steady state. A small capacitor, *C*, is used for the low-pass loop filter. The current level of the charge-pump and the charge delivered/accepted at every rising/falling clock edge transition are set to a small value [5]. This allows the implementation of the loop capacitor on chip.

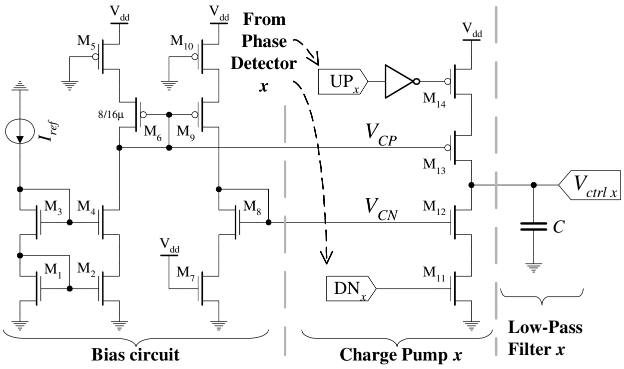


Fig. 7. Current pump and loop filter

The bias circuit provides correct operation of the charge pump. Its structure is given on the left side of Fig. 7. This circuit generates two control voltages, V_{CP} and V_{CN}. These voltages define the charge and discharge currents of loop capacitor, C, that pass through transistors M_{12} and M_{13}.

V. SIMULATION RESULTS

The DES-DLL sketched in Fig. 2 is implemented in 1.2 µm CMOS technology. It is supplied with V_{dd}=5V. SPICE simulation results that relate to referent clock excitation f_{ref}=80MHz are given in Fig. 8. The DES-DLL is operatives within the frequency range from 55 MHz up to 166 MHz.

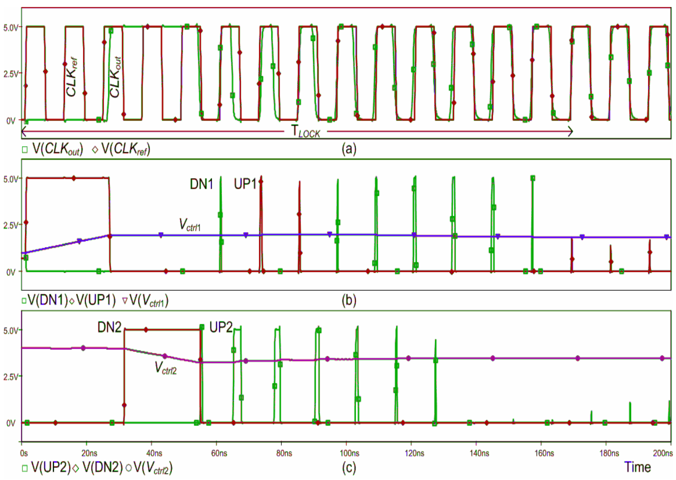


Fig. 8. Simulation of DLL with Double Edges Synchronization

Timing diagrams that correspond to referent input clock pulses, CLK_{ref}, and buffer output, CLK_{out}, are given in Fig. 8(a). This Fig. shows that the locking time, T_{LOCK}, between the referent CLK_{ref} and output CLK_{out} pulses is less than 200ns. We define T_{LOCK} as a time interval starting from initial condition up to the instant when total coincidence of rising and falling edges between both pulses, CLK_{ref} and CLK_{out}, exists, (see Fig. 8(a)). If we assume that CLK_{ref} is symmetrical then the coincidence corresponds to 50% duty-cycle of CLK_{out}. According to numerous simulations, within the operating range, duty cycle deviations are less then 1%. Compared to the results presented in [6-9] we obtain identical or better accuracy in duty cycle regulation.

Fig. 8(b) (8(c)) deals with waveforms that are obtained at the outputs UP1 (UP2), DN1 (DN2), and V_{ctrl1} (V_{ctrl2}). As can be seen from Fig. 8(b) (8(c)) UPx and DNx signals define the control voltage V_{ctrlx} during the transition period ($0 < t < T_{LOCK}$). After that the system enters in stable state and UPx and DNx signals disappear and $V_{ctrl\,x}$ takes a constant value.

According to the obtained results we can conclude that the proposed DES–DLL can be seen to have a wide-operational range and good duty–cycle correction capability.

VI. CONCLUSION

In this paper a new DLL architecture with clock alignment capability of both leading and trailing edges is described. The proposed circuit was simulated using models for 1.2 µm CMOS technology and SPICE simulator. The clock aligner has been designed specifically to correct precisely the duty-cycle factor in a multistage clock buffer to (50±1) % within the operating frequency range from 55 MHz up to 166 MHz. Timing diagrams are measured by simulation. The proposed DLL based clock aligner keeps the same benefits of conventional DLL's such as good absolute stability, fast-response, and low-level output jitter for both (rising and falling) edges. Such circuits serve in many applications including clock distribution network within the VLSI ICs, high-speed DRAM, and core-to-core interconnects within a system-on-chip designs.

REFERENCES

[1] Oklobdzija V., Stojanovic M., Markovic D., and Nedovic N., *Digital System Clocking: High-Performance and Low-Power Aspects*, John Wiley & Sons Pub., Hoboken, New Jersey, 2003.

[2] Razavi B., ed., *Phase-Locking in High-Performance Systems: From Devices to Architectures*, John Wiley & Sons Pub., Hoboken, New Jersey, 2003.

[3] Yuan J., and Svensson C., *High-Speed CMOS Circuit Technique*, IEEE JSSC, vol.24, No.1, pp.62-70, February 1989.

[4] Garleep B., et al., *A Portable Digital DLL for High-Speed CMOS Clock Buffers*, IEEE JSSC, Vol. 34, No. 5, pp. 632-643, May 1999.

[5] Moon Y., et al., *An All-Analog Multiphase DLL Using a Replica Delay Line for Wide-Range Operation and Low-Jitter Performance*, IEEE JSSC, Vol. 35, No. 3, pp. 377-384, March 2000.

[6] Fenghao M., Svensson C., *Pulse Control Loop in High-Speed CMOS Clock Buffers*, IEEE JSSC, Vol. 35, No. 2, pp. 134-141, February, 2002.

[7] Han S.-R., Liu S.-I., *A 500MHz- 1.25GHz Fast-Locking Pulsewidth Control Loop with Presettable Duty-Cycle*, IEEE JSSC, Vol. 39, No. 3, pp. 463-468, March 2004.

[8] Yang P.-H., Wang J.-S., *Low Voltage Pulsewidth Control Loops for SoC Applications*, IEEE JSSC, Vol. 37, No. 10, pp. 1348-1351, October 2003.

[9] Liu W.- M., Huang H.- Y., *A Low-Jitter Mutual-Correlated Pulsewidth Control Loop Circuit*, IEEE JSSC, Vol. 39, No. 8, pp. 1366- 1369, August 2004.

[10] Kim C., Kang S., *A Low-Swing Clock Double-Edge triggered Flip-Flop*, IEEE JSSC, Vol. 37, No. 5, pp. 648-652, May, 2002.

Threshold Logic Circuits Implementation Using FPAA

Boyan P. Lyubenov[1], Emil D. Manolov[2]

Abstract – **The paper presents the results from the design and investigation of basic threshold logic circuits using Field Programmable Analog Array (FPAA) of Anadigm Inc. To this aim, the functional model of the threshold logic gate is discussed and some approaches in building logic functions are presented. The results from implementation of different FPAA variants of threshold logic structures are described. The use of the FPAA ensures possibilities for simple programming and dynamic reconfiguration of different values of the weights on the inputs as well as a flexible realization of different logic functions. The results clearly present the practical use of the discussed approach and could find application for fast prototyping of threshold logic based systems with possibilities for flexible real-time programming and reconfiguration of the parameters and functions.**

Keywords – **Threshold logic, Neural networks, Programmable analog circuits, Field Programmable Analog Array, FPAA**

I. Introduction

The threshold logic (TL) concept was introduced as theory of logic gates [1], [2], and over the years has promised much in terms of reduced logic depth and gate count, compared to conventional Boolean logic based design. Since the basic TL-gate is functionally more powerful than those of the conventional logic, many complex functions can be synthesized in TL with lesser number of gates in a shorter logic depth. Despite the theoretically obvious merits, TL has never had a significant impact in practice, most probably due to the lack of efficient physical realization.

In the last decade the fast development of VLSI technology has made neurocomputer design not only a research topic but several chips have been developed [3], [4], [5]. Research on hardware implementations of neurons has recently been very active. In [6] and [7], for fast prototyping of neural networks, is used Field Programmable Analog Array (FPAA).

The FPAA circuits have been introduced in 1990's by some of the biggest chips' suppliers. They are analog equivalents of the Field Programmable Gate Array (FPGA). The main reason behind FPAA was to help analogue designer to debug their systems long before real silicon comes out the fabs so that significant time-to-market reduction can be achieved. Nowadays, the FPAA technology is very flexible and powerful technology for fast prototyping of different electronic systems [8].

In the presented research is used FPAA chip introduced by Anadigm® [9]. The Anadigm® AN220E04 is a reconfigurable analog device based on switched capacitor technology.

[1]Boyan P. Lyubenov is Ph.D. student at Faculty of Electronic Engineering and Technologies, Technical University – Sofia, 1000 Sofia, Bulgaria

[2] Emil D. Manolov is with the Faculty of Electronic Engineering and Technologies, Technical University – Sofia, 1000 Sofia, Bulgaria, E-mail: edm@tu-sofia.bg

The paper demonstrates a practical approach for building, simulation, implementation and verification of different threshold logic circuits using FPAA.

II. Basic Threshold Logic Theory

A threshold gate is defined as an n-input logic gate, functionally similar to a hard-limiting neuron without learning capability [1]. The gate takes n binary inputs x_1, x_2 … x_n and provides a single binary output y as it is shown in Fig. 1.

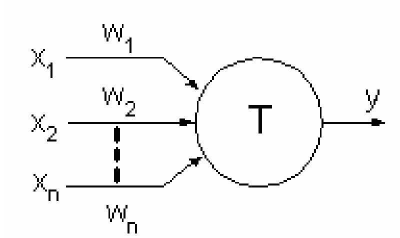


Fig. 1. Threshold logic gate

The output of the gate is determined by the following set of relations:

$$y = \begin{cases} 1, & if \sum_{i=1}^{n} w_i x_i \geq T \\ 0 & otherwise \end{cases}, \qquad (1)$$

where T is the threshold and the w_i is the weigh associated with the i-th input variable x_i. The function can be written in more compact form:

$$y = \operatorname{sgn}\left(\sum_{i=1}^{n} w_i x_i - T \right), \qquad (2)$$

where the sgn() function is defined as follows:

$$\operatorname{sgn}(x) = \begin{cases} 1, & if\ x \geq 0 \\ 0 & otherwise \end{cases} \qquad (3).$$

In order to increase the robustness of the threshold logic gate, fault tolerances can be added:

$$y = \begin{cases} 1, & if \sum_{i=1}^{n} w_i x \ \geq T + \Delta_1 \\ 0, & if \sum_{i=1}^{n} w_i x_i < T - \Delta_2 \end{cases}, \qquad (4)$$

where parameters Δ_1 and Δ_2 are tolerances that ensure stability with respect to technology and temperature variations which can violate functionality of the system.

A device which implements this theoretical model must compute the linear weighted sum of the binary inputs, store the threshold value and compare the weighted sum to this threshold. TL can be programmed to realize many distinct

Boolean functions by adjusting the threshold T and/or the weights w_i. For example, an n-input TL gate with $T=n$ will implement n-input AND gate, but only by setting $T=n/2$, the gate will compute a majority function. This versatility means that TL offers significantly increased computational capabilities over the conventional AND/OR/NOT logic. Moreover, the reduced area, increased speed and larger number of input variables are part of advantages of TL over conventional Boolean logic.

III. IMPLEMENTATION

Standard logic functions can be easily implemented using threshold logic gates instead of Boolean. A TL gate can directly replace every conventional logic gate. If consider 2-input threshold gate from Fig. 1 and set the weights equal to '1', we can control the property of that gate only by varying the threshold value.

For instance, to obtain logical operation "AND" the gate have to be trained for the input pairs (x_1 x_2) and the output respond y so that (0 0):0, (1 0):0, (0 1):0, (1 1):1. Analysis of Eq. (1) in case of $w_1=w_2=1$ show that if $1<T\leq 2$ the threshold gate is AND. Then if the both inputs are '1' the weighted sum is equal or greater than the threshold T.

The similar approach can be used to implement logical function "OR". In this case the inputs signals and output response are (0 0):0, (1 0):1, (0 1):1, (1 1):1. Analysis of Eq. (1) in case of $w_1=w_2=1$ show that if $0<T\leq 1$ the threshold gate is OR. The results clearly prove that only by changing the threshold value T, AND-gate can be transformed to OR-gate and vice versa. Graphical explanation of the possibility for threshold adjustment is shown in Fig. 2. Threshold value T changes the position of the line that separates the plane into two regions.

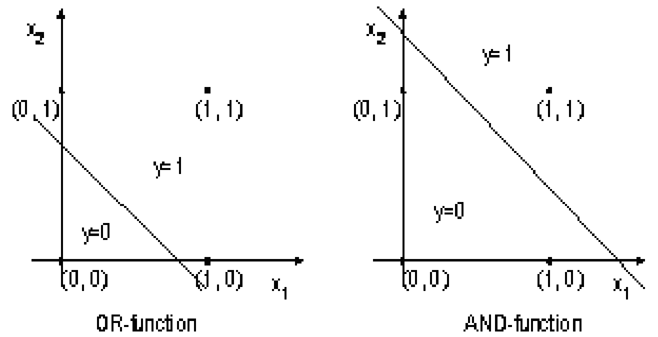


Fig. 1. AND/OR function representation

To implement and simulate AND/OR threshold gates Anadigm Desiger2 program is used. The practical verification is carried out using Anadigm Evaluation board. The circuit is shown on Fig. 2: the summation is done by SumDiff block (Σ) and the comparison - by Comparator. Input levels for x_1 and x_2 are 0V and 1V respectively and the threshold T changes from 0V to 2V. The circuit simulation and verification confirm the transformation of OR-gate to AND-gate and vice versa only by changing the threshold level T (Fig. 3).

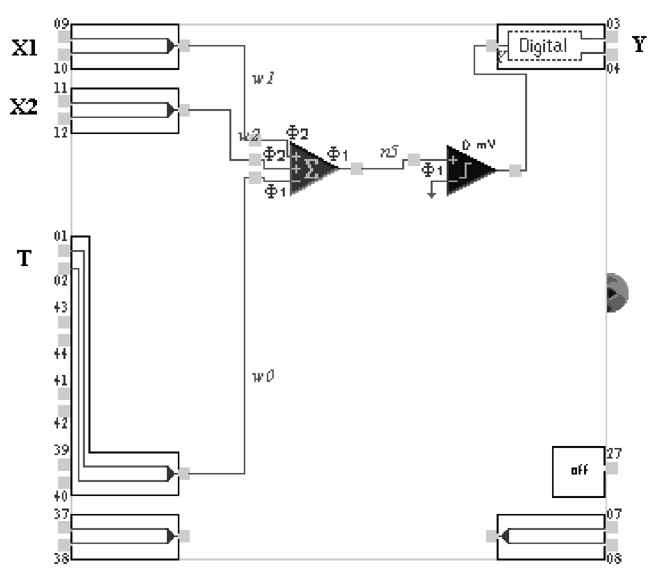


Fig. 2. FPAA implementation of AND/OR gate

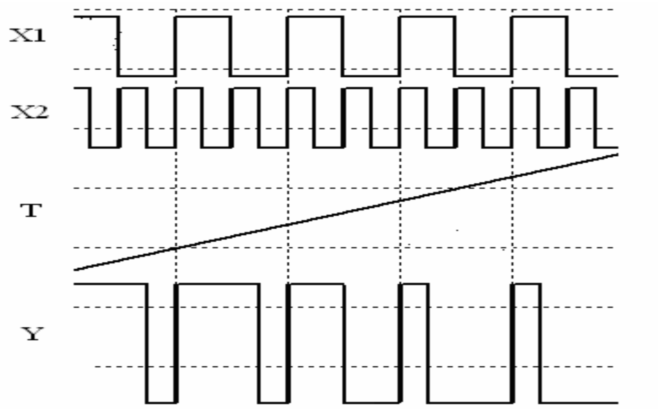


Fig. 3. Simulation of AND/OR function

Implementation of NAND/NOR gate also is easy task – now the basic gate output Eq. (1) is changed into Eq. (5):

$$y = \begin{cases} 0, & if \sum_{i=1}^{n} w_i x_i \geq T \\ 1 & otherwise \end{cases} . \tag{5}$$

In this case, the regions, where $y=1$ and $y=0$, are swapped in comparison with AND/OR function. FPAA implementation is identical to the shown on Fig. 2. Only the Comparator block is changed from non-inverting to inverting mode.

In the case of 3-input logic gate, if $2<T\leq 3$ - the threshold gate is AND, if $0<T\leq 1$ - the threshold gate is OR. Increasing the number of the inputs goes not affect the speed of the gate – this is a very important property that attracts designer's attention to threshold logic. The 3-input gate is presented on Fig. 4. The summing element is 4-input SumDiff (Σ), the input levels for x_1, x_2 and x_3 are 0V and 1V respectively, and the threshold T changes from 0V to 3V. The simulations are shown in Fig. 5. Threshold curve is not presented on the picture because of the fact that only 4 cursors are available in Anadigm Desiger2 software [9]. The gate changes its function according to the value of threshold T.

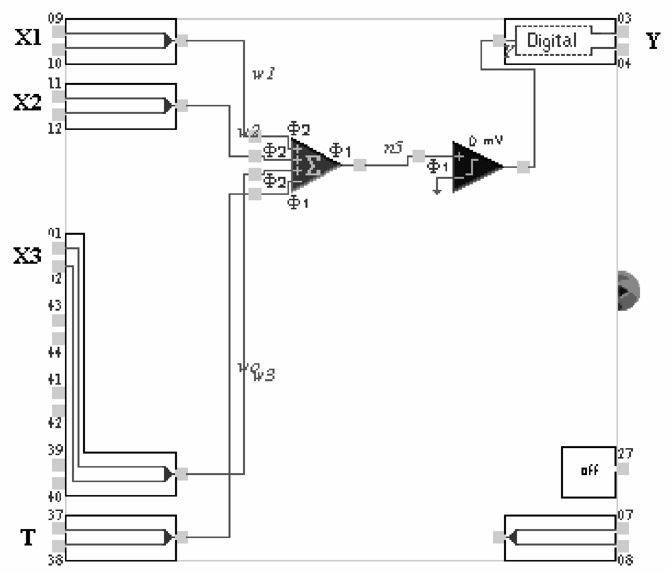


Fig. 4. FPAA implementation of 3-input AND/OR gate

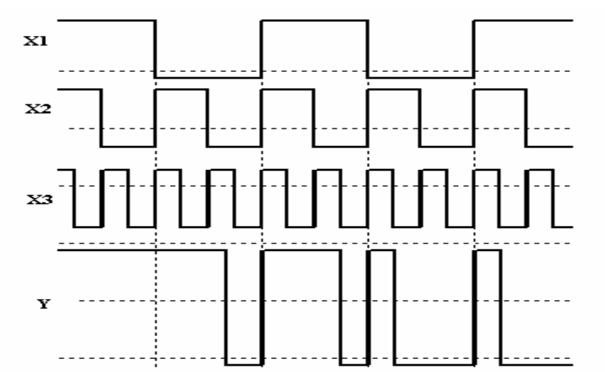


Fig. 5. Simulation of 3-input AND/OR gate

Let consider more complex network such as the logical operation "EXCLUSIVE-OR". The creation of XOR gate is a bit complex task. In this case two-level threshold logic is needed. The logic equation of XOR gate is $y = x_1\bar{x}_2 + \bar{x}_1 x_2$.

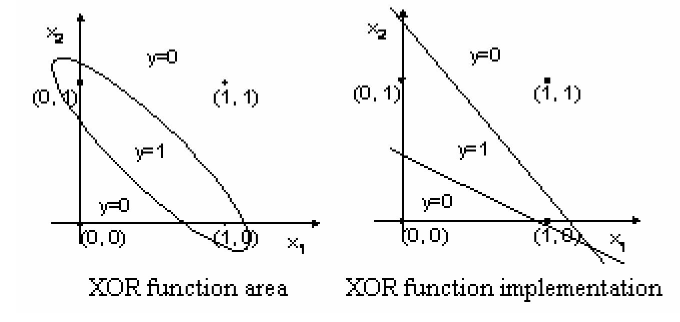


Fig. 6. XOR function representation and implementation

There are several approaches to implement such gate. Graphical explanation is shown in Fig. 6. The left picture shows the area of representation of XOR-function. The right picture present the implementation of the function using AND and OR logical components. The goal is to build a system that performs summation of the outputs of an AND-gate and OR - gate (Fig. 6). The function can be implemented by using two-layer network (Fig. 7). The presented network have to be trained to obtain the following states $(0\,0):0$, $(1\,0):1$, $(0\,1):1$, $(1\,1):0$. Setting $w_{11} = w_{12} = w_{21} = w_{22} = 1$, the system have to be solved for T_1, T_2, T_3 and w_{Y1}, w_{Y2} as well. $Y1$ and $Y2$ are the outputs of the hidden layer. One possible solution is to set $0<T_2\leq 1$, $0<T_3\leq 1$ and $1<T_1\leq 2$ and $w_{Y1}= -1$, $w_{Y2}=1$.

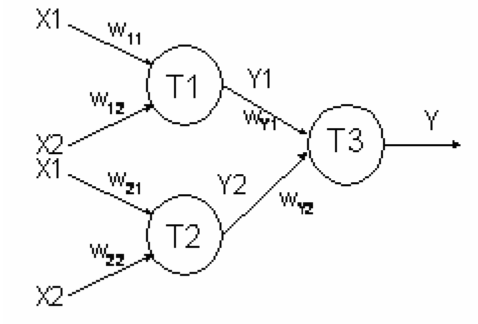


Fig. 7. 2-input 2-layer network

Another possible way to implement a XOR-gate is shown on Fig. 8. In this case: $1<T_1\leq 2$, $0<T_2\leq 1$, $w_{11}=w_{12}= w_{21}=w_{22}=1$, and $w_{Y1}= -2$.

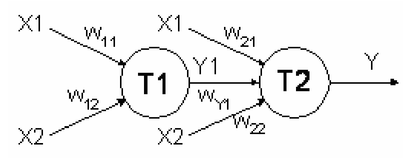


Fig. 8. Another 2-input 2-layer network

The circuit shown in Fig. 8 is more compact and was implemented using FPAA (Fig. 9).

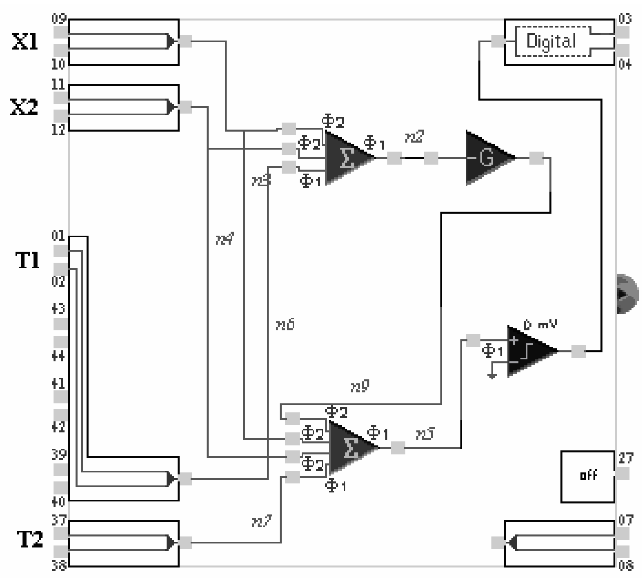


Fig. 9. FPAA implementation of 2-input XOR gate

The gate T1 is represented by 3-input SumDiff (Σ) and the gate T2 by 4-input SumDiff (Σ). GainInv (-G) cell represents the weight w_{Y1}=-2 as was derived above. The thresholds of the gates are set T_1=1.5V and T_2=0.5V. At these conditions the simulations (Fig. 10) clearly show that the presented FPAA implementation is XOR-gate.

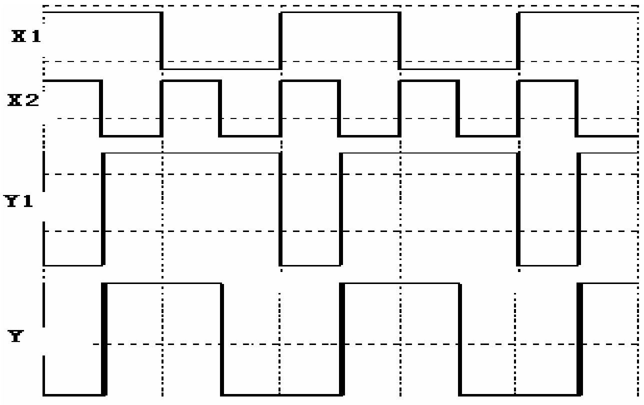


Fig. 10. Simulation of 2-input XOR function

Another experiment is to build more complex functions. There are two basic approaches: either to replace every logic function with the appropriate threshold logic gate or to train a special gate to satisfy the wanted function. Two examples are presented in order to demonstrate these basic options.

Let the wanted function is $y = x_1 x_2 + x_3$. Using the first method every operation have to be implemented by a single TL-gate: firstly $y_1 = x_1 x_2$ and then $y = y_1 + x_3$. That means two gates and two threshold values. The way to build 'AND' and 'OR' gates was presented above.

The second approach is to use 3-input TL gate. For the input vectors (x_1 x_2 x_3) the following response y is needed: (0 0 0):0, (0 0 1):1, (0 1 0):0, (0 1 1):1, (1 0 0):0, (1 0 1):1, (1 1 0):1, (1 1 1):1. Analysis of Eq. (1) in the case of w_1=w_2, gives the following:

$$T>0, \qquad 0.5T< w_1=w_2<T, \qquad w_3>T \qquad (6)$$

Every set of values that satisfy the system (6) represents the given logic function. This approach reduces the final gate number and the complexity of the entire system.

FPAA implementation of the discussed function is very easy. As summing element is used 4-input SumDiff (Σ) and the weights are set as coefficients within that block (Fig. 4). One possible solution of (6) to achieve $y = x_1 x_2 + x_3$ is T=1V, w_1=w_2=0.6, w_3=1.2, input levels for x_1, x_2 and x_3 - 0V and 1V respectively for logic '0' and logic '1'. The simulation results are given on Fig. 11.

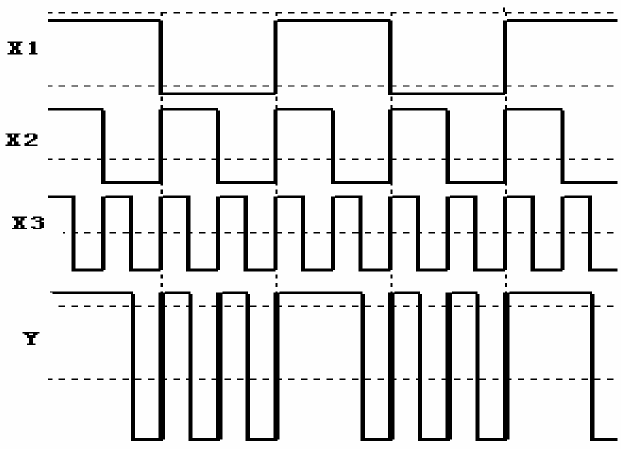


Fig. 11. Simulation of y=x_1x_2+x_3 function

Another logical function is $y = (x_1 + x_2)x_3$. Now the input vectors and the output response are (0 0 0):0, (0 0 1):0, (0 1 0):0, (0 1 1):1, (1 0 0):0, (1 0 1):1, (1 1 0):0, (1 1 1):1. The analysis of Eq. (1), for w_1=w_2, gives:

$$T>0, \qquad 0< w_1=w_2<0.5T, \qquad T>w_3>T-w_1 \qquad (7)$$

The function was implemented using FPAA in the same way as the previous. The conditions are: T=1V, w_1=w_2=0.4, w_3=0.8, input levels for x_1, x_2 and x_3 - 0V and 1V respectively for logic '0' and logic '1'. The results from simulation confirm the correctness of the presented deductions.

More complex combinatory functions such multiplexers, demultiplexers and many other can be implemented using the approach and basic threshold logic gates proposed above.

IV. CONCLUSIONS

The paper presents an approach for prototyping and examination of basic threshold logic functions using FPAA. To this aim, the functional model of the threshold logic gate is discussed and some approaches in building logic functions are presented. The results from implementation of different FPAA variants of threshold logic structures are described. The use of the FPAA ensures possibilities for simple programming and dynamic reconfiguration of different values of the weights on the inputs as well as a flexible realization of different logic functions. The results clearly present the practical use of the discussed approach and could find application for fast prototyping of threshold logic based systems with possibilities for flexible real-time programming and reconfiguration of the parameters and functions.

REFERENCES

[1] McCulloch, W.S., Pitts, W., "A logical calculus of the ideas implement in nervous activity," Bull. Math. Biophysiol., Vol. 5, 1943, pp. 15-33.

[2] Rosenblatt, F., "The perceptron a probabilistic model for information storage and organization," Brain Psych. Rev., Vol. 62, 1958, pp. 368-408.

[3] Huertas, J.L., Sanchez-Solano, S., Baturone, I., Barriga, A., "Integrated circuit implementation of fuzzy controllers" IEEE JSSC, Vol. 31, NO. 7, July 1996, pp. 1051-1058.

[4] Espejo, S., Domingues-Castro, R., Rodrigues-Vazques, A., "A 16x16 cellular neural network chip for connected component detection", 1993

[5] Leong, P.H.W., Jabri, M.A., "A VLSI neural network for morphology classification", International Joint Conference on Neural Networks, IJCNN, 1992, Volume: 2, pp. 678-683.

[6] Berenson, D., N. Estevez, H. Lipson, "Hardware Evolution of Analog Circuits for In-situ Robotic Fault-Recovery," 2005 NASA/DoD Conference on Evolvable Hardware (EH'05), 2005 , pp. 12-19.

[7] Manolov E.D., B.P. Lyubenov, Design and investigation of two-parameter space classification circuits using FPAA. Proceedings of the 14th International Scientific and Applied Science Conference Electronics ET'2005, book 5, pp.99-104.

[8] Manolov, E.D., "Research and Educational Experiments with FPAA", 12th International Conference Mixed Design of Integrated Circuits and Systems MIXDES'2005. Krakow, Poland, vol.2. pp. 975-980.

[9] Anadigm Inc. Technical Documentation. www.anadigm.com

Controlled by Optoelements Oscillators

Tsanko V. Karadzhov[1]

Abstract – **The controlled oscillators have usually electrical control input. The disadvantage of this oscillator type is that there is no galvanic separation between the control signal and the oscillator as well the control can be fulfilled only by the means of electrical stimulus. Thus electrical circuits are designed as they have both electrical and optical control input as there is galvanic separation between the oscillator and control stimuli in the same time. The circuits of controlled by optoelectronic elements oscillators for impulse and sine wave signal are developed. The frequency of these oscillators shifts in wide range as the illumination changes. In the current article two oscillators fulfilled with photo FETs and operational amplifiers are investigated. Photo FETs are connected in time setting circuit of the oscillators which allows to control them optically**

Keywords – **Oscillator, Optoelement, photo FETs.**

I. Controlled by Optoelements Impulse Oscillator Fulfilled with Photo JFETs and Operational Amplifier (OPAMP)

A. Work principle of the circuit – fig.1

The circuit represents symmetrical multivibrator implemented on OPAMP which is used as Schmidt trigger. The Schmidt trigger thresholds are determined as the resistance ratio between the drain and the source of the two photo JFETs. The computer modeling results are shown on fig. 2 and the real experimental results are shown on fig. 3

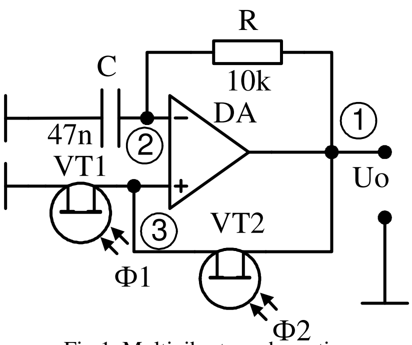


Fig.1. Multivibrator schematic

B. Circuit analysis

The Schmidt trigger threshold voltages are defined accordingly to following equations:

$$U_I = +\frac{R_{VT1}(\Phi 1)}{R_{VT1}(\Phi 1) + R_{VT2}(\Phi 2)} Uo = \beta . U_O \qquad (1)$$

$$U_{II} = -\frac{R_{VT1}(\Phi 1)}{R_{VT1}(\Phi 1) + R_{VT2}(\Phi 2)} Uo = \beta . U_O \qquad (2)$$

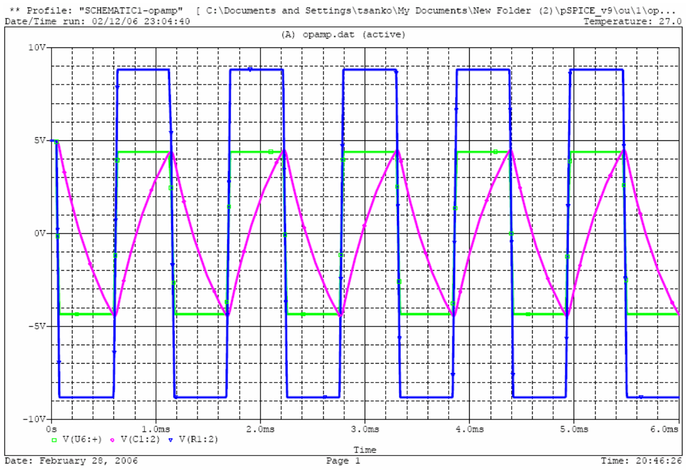


Fig. 2. Simulation results for impulse oscillator with photo FETs

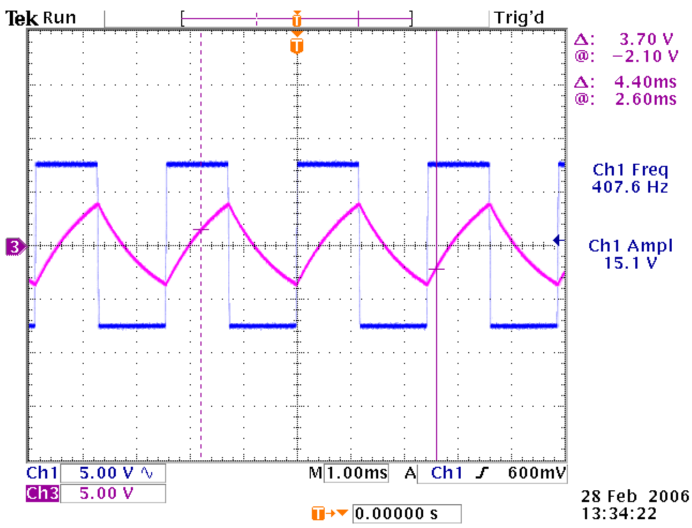


Fig. 3. Experimental results for impulse oscillator with photo FETs

The recharge cycle of capacitance C is described by the formula:

$$u_{C(t)} = U_{CC}(1 - e^{\frac{t}{\tau}}) \text{, where}$$

$$U_{CC} = (1+\beta)U_O \text{ end } \tau = RC \qquad (3)$$

The capacitance voltage changes its values in the range from $-\beta U_O$ to $+\beta U_O$ – the variation interval is $2\beta U_O$.
Then for $2\beta U_O$ the following formula is obtained:

$$2\beta U_O = (1+\beta)U_O\left(1 - e^{-\frac{T}{2RC}}\right) \qquad (4)$$

[1]Tsanko V. Karadzhov. Department of Electronics, Technical University – Gabrovo, Street "H. Dimiter" No. 4, 5300 Gabrovo, Bulgaria, phone: +359 66 801064, e-mail: karadjov_st@abv.bg

And therefore the period T can be presented as follows:

$$T = 2RC\ln\frac{1+\beta}{1-\beta} \tag{5}$$

$$T = 2RC\ln\left(1+2\frac{R_{VT1}(\Phi1)}{R_{VT2}(\Phi2)}\right) \tag{6}$$

When both of the photo JFETs are equally illuminated, i.e. $\Phi1 = \Phi2 = \Phi$, then $R_{VT1}(\Phi1) = R_{VT2}(\Phi2) = R_{VT}(\Phi)$.

In this case the repeat period of the generated impulses will be as follows:

$$T = 2RC\ln\left(1+2\frac{R_{VT}(\Phi)}{R_{VT}(\Phi)}\right) = 2RC\ln 3 = 2{,}2RC = \tag{7}$$
$$2{,}2.10.10^{3}.47.10^{-9} = 1{,}034.10^{-3} \approx 1ms$$

$$f = \frac{1}{T} = \frac{1}{2{,}2RC} = \frac{1}{1{,}034.10^{-3}} \approx 1kHz \tag{8}$$

The output frequency depends on the correlation between the degrees of illumination each of the photo FETs is exposed to. The experimental research determined that the threshold voltage varies in very wide range.

When the resistance ratio is $\frac{R_{VT1}}{R_{VT2}} = 100$, then the period is $T_{max} = 4{,}98$ ms.

When the resistance ratio is $\frac{R_{VT1}}{R_{VT2}} = 1$, i.e. the photo FETs are equally exposed to the light, then the period is $T_{med} = 1$ ms.

When the resistance ratio is $\frac{R_{VT1}}{R_{VT2}} = 0{,}1$, then the period is $T_{min} = 171$ µs.

The rate of changeable period is:

$$\frac{T_{max}}{T_{min}} = \frac{4{,}98.10^{-3}}{171.10^{-6}} = 29 \tag{9}$$

II. Controlled by Optoelements Sine Wave Oscillator

A. Work principle of the circuitt

The circuit of sine wave generator is shown on fig. 4. It represents RC oscillator with utmost transference coefficient and zero lag in the feedback circuit, which is realized with the Wien bridge circuit. In the Wien bridge a photo JFET is added to each of the serial RC group and parallel RC group.

With the classical Wien bridge the output voltage is brought back to the input at maximum efficiency for single certain frequency f_0 without the change in phase φ_β. The Wien bridge is connected to the non-inverse input of the OPAMP. The advantage of the circuit shown on fig. 4 is in the fact that generated frequency changes proportionally to the applied illumination as the only requirement for stable work of generator is both photo FETs to be approximately equally illuminated. The results from computer modeling and laboratory research are respectively shown on fig. 5 and fig. 6.

$\Phi1 = \Phi2$ and therefore $R_{VT1} = R_{VT2}$

This requirement is more easily implemented when in the circuit is used differential photoreceiver which consists of two photo FETs packed together.

The stable work of generator is additionally assured as the negative feedback represented by the resistors R_3 and R_4 is connected in the circuit.

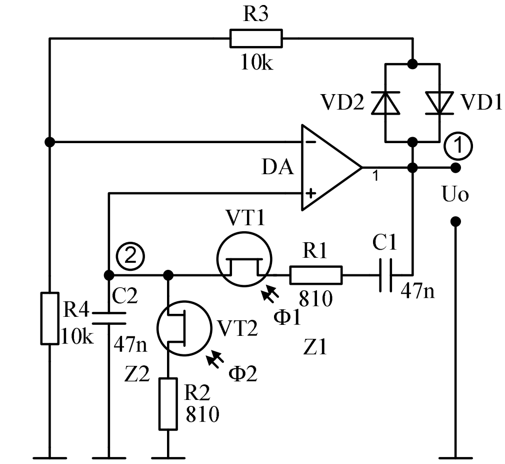


Fig. 4. Sine wave oscillator with photo FETs

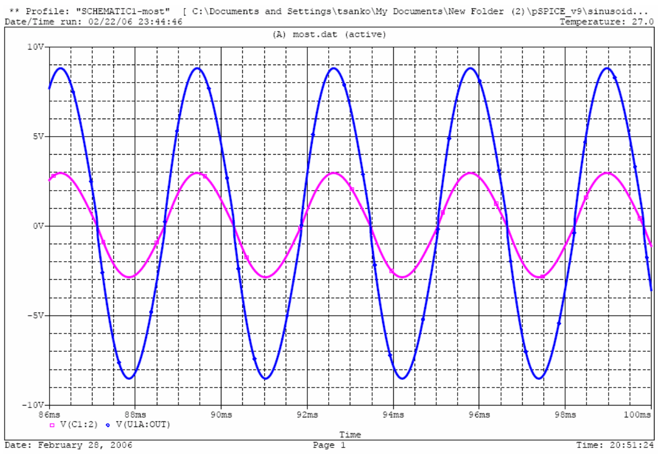


Fig. 5. Simulation results for sine wave oscillator with photo FETs

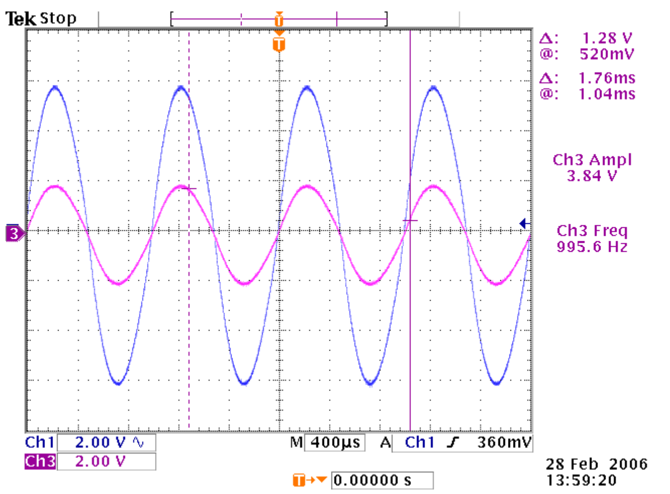


Fig. 6. Experimental results for sine wave oscillator with photo FETs

B. Circuit analysis

As the sinusoidal generator has two feedbacks in its circuit, the voltage U_β is equal to the difference between feedback voltages.

$$U_\beta = U_{\beta+} - U_{\beta-} \tag{10}$$

If assumption is made that the circuit works with ideal voltage amplifier ($R_I = \infty$ and $R_O = 0$), then the bridge will be balanced if the following requirement is fulfilled:

Z1.R4 = Z2.R3, where

$$Z1 = R1 + R_{VT1}(\Phi 1) + \frac{1}{j\omega C1},$$

$$Z2 = \frac{R2 + R_{VT2}(\Phi 2)}{1 + j\omega C2\left[R2 + R_{VT2}(\Phi 2)\right]} \tag{11}$$

Therefore the feedback coefficient β will be:

$$\beta = \frac{U_\beta}{U_O} = \frac{U_{\beta+}}{U_O} - \frac{U_{\beta-}}{U_O} = \frac{Z2}{Z1+Z2} - \frac{R4}{R3+R4} \tag{12}$$

After substitution the following equation is obtained:

$$\beta = \frac{1}{1 + \frac{R1+R_{VT1}(\Phi 1)}{R2+R_{VT2}(\Phi 2)} + \frac{C2}{C1} + j\left\{\omega C2\left[R1 + R_{VT1}(\Phi 1)\right] - \frac{1}{\omega C1\left[R2+R_{VT2}(\Phi 2)\right]}\right\}} - \frac{R4}{R3+R4} \tag{13}$$

The minimal value of OPAMP amplify coefficient which guarantees the stable generations A_{min} is:

$$A_{\min} = \frac{1}{\beta} = \frac{1}{\frac{1}{1 + \frac{R1+R_{VT1}(\Phi 1)}{R2+R_{VT2}(\Phi 2)} + \frac{C2}{C1}} - \frac{R4}{R3+R4}} \tag{14}$$

The lag between output signal U_O and input signal U_I is zero.

The frequency of the sinusoidal generator is:

$$f = \frac{1}{2\pi\sqrt{\left[R1 + R_{VT1}(\Phi 1)\right].\left[R2 + R_{VT2}(\Phi 2)\right].C1C2}} \tag{15}$$

Since in the presented case $R_1 = R_2 = R$, $C_1 = C_2 = C$ and $R_{VT1}(\Phi 1) = R_{VT2}(\Phi 2) = R_{VT}(\Phi)$, thus the frequency will be as follows:

$$f = \frac{1}{2\pi\left[R + R_{VT}(\Phi)\right].C} \tag{16}$$

$$f = \frac{1}{2\pi.\left[810 + 10.10^3\right].47.10^{-9}} \approx 157 Hz$$

The Wien bridge oscillator works more steadily when the negative feedback circuit has nonlinear elements in it. Thus the diodes VD_1 and VD_2 are connected in parallel with the resistor R_3.

III. Conclusion

The developed oscillators which are controlled by opto-elements can be applied as illumination-frequency converters, for illumination measurement or to measure difference between illuminations. Also proposed oscillators can be used in the frequency adjustment circuits which are controlled optically.

References

[1] Philips. Optoelectronic devices, 1986.
[2] Texas Instruments, Intelligent Opto Sensors, 1999.
[3] Toshiba. Optoelectronics, 1999.
[4] Zetex. Discrete Semiconductors. Quickreference Guide, 93.

Optocouplers and Optoelectric Elements Controlled by Sensors

Elena N. Petkova[1], Ivan S. Kolev[2]

Abstract – **It is well known that the control over optocouplers can be achieved in two ways – electronically and optically. In the current article a sensor control of both input and output optocoupler.**

Keywords – **Sensor, Optocoupler, Control, Optoelectronic elements**

I. INTRODUCTION

In the known from the practice sensor buttons [5] there is no galvanic separation between executive circuit and sensor plates. There is a risk of electrical damage for the operator when the high voltage circuits are controlled. This disadvantage is avoided through the proposed sensor buttons with galvanic separation, which is fulfilled with optocouplers. A classification of the sensor optocoupler buttons

According to the work principle:

– by means of the skin resistance (resistance which is introduced in the circuit by touch);

– introduced capacitance in the circuit (can be remotely applied);

– antenna effect – an introduced voltage which is generated by the human body

According to the number of sensor plates:

– with one sensor plate (touch button)

– with two sensor plates (contact button)

According to the used elements:

Circuits with discrete elements

– bipolar transistors

– JFETs

– MOSFETs

Circuits with ICs

According the place where sensor element is connected:

– in the light source circuit

– in the photo-element circuit

– in both circuits

The sensor buttons can be produced in two variants – with memory and without it.

The resistance between skin and electrode has inconstant characteristics and it depends on the contact area, distance between contact electrodes, skin conditions, human conditions, type of used voltage – DC or AC, frequency, etc.

II. INPUT CIRCUIT SENSOR CONTROL OF OPTOCOUPLERS

The circuit with p – channel JFET is presented on figure 1. The circuit work principle is based on the introduced resistance of human skin between the two sensor areas S.

[1]Elena N. Petkova, Department of Electronics, Technical University of Gabrovo, street "H. Dimiter" N 4, 5300 Gabrovo, Bulgaria, fone: +35966223433, e-mail: elena_ndpt@yahoo.com

[2]Ivan S. Kolev, Department of Electronics, Technical University of Gabrovo, street "H. Dimiter" N 4, 5300 Gabrovo, Bulgaria, fone: +35966223433, e-mail: ipk_kolev@yahoo.com

The JFET VT_1 is off as well as the phototransistor VT_2 of the optocoupler O_1 when human finger is not touching the sensor plates. When touch is applied to the sensor, VT_1 goes on and there is current flow through LED.

In result VT_2 goes also on and the output voltage changes from high logical level to low. The threshold voltage of JFET (КП103Л) VT_1 is in the range from 2 to 6 volts ($U_{GO} = (2 \div 6)$ V.

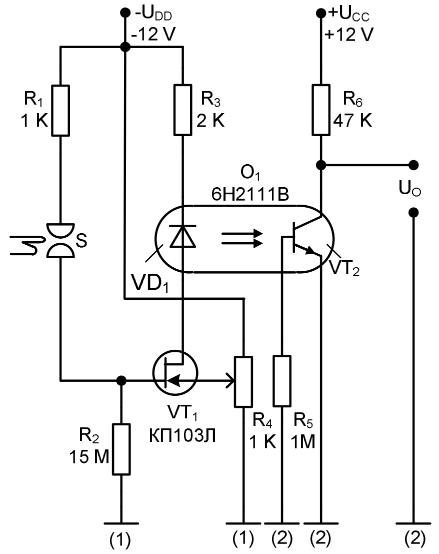


Fig. 1. Sensor button with JFET transistor

The sensor button shown on figure 2 works in the similar way, only in the circuit transistor VT_1 is n – channel MOS transistor.

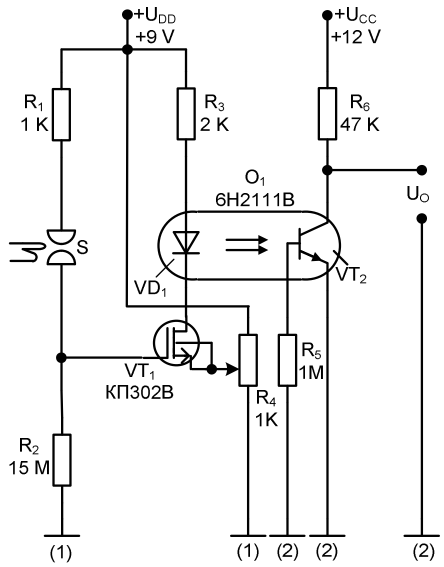


Fig. 2. Sensor button with MOS transistor

III. OUTPUT CIRCUIT SENSOR CONTROL OF OPTOCOUPLERS

In the circuit shown on figure 3, when the current I_F is zero and the sensor S is not touched, the transistors VT_1, VT_2 and VT_3 are off and the value of output voltage is little smaller than the supply voltage U_{CC}. When the sensor is touched, all of transistors in the circuit are on and the output voltage value is about 0,7 V ($U_O \approx 0{,}7$ V).

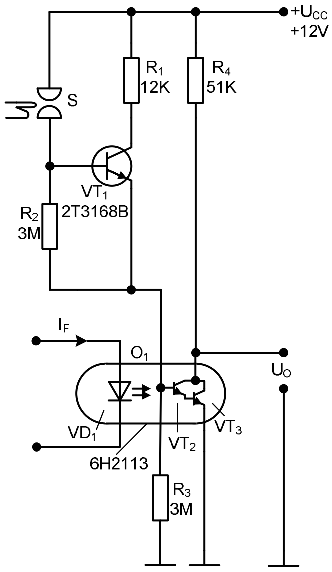


Fig. 3. Sensor button with bipolar transistor

There is a necessity of Schmidt trigger connected to the sensor button output which will provide steadily turn out of the sensor button – figure 4.

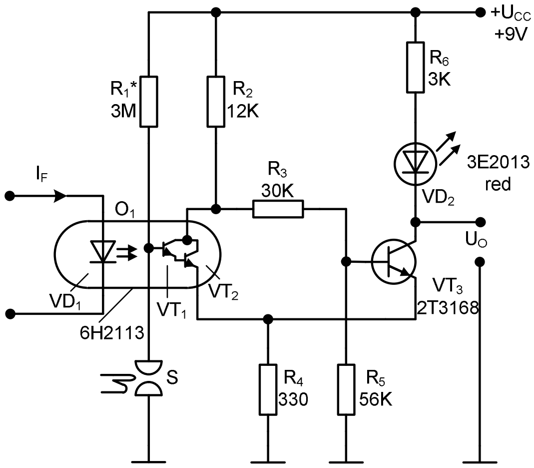


Fig. 4. Sensor button with trigger Schmidt

There is a necessity of Schmidt trigger connected to the sensor button output which will provide steadily turn out of the sensor button – figure 4. Such trigger is fulfilled with Darlington phototransistor VT_1 and transistors VT_2 and VT_3. Darlington phototransistor is on in the initial state when I_F is zero and the sensor is not touched, which results in off state for the outlet transistor VT_3 and high level for output voltage. When touch is applied to the sensor, the output voltage shifts to low logical level ($U_O \approx 0{,}7$ V). The LED VD_1 emits in the red part of the spectrum.

The proposed above sensor buttons have two contact plates.

E. Sensor button with one contact plate – on figure 5.

The work principle of proposed sensor uses inducted by human body voltage (antenna effect). All of the transistors in the circuit are off and output voltage is in high level state when the current I_F is zero and the sensor S is not touched. When touch is applied the current I_F remains zero but output voltage shifts to low state $U_O \approx 0{,}15$ V and LED VD_1 emits in the spectrum red portion.

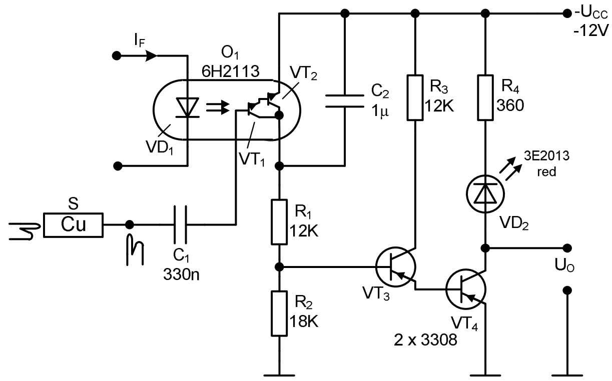


Fig. 5. Sensor button with one contact plate

When sensor S is not touched circuit is switched by the current I_F = 10 mA which flows through LED and in result output voltage transits in low state $-U_O \approx 0{,}15$ V. The foiled paper based laminate with sizes 10x15 cm is used to construct contact plate. When lower sensibility is needed the capacitance C_1 can be used as only part where touch is applied.

IV. CALCULATION METHODS FOR THE SENSOR – figure 3.

A. DC point calculation

The Darlington phototransistor ($VT_2 \div VT_3$) is kept in off state by the resistor R_3 and the transistor VT_1 is kept off by the resistor R_2. Output voltage U_O is in high level state when the current through LED I_F is zero.

The only current that flows through the base of Darlington phototransistor is the transistor VT_1 reverse current.

The condition for the transistor VT_1 saturation is:

$$R_S < h_{21E_1}.R_1 \qquad (1)$$

where R_S is the finger skin resistance, h_{21E_1} is the transistor VT_1 current amplify coefficient.

The coefficient h_{21E_1} for the transistor 2T3168B is in the range from 180 to 460 and thus the resistor $R_S < 180.12.10^3$ kΩ = 2,16 MΩ.

The condition for the Darlington phototransistor saturation is:

$$R_1 < h_{21E_2}.h_{21E_3}.R_4 \qquad (2)$$

The LED current I_F which sets the Darlington phototransistor is on state can be described by the following equation:

$$I_F > \frac{U_{CC} - U_{CEsat}}{K_1.R_4} \qquad (3)$$

where K_1 is current transfer ratio of the optocoupler;

U_{CEsat} is the voltage between the Darlington phototransistor collector and emitter in the saturation mode.

The conditions (2) and (3) are implemented as the requirements are extremely exceeded by the resources used in the circuits.

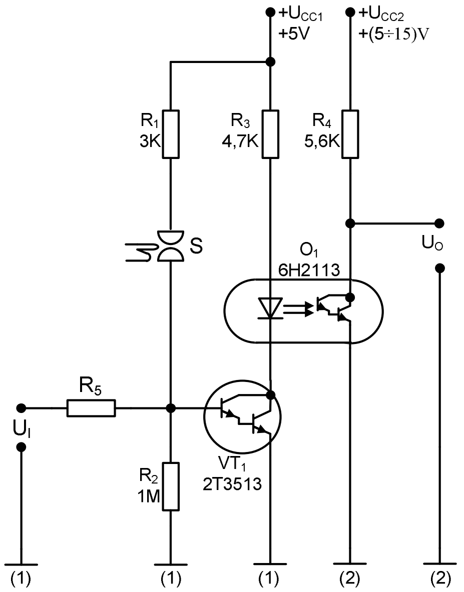


Fig. 6. Sensor button with Darlington transistor

In the circuits shown on figure 1 and 2, JFETs and MOS transistors are used in the inlet circuit to the sensor buttons. The circuit on figure 6 does not use such transistor, only it is also controlled by the inlet circuit. The output voltage transits to low logical level when touch is applied to the two sensor plates.

V. Application

Sensor control of the light sources, photoreceivers, optocouplers, optocoupler ICs as the two methods are used – by means of the introduced resistance of human skin and the inducted by human body voltage over the sensitive part of sensors.

Fig. 1, fig. 2 and fig. 6 – sensors buttons with galvanic decoupling. Fig. 3, fig. 4 and fig. 5 – sensor buttons with photodetector controlled of optocouplers.

VI. Conclusion

The current research intends to additionally expand the functional capabilities to control optoelectrical elements. For the present the known controls over optoelectric elements are by means of electric, optic, magnetic, thermal and mechanical impacts. The current research adds the sensor control to these as the main optocoupler function – the galvanic insulation is intact.

References

[1] Бочаров, Л. Н. *Полевые транзисторы.* М., Энергия, 1976.

[2] Колев, И. С. *Инфрачервена оптоелектроника.* Габрово, Унив. изд. "В. Априлов", 2004.

[3] Колев, И. С. *Оптоелектроника. III* прераб. и доп. изд. Габрово, Унив. изд. "В. Априлов", 2004.

[4] Справочник по полупроводниковым диодам, транзисторам и интегральным схемам. Под ред Н. Н. Горюнова, М., Энергия, 1979.

[5] Цанев, В. Д. *Сензори и сензорни устройства,* С., Техника, 1980.

[6] Siemens. *Optical Information Technology.* Data Book, 1996.

[7] Texas Instruments. *Intelligent Opto-Sensor.* Data Book, 1995.

DDS Method for Generating a Frequency Grid at Systems for Test Control and Automated Regulation

Hristo Z. Karailiev[1] and Valentina V. Rankovska[2]

Abstract – **In the present paper the Direct Digital Synthesis (DDS) method application is suggested. Its main features and a way of operation are examined, and the basic analytical equations characterizing a frequency grid oscillator have been derived.**

Exemplary solutions of an oscillator and a block forming two-phase control signals with programmable duty cycle, based on the above method, have been suggested.

Keywords – **Phase Locked Loop (PLL), Direct Digital Synthesis (DDS), frequency grid, resonant inverters, induction heating.**

I. Introduction

DDS method is widely applied for generating frequencies in many areas as: Creating multi-channel frequency synthesizers at wireless modems, digital transceivers, etc.; Developing functional generators (with sinusoidal, triangular, square, and arbitrary forms of the generated signals); implementing various types of modulations applied in the communication systems, as FSK, QPSK, BPSK, etc.

One of the applications of the method is in the control systems for testing, measurement and regulation the modes of operation of converter devices.

At transistor converter devices control one of the basic functions of the control system is to generate control signals with various parameters (frequency, waveform, etc.) during the technological process. Essential moments of the system operation, which represent interesting problems for solving, are to find the optimal operation mode at the initial start and its assertion during the operation.

At the initial start of the converter or at an enough short arbitrary time interval along its operation we can assume that its Q-factor is a constant. Then we can apply the frequency grid approach consisting of consecutively supplying frequencies from the control system to the converter, measuring the active power detached in it for every of the frequencies and its comparing with the previous one till defining the maximum detached power, i.e. finding the optimal operation mode.

Objective of the report:

Developing a variant of frequency grid oscillator used in the systems for test control and automated regulation.

Main problems of the report:

- Analysis of the Direct Digital Synthesis (DDS) method in relation to generating a frequency grid.
- Developing architecture of a generator. Deriving the basic analytical equations, characterizing the frequency grid oscillator.
- Developing circuits of a frequency grid oscillator and a block forming two-phase control signals using FPGA.

II. Classical approach for generating a frequency grid based on Phased-Lock Loop

A typical solution for generating a frequency grid at the systems for test control and automated regulation of the generated frequency in a definite range and with a definite step is the automated phase regulation of the frequency (APR), implemented with a classical PLL circuit [1, 3].

Very often in practice the output frequency $f_{out} \neq f_{in}$, also it is necessary to be quite higher than the input one and to be automatically asserted constant. In such cases a variant of the classical PLL can be used, called automated phase regulation (APR) method.

In the systems, using frequency grid oscillators, the step of changing the frequencies Δf_{clk} for various objects or one and the same objects but with different characteristics varies. That is why a programmable divider is implemented between the clock oscillator and the phase detector. In order to make an output frequency grid for a definite range from $f_{out}^{\min}$ to $f_{out}^{\max}$ a second frequency divider connecting the oscillator output with the second output of the phase detector have to be implemented.

The drawbacks of the classical structure of PLL and APR for frequency generation due basically to the analogue implementation of some blocks and they can be summarized as follows [1, 3]:

- Output frequency stability dependence on the ambient temperature and aging of the components;
- Limited range of output frequency change ($f_{out}^{\min}$ to $f_{out}^{\max}$), depending on the values of external components of the PLL circuits;
- Low resolution of the generated output frequencies. Increasing the resolution leads to increasing the used hardware and complicating the circuit.

III. Principle of operation and applications of the DDS approach

The new technologies for the integrated circuits development have extremely increased the scale of their integration (microcontrollers, DSP, PLD). At the same time the methods for digital processing of the signals have

[1] Hristo Z. Karailiev is with the Faculty of Electrical Engineering and Electronics, Technical University, 4 H. Dimitar str. 5300 Gabrovo, Bulgaria, E-mail: hkarailiev@gmail.com

[2] Valentina V. Rankovska is with the Faculty of Electrical Engineering and Electronics, Technical University, 4 H. Dimitar str. 5300 Gabrovo, Bulgaria, E-mail: rankovska@tugab.bg

developed too, for instance Direct Digital Synthesis (DDS). The DDS method provides frequencies with a high resolution, depending on the number of bits *n* of *PIR, Σ, PhR*. [1,4,5].

The architecture of a frequency grid oscillator based on the DDS approach is shown on Fig. 1.

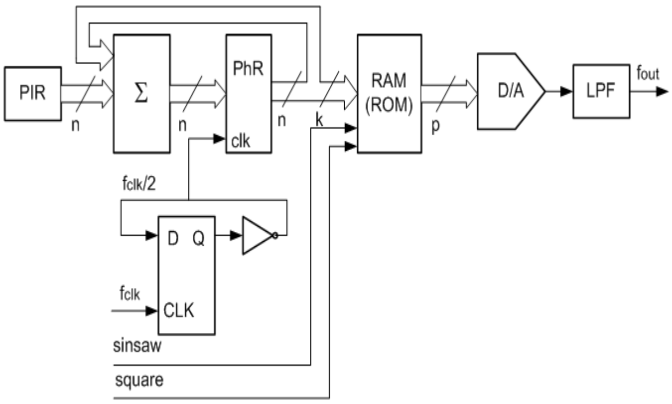


Fig. 1. Architecture of a frequency grid generator based on DDS

PIR	Phase Increment Register – n-bit wide register containing the digital code *M* of the generated output frequency f_{out};
$\sum$	n-bit wide adder, forming the current address of a RAM (ROM) location, containing the current value of the generated frequency
PhR	Phase register which latches the address of the current RAM (ROM) location;
RAM (ROM) –	A memory containing the amplitude values of the generated frequency;
D/A	Digital to analogue converter;
LPF	Low pass digital filter, intended to pass the main frequency and to reject the higher harmonics;
D	Trigger, intended to form twice low frequency then f_{clk}, supplied to the phase register *PhR*;
sinsaw *square*	Control signals, defining the waveform of the generated output signal f_{out};
f_{clk}	Clock signal, forming the sampling step at DDS approach;
L	Digital code, setting the value of the generated output frequency f_{out};
k	Number of master bits of *PhR* used as a current address of the RAM (ROM) location, i.e. $k \in n-1, n-2, ..., n-k\,]$
p	Length of the RAM (ROM) location according to the resolution of the DAC.

The principle of operation of the frequency grid oscillator based on DDS is the following:

In the *n*-bit wide register *PIR* the digital equivalent *L* of the generated output frequency f_{out} has been set. The range of the number, written in *PIR* is:

$$0 \le L \le 2^n \tag{1}$$

where $n \in 0,1,2,...,n-1$ are the numbers of the particular bits of the register *PIR* and adder Σ.

The addresses supplied to RAM, have been formed by adding the number *L*, written into the register *PIR*, to the contents of the phase accumulator *PhR*. A table consisting of *M=2κ* locations had been written into the RAM. The locations are *p*-bits wide and contain information about the form of the generated signal.

At generating of the signal at every $f_{clk}/2$ new current value has been formed in the phase register *PhR*, representing the new address of a RAM location, where is the current value of the generated output signal. Its frequency depends on the speed of changing the addresses in RAM.

In order to separate the basic (useful) signal from the unnecessary harmonics, the output signal from the DAC passes trough a *LPF*, which rejects the high harmonics of non-square signals.

IV. DDS APPROACH USED FOR GENERATING SQUARE WAVES

In the structure of Fig.1 for generating square waves it is not necessary to use *RAM, D/A,* and *LPF*. The output signal is issued by the most significant bit *n-1* of the phase register *PhR.*

It is well known from the Nyquist theorem that at frequency sampling of an analogue signal, the minimal allowed sampling frequency f_{clk}, assuring the correct reconstruction of the primary analogue signal, have to meet the requirement [2]:

$$f_{clk} = 2.f_{out}^{\max} \tag{2}$$

This determines the presence of the D latch in Fig. 3, so the maximum output frequency of square waves will be:

$$f_{out}^{\max} \le f_{clk} / 2 \tag{3}$$

The number on which the output frequency depends, and which is written into the register *PIR*, can be defined by:

$$L = f_{out}.2^n /(f_{clk} / 2) \tag{4}$$

When the number written in *PIR* is known*,* the output frequency value can be defined:

$$f_{out} = L.(f_{clk} / 2) / 2^n = L.f_{clk} / 2^{n+1} \tag{5}$$

The step Δf with which the output signal changes is defined by the equation:

$$\Delta f = f_{clk} / 2^{n+1} \tag{6}$$

We have to consider the following condition for the square waves output frequency f_{out} to be an integer:

$$\Delta f = f_{clk} / 2^{n+1} \tag{7}$$

where Δf =1,2,4,8,… [Hz].

Practically this means that the clock oscillator must be of the type that generates frequencies proportional by power of 2.

When generating square waves it is not sure that the period of the output signal $T_{out}=1/f_{out}$, will be integer, because:

$$\begin{aligned} f_{out} &= 1/T_{out} = L.f_{clk} / 2^{n+1} \\ T_{out} &= (1/L).(2^{n+1} / f_{clk}) = T_{clk}.2^{n+1} / L \end{aligned} \tag{8}$$

V. DDS APPROACH USED FOR GENERATING ARBITRARY FORM SIGNALS

It is necessary to use M number of points to depict one period of the signal at generating signals with non-square form (sinusoidal, triangular, saw, trapezium, etc.). Then the output frequency depends on:

$$f_{out} = 1/M.f_{clk} \quad (9)$$

where: $M=2^k$ is the number of locations of RAM which length is p bits. Then the output frequency will be:

$$f_{out}=1/2^k.f_{clk} \quad (10)$$

The nonlinear distortion of the output signal has been defined by:

$$\eta = F(S, p) \quad (11)$$

where:

S number of points depicting a period of the generated signal;

p resolution of the DAC.

At the principle of operation of the DDS approach usually the speed of the summations is a constant, as it has been defined by the clock frequency f_{clk}. As the number L written in the register *PIR* sets the value of the output frequency, if N is a constant, the output frequency will be constant too:

$$f_{out}=F_1(L,\ f_{clk}, M) \quad (12)$$

As at most circuits the clock frequency $f_{clk}=const$ and the number of points depicting one period of the generated signal $M=const$, then the output frequency will be $f_{out}=F1(L)$.

The relative resolution of the output frequency f_{out} depends on the number of bits n of the register *PIR* and it is defined by:

$$R_{fr} = 1/2^{n-1} \quad (13)$$

The absolute resolution of the output frequency f_{out} is defined by the equation:

$$R_a = R_{fr}.f_{out}^{\max} = 1/2^{n-1}.f_{out}^{\max} \text{ [Hz]} \quad (14)$$

In the industrial converter devices as resonant inverters for hardening and melting metals with the change of the temperature of the object the equivalent parameters of the resonant circuit vary and hence its Q-factor varies in the range from 10 to 2 [6]. According to that fact the limits of the resonant frequency in which the output frequency f_{out} has to be supplied – from $f_{out}^{\min}$ to $f_{out}^{\max}$, vary:

$$\Delta f_{out} = f_{out}^{\max} - f_{out}^{\min} \quad (15)$$

If we consider that absolute resolution R_a from 1 Hz to 10 Hz completely meets the requirements for generating a frequency grid for industrial converters control, using Eq. (14), we can write:

$$\begin{aligned} &2^{n-1} = 1/R_a.f_{out} \qquad \text{or} \\ &n-1 = \log_2[1/R_a(f_{\max} - f_{\min})] \\ &n = \log_2[1/R_a(f_{\max} - f_{\min})] - 1 \end{aligned} \quad (16)$$

Using Eqs. (16), and knowing the range of the frequency band of the resonant circuit of the converter if we take a concrete value for the absolute resolution, we can define the necessary number of bits n of the register *PIR*, Σ and *PhR* to implement the frequency grid oscillator based on DDS approach.

During the technological process (hardening or melting of metals) in the course of time the temperature of the object changes, hence the carried into the resonant circuit of the converter active and reactive component changes too, which necessitates to monitor and regulate automatically the detached power by changing the control frequency [7].

When using the DDS approach to generate a frequency grid it is necessary to estimate the time for resetting the frequency with one step and also the regulation time in order to achieve an optimal power [8].

The time for setting the frequency with one step can be expressed by:

$$\Delta t_{reg} = t_P + t_{SW} + t_{DDS} + t_{DR}, \quad (17)$$

where:

t_P time for measuring the active power

t_{SW} software time for estimating the measured power

t_{DDS} time defined by the hardware used for the DDS method

t_{DR} time for starting the drivers and switching components of the converter

The regulation time is as follows:

$$t_{reg} = n.\Delta t_{reg}, \quad (18)$$

where n is the number of regulation steps to achieve an optimal power.

VI. CIRCUITS OF SOME BLOCKS IN THE MICROPROCESSOR SYSTEM FOR MONITORING AND CONTROL OF RESONANT INVERTERS IMPLEMENTED BY FPGA [7]

The architecture of the blocks DDS1 and DDS2 generating a frequency grid is shown in Fig.2. These blocks are intended to generate square waves with programmable frequency, which value depends on the number, written in the shift and bit registers. The basic elements of the blocks are the shift registers, multiplexer, adder and phase register.

The architecture of a block forming two-phase control signals with programmable duty cycle shown in Fig.3. The block includes a shift register, programmable counter forming the pause duration, triggers, and logical gates.

The suggested blocks have been implemented using Altera FPGA EP1C6T144C8 and integrated development environment Quartus (schematic editor, compiler, simulator and programmer). All logic elements of the chip are 6000 and about 700 have been used. MCU is Atmel microcontroller Atmega128L

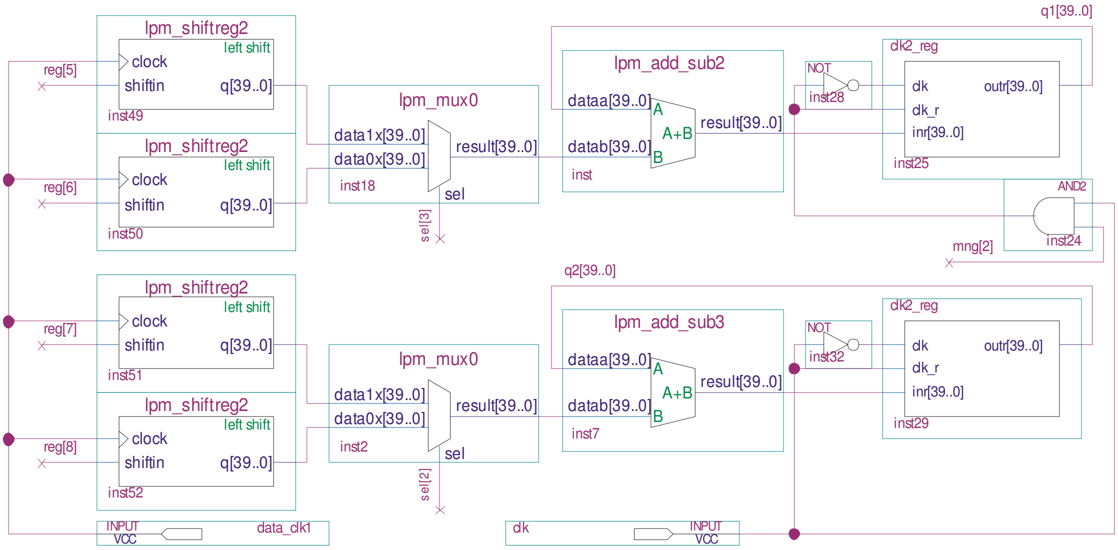


Fig. 2. Architecture of blocks DDS1 and DDS2 for generation of frequency grid

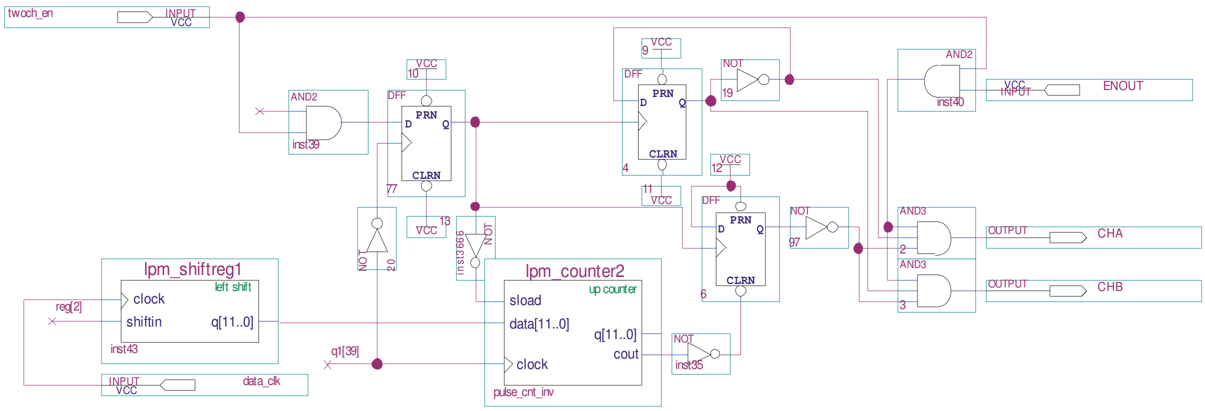


Fig. 3. Architecture of block forming two-phase control signals with programmable duty cycle

VII. Conclusion

The contributions in the present report are as follows:

- A method for generating a frequency grid based on the DDS approach has been suggested. Architecture of frequency grid oscillator allowing generating of square waves, sinusoidal or arbitrary forms has been presented.
- Analytical equations defining the basic characteristics and parameters of a frequency grid oscillator for square waves or arbitrary waveforms have been derived.
- Circuits of DDS frequency grid oscillator allowing a wide range of the generated frequencies and high resolution, and a block forming two-phase control signals using FPGA, have been suggested.

References

[1] Ridiko, L. DDS: Direct Digital Synthesis of Frequencies. "Components and Technologies", № 7, 2001, www.compitech.ru (in Russian)

[2] Ferdinandov, E. Signals and Systems. Sofia, SIELA, 1999r. (in Bulgarian)

[3] Adrian Fox. PLL Synthesizers. Analog Dialogue, vol. 36, part 1, 2002, pp.13-16

[4] A Technical Tutorial on Digital Signal Synthesis. Analog Devices, Inc. 1999.

[5] Eva Murphy, Colm Slattery. All about Direct Digital Synthesis. Analog Dialogue, vol. 38, No3, 2004, pp.8-12

[6] Ch. Hinow, Georg Hinow, "Betrieb eines Wechselrichters mit inductiver Kummulation", Elektrische Bahnen., Heft 11/2004, 28.08.2004, pp.490-493

[7] H. Z. Karailiev, V. Rankovska, V. K. Karailiev. Embedded Microprocessor System for Monitoring and Control of Resonant Inverters. ICEST 2006 Conference Poceedings, Sofia, Bulgaria, 2006 (forthcoming)

[8] H. Z. Karailiev, H. P. Hinov, V. V. Rankovska. Modeling and Investigation of an Inductive Cumulation Inverter. ICEST 2005 Conference Poceedings, Nish, Serbia and Montenegro, 2005, pp.237-240.

29 June – 1 July 2006, Sofia, Bulgaria

Embedded Microprocessor System for Monitoring and Control of Resonant Inverters

Hristo Z. Karailiev[1], Valentina V. Rankovska[2] and Vladimir H. Karailiev[3]

Abstract - **Analyses of some aspects of the processes in induction heating has been done in order to define the necessary parameters of the object (resonant inverter), which have to be measured and controlled in real time.**

A variant of architecture of embedded microcontroller system for monitoring and control of a class of resonant converters is suggested.

The blocks generating the frequency of the two-phase sequence of control signals and forming the duration of the pause when switching the transistors in resonant converter implement the Direct Digital Synthesis approach and a programmable logic device. They allow essential extending the range of control frequencies and increasing the resolution.

Keywords - **embedded microprocessor system, monitoring, resonant inverter, direct digital synthesis, programmable logic devices.**

I. INTRODUCTION

The resonant and quasi-resonant inverters have been wildly spread in induction heating of various materials for surface hardening, surface melting, sticking, casting etc. The proper operation of the system inductor-detail depends on both the parameters and modes of operation of the passive and active components of the inverter and the equivalent parameters of the whole system. During the technological process together with changing the heating temperature these parameters also change in some range. In order to measure the characterizing parameters of the operational modes and to control and/or regulate to achieve optimal operation, it is necessary to know the nature of their change [1].

A variety of architectural and schematic solutions for control of such systems based on various components – analogue and digital – have been published. In order to increase the quality of the control and operational capacities becomes more necessary using intelligent control systems based on microcontrollers, digital signal processors, programmable gate arrays, or combination. [2,3,5,6]. Preconditions for that are the rapid progress in the last decade both of the digital, microprocessor and programmable components and also improving the technologies and the means, which the companies provide to facilitate and to decrease the time for development, testing, debugging and programming of such systems [7].

Implementing intelligent digital systems to control converter devices results in many advantages as:

- More flexible systems, where it is possible to add a wide variety of features and characteristics;
- Implementing some features by software results in decreasing the size, consumption and cost of the system;
- Storing, processing and monitoring of various data about operation of the object and the whole system;
- Low energy consumption by the control system;
- Enhanced opportunities for automated control and diagnostics of the whole system, increasing in such a way the reliability and improving the quality and operational features of the system.

Aim of the report: Development of universal variant architecture of an embedded microprocessor system (EMS) intended for monitoring and control. The system has to be adaptable to operate in a wide frequency range and with high resolution and will be applied to control the signal parameters of a class of resonant inverters (RI).

Main problems of the report:

- Brief analysis of some aspects of the induction heating technological process in order to define the parameters which have to be measured, monitored, controlled and regulated;
- Defining the most common features and operation modes of the EMS;
- Synthesizing the architecture of the EMS using the DDS approach and programmable logic devices to form the control signals of the system.

II. ANALYSES OF SOME ASPECTS OF THE INDUCTION HEATING TECHNOLOGICAL PROCESS IN CONNECTION WITH ITS CONTROL

The bridge and half-bridge resonant inverters are most common power sources (converters) in induction heating of materials with small mass. A classical circuit of a bridge resonant inverter using transistors as switching devices is shown in Fig. 1.

Y1, Y2, Y3, Y4 – Control signals, applied to the transistor gates

CS1, CS2, CS3, CS4, CS5 – Current Sensors, submitting data about the currents trough the transistors and the current of the resonant circuit

VS1 Resonant circuit Voltage Sensor

$\boldsymbol{Z_T}$ Complex impedance of the heated material

I Inductor, implementing the galvanic isolation and the link of Z_T with the parallel resonant circuit of RI.

[1] Hristo Z. Karailiev is with the Faculty of Electrical Engineering and Electronics, Technical University, 4 H. Dimitar str. 5300 Gabrovo, Bulgaria, E-mail: hkarailiev@gmail.com

[2] Valentina V. Rankovska is with the Faculty of Electrical Engineering and Electronics, Technical University, 4 H. Dimitar str. 5300 Gabrovo, Bulgaria, E-mail: rankovska@tugab.bg

[3] Vladimir H. Karailiev is with the Faculty of Electrical Engineering and Electronics, Technical University, 4 H. Dimitar str. 5300 Gabrovo, Bulgaria, E-mail: vkarailiev@gmail.com

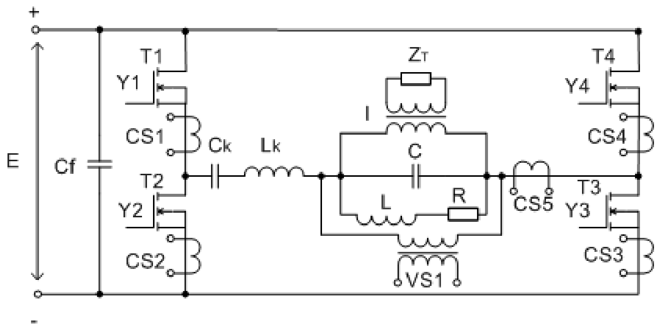


Fig. 1. Classical circuit of a transistor RI

E Source voltage

T1, T2, T3, T4 – Transistors, operating in switch mode, switching the resonant circuit *L,R,C* to the power source *E*.

The necessary active power for heating the material to the previously defined temperature according to the purpose of the induction heating (hardening, soldering, melting, etc.) is a multi-factor value and depends on the following variables:

$$P_M = k_1 (G.\Delta T.C)/\Delta t\,, \tag{1}$$

where:

- P_M Detached into the material heat power
- G Mass of the processed material
- ΔT The difference between the initial and final temperature of the material
- C Special thermal capacity of the material
- Δt The time necessary to reach the final temperature
- k_1 Dimensions equalizing coefficient

The power $P_И$, generated by the RI and detached into the inductor, can be defined taking into account the active losses and the power P_M, detached into the heated material. Using the efficiency of the inductor η:

$$\eta = F_1(\Delta T, S, \Delta t)\,, \tag{2}$$

we can write:

$$P_И = P_M/\eta\,, \tag{3}$$

where *S* is the whole surface of the heated material.

With the increasing of the temperature *T* of the heated material during the technological process its properties vary and hence the active and reactive components of its impedance [1,4] are:

$$Z_T = R_T + j\varpi L_T = F_2(\Delta T) \tag{4}$$

Usually the change of the active and the inductive components passes through three basic phases when heating metals.

It is clear from the study that the active component of the load R_T increases nearly linear with the temperature increasing; hence the equivalent power factor of the resonant circuit during the three phases tends to decreasing and the frequency band of the circuit expands.

In metal heating it is interesting the change of the inductive component of the load L_T. It is clear from the study that during the first and the third phase L_T, although with different values, keeps the characteristics of the inductive load nearly constant with the temperature change. During the second phase when the temperature of the object reaches the Cury point the relative permeability μ_r rapidly becomes 1, hence the load inductivity L_T rapidly decreases and the equivalent inductivity of the circuit rapidly increases:

$$Le = L - L_T' \tag{5}$$

As:

$$Qe = \rho/\mathrm{Re} = \sqrt{Le/C}/\mathrm{Re}\,, \tag{6}$$

the Q-factor of the resonant circuit at Cury point rapidly increases.

As the capacitance *C* in the resonant circuit is a constant when increasing the temperature, the equivalent inductance L_e at the Cury point quickly increases, and the equivalent resonant frequency increases too:

$$\varpi_e = \sqrt{Le.C} \tag{7}$$

The conclusion from the above reasonings is the need of measuring and regulating the control frequency for the RI operation in order to keep the resonant or quasi-resonant mode of operation.

It is necessary to measure the currents through the transistors and the voltage of the resonant circuit to receive information about the power in the resonant circuit.

To control the operation of the RI it is necessary to generate two-phase control signals supplied to Y1 and Y3 with one phase and to Y2 and Y4 with an inverse phase. At that the frequency of the control pulses f_Y have to be regulated in a wide range with maximum resolution for more easy regulating of the power.

Fig. 2 shows the relation of the two control signals.

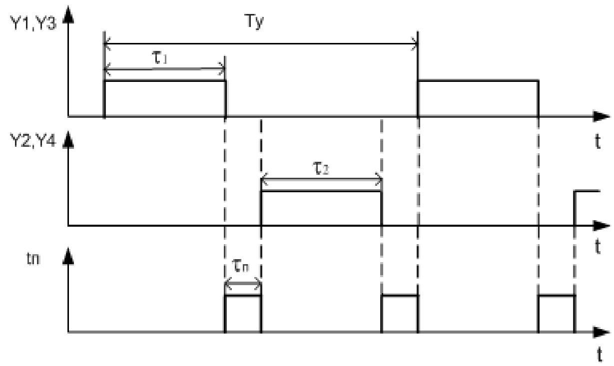


Fig. 2. Control signals, supplying the gates of the switching transistors

where:

$$f_Y = 1/T_Y \tag{8}$$

$$\tau_П = [T_Y - (T_1 + T_2)]/2 \tag{9}$$

As the duration of the control impulse usually is $\tau_1 = \tau_2 = \tau$, the equation becomes:

$$\tau_П = T_Y/2 - \tau \tag{9'}$$

The value of $\tau_П$ can be defined, taking into account two considerations. The first one is that two consecutively connected transistors have not to be turned on, because otherwise the power source *E* will be sort circuited. From the datasheets of the switching transistors it is clear that the following equation is true:

$$t_{on} \leq t_{off} \tag{10}$$

where t_{on} is the turn-on time of the transistor, and t_{off} – its turn-off time. These two times depend on the frequency features and the operation mode of the transistor.

To take into account the first consideration it is necessary to realize the following equation:

$$\tau_{\Pi}^{\min} > 2.k_2(t_{off} - t_{on}), \tag{11}$$

where k_2 is a safety coefficient, which can vary in the following range $1{,}1 \leq k_1 \leq 1{,}5$.

The second fact to consider with when defining the duration of the pause is that with its increasing the detached power in the inductor will decrease, so the power regulation coefficient k_P will be:

$$k_P = \tau/(T_y/2) = 2/T_y(T_y/2 - \tau_{\Pi}) = 1 - 2\tau_{\Pi}/T_y \tag{12}$$

The range of variation of k_P will be:

$$0 \leq k_P \leq 1 - 4k_P(t_{off} - t_{on})/T_y \tag{13}$$

Eq. (12) shows, that except automated regulation of the control frequency, the duration of the pause τ_{Π} has to be controlled too in order to regulate the power in the inductor of the RI.

As a conclusion with a reference to the first problem it is necessary to measure the currents of the transistors and the resonant circuit, the voltage of the resonant circuit, the resonant frequency of the circuit in every moment of the technological process and the temperature of the heated material. In order to control the RI and to allow the possibility to regulate the detached into the heated material active power it is necessary to regulate the controlled frequency f_V and the pause duration τ_{Π}.

III. ARCHITECTURE OF AN EMBEDDED MICROPROCESSOR SYSTEM FOR MONITORING AND CONTROL

Features and operational modes of the embedded system

In order to perform its purpose for a class of resonant inverters in a wide frequency range and a possibility to regulate the power, the embedded microprocessor system for monitoring and control must be able to operate in two basic modes and to have some definite features:

<u>Main operation modes</u> of the system:

- Programming and debugging the software of the control unit MCU. A development system with JTAG interface, PC and proper programming and debug software can be used.
- Setting the mode of regulation, the variation range and the values of the parameters of RI and the technological process.
- Implementing a frequency grid to find the optimal mode of operation of RI, using the DDS block in FPGA. It forms a frequency sequence with a definite step of variation in a previously defined range. The active power detached into the object has been measured automatically by measuring the currents and the voltages in the RI. The optimal power value has been defined using selected sorting approach.
- Operational mode of the control system including three sub-modes:
  - Working mode with automated assertion of the regulation law without statistics and with local monitoring;
  - Working mode with automated assertion of the regulation law with accumulating and processing of statistical data from the object and the technological process;
  - Emergency sub-mode. The system will be put in it manually or automatically when the differential defense turns out and the generation of control signals for the RI will has been stopped.

<u>Main features</u> of the system:

- Setting the ranges of the control frequency, duty cycle, time and temperature parameters of the technological process using keypad. A choice a of regulation law for the technological process.
- Automatic generation of two-channel control frequency sequence galvanically isolated from the object. The two frequencies must be dephased at 180^0 with a changeable duty cycle depending on the power regulation law.
- Using the frequency grid approach at the start mode of the RI and a proper optimization method to define the maximum active power detached in the RI.
- Indicating on LCD or the monitor the operation frequency, the pause, currents, voltages, active power, the processed material temperature, time intervals, etc. Indicating on LEDs the operational mode, emergency sub-mode and the status of the power source.
- Accumulating, processing and displaying statistical information from RI and the technological process.

Architecture of embedded microprocessor system

Description of the system blocks:

JTAG - Standard interface of the development system for programming and debugging the monitor program in MCU and tracing the connections between the blocks in the FPGA

PC Personal Computer for accumulating, statistical processing and monitoring of the measured values from RI and system operation control.

LCD, LEDs &KeyPad – Keypad, display and indicators for setting the operation modes and initial values and indicating the measured values and operation modes

DC/DC ISO – Normalization and galvanic isolation block of the incoming data about the current and voltage from the RI

ISO DR - Isolating Driver forming the control signals with necessary power for the switching components in RI.

CG Clock Generator – generates the switching frequency of the memory components in FPGA and MCU

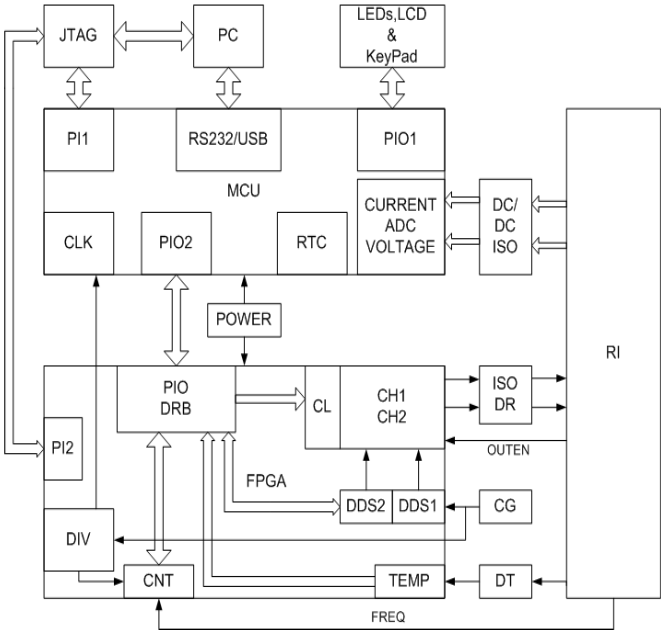


Fig. 3. Architecture of the system for monitoring and control

DT Digital Thermometer, measuring the temperature of the heated material

TEMP – Temperature digital value processing block.

POWER – Power source for MCU and FPGA

OUTEN - Signal enabling/disabling the control signals for the RI. It is formed by the differential defense of RI.

MCU Microcontroller – initializes the FPGA, processes the information received from the RI, implements programming and debugging of the system and interfaces with KeyPad, LCD, LEDs, PC.

PI1, PI2 – Interfaces for programming and debugging of the monitor program in MCU and programming the FPGA

RS232/USB – Interface for connection with PC

PIO1 Parallel Input-Output interface with the KeyPad, LCD and LEDs

ADC Multi-channel ADC receiving the digital code of the currents and voltages in RI.

CLK Clock input for the MCU

PIO2 Parallel Input-Output interface with the FPGA

RTC Real Time Clock

FPGA Field Programmable Gate Array, implementing the DDS control signals oscillator for channel 1 (CH1) and channel 2 (CH2)

PIO DRB – Input-output driver buffer, used by the DDS oscillator, for measuring the temperature of the heated material and/or generated frequency.

CL Digital logic, setting the pause duration at switching the consecutively connected switching devices (transistors).

CH1, CH2 – Channel 1 and channel 2 forming control signals for RI

DDS1, DDS2 – Digital blocks forming square waves

DIV Frequency divider blocks for synchronizing the MCU and CNT

CNT Counter block measuring the frequency of RI

RI Resonant Inverter for induction heating

In the suggested in the present article architecture of the system for monitoring and control the following blocks have been implemented in Altera FPGA EP1C6T144C8 using the integrated development environment Quartus – DDS1, DDS2, CH1, CH2, and CH3. Their functional circuits as a result of the development have been given in [8].

IV. CONCLUSION

The contributions in the present report are as follows:

- A brief analysis of the induction heating technological process and the operation of the RI as a power source has been done. The parameters to be measured, monitored, controlled and regulated have been defined.
- The main features and operational modes have been defined. The architecture of an embedded microprocessor system has been synthesized.
- Architectures of the blocks built in FPGA have been synthesized. DDS approach for generating a frequency grid with a high resolution has been used.

REFERENCES

[1] Sluhotskii, A. E., S. E. Riskin. Inductors for Induction Heating. Energiia, Leningrad, 1974г. (in Russian)

[2] Cassiano Rech, Humberto Pinheiro, Hilton Abilio Grundling, Helio Leaes Hey, and Jose Renes Pinheiro. Comparison of Digital Control Techniques with Repetitive Integral Action for Low Cost PWM Inverters. IEEE Trans. on Power Electronics. Vol. 18. No. 1, 2003, pp. 401-410

[3] Christian Hattrup, Heinz W. van der Broek, and Martin Ossmann.Fast Estimation Techniquies for Digital Control of Resonant Coverters. IEEE Trans. on Power Electronics. Vol. 18. No. 1, 2003, pp. 365-372

[4] H. Karailiev, H. Hinov, V. Rankovska. Modeling and Investigation of Inductive Cummulation Inverter. ICEST 2005 Conference Proceedings, Nish, Serbia and Montenegro, 2005, pp. 237-240

[5] H. Karailiev, V. Rankovska, N. Madjarov. Microprocessor System for Control, Testing, and Regulation of Transistor Converter Devices. ICEST 2004 Conference Proceedings, Bitola, Macedonia, 2004, pp. 457-460

[6] Rafael R. Ramos, Domingo Biel, Enric Fossas, and Farncesc Guinjoan. A Fixed-Frequency Quasi-Sliding Control Algorithm: Application to Power Inverters Design by Means of FPGA Implementation. IEEE Trans. on Power Electronics. Vol. 18. No. 1, 2003, pp. 344-355+

[7] Quartus II Version 5.0 Handbook. Altera Corp., 2005

[8] H. Z. Karailiev, V. Rankovska. DDS Method for Generating a Frequency Grid at Systems for Test Control and Automated Regulation. ICEST 2006 Conference Poceedings, Sofia, Bulgaria, 2006 (forthcoming)

Contactless Charge of an Accumulator and High Voltage Supply of Lighting Lamps from Invertor class "E"

Hristo P. Hinov[1] and Dimitar Tsv. Dimitrov[2]

Abstract - **Tests have been carried out with an inductive cummulation generator, aimed to charge accumulator batteries of small capacity (Mobil phones and portable instruments with accumulator batteries). An aid invertor of the same type supplies high voltage to a lighting lamp. The generator is a one-switch invertor which belongs to the category of generators with inductive cummulation and zero commutational voltage. An inductive link with the load has been used.**

Key words: **Accumulator 6V; 4,5Ah, Mobil device PHILIPS – NiMH, Portable boring machine with an accumulator battery 3,6V; 1,2Ah, High frequency supply of high voltage to lighting lamps.**

I. Introduction

Invertors are sources of high frequency power and the total effect and the efficiency of the supplied by them devices depend on their power. The traditional invertors accumulate after their start power in capacity or capacitance predominating circuit. In the used invertor the power is cumulated in inductivity and capacity fulfils other power and stabilizing functions. The typical for the other invertors transient process lacks in it because it is started and works directly in a steady state regime. The generator is a one-switch inventor.

II. Scheme

The scheme of the charging generator is shown in fig.1. A timer controlled transistor is used as a key element. The inductive and capacitive elements of the circuit are in conformity with the respective load and working frequency of the timer.

III. Tests

A. *Акумулатор* 6V , *4,5Ah*

A contactless charge at $I_{ch.eff}$ = 0,48A has been carried out.

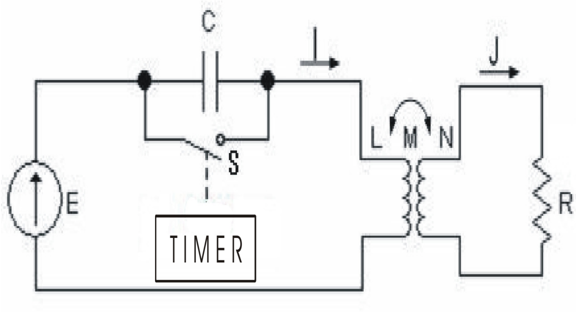


Fig. 1

The time diagrams of the voltage are shown in fig. 2.

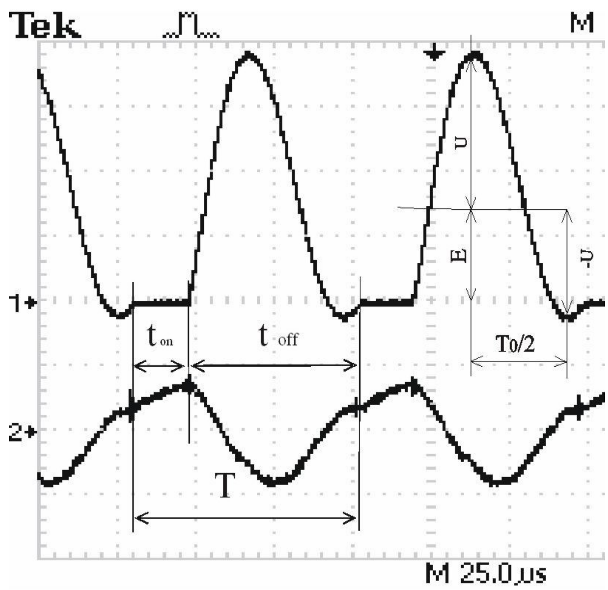


Fig. 2

The working frequency is f_p = 26,3kHz, supply voltage is U = 32,3V.

B. *Mobil phone*

Mobilphone – PHILIPS with a battery NiMH – 4,8V. Supply parameters: U_{ch} = 6,5V, at $I_{ch.eff}$ = 150mA. Working frequency f_p = 33kHz. The time diagrams of the voltages are shown in fig. 3.

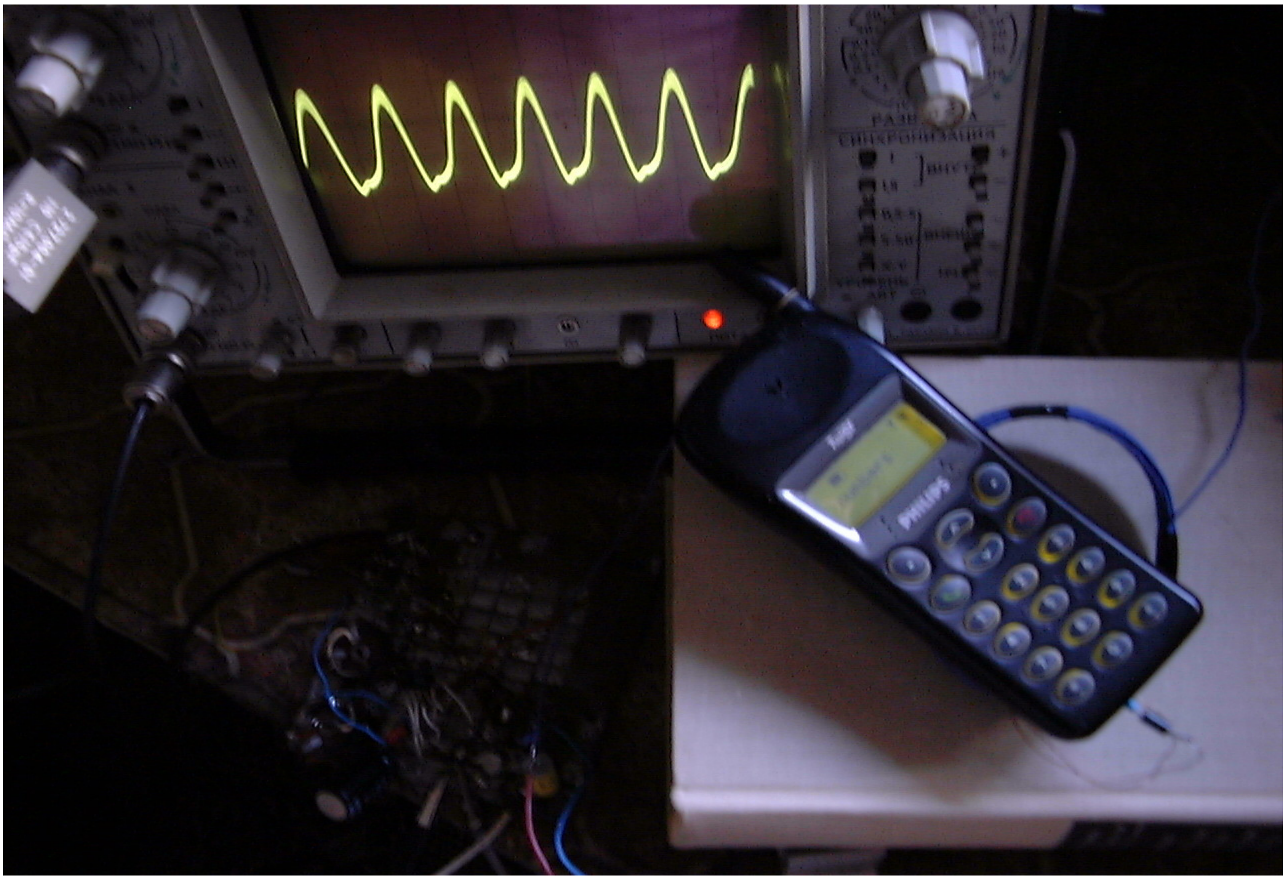


Fig. 3

In the picture of fig.3 the handy device is in a charge regime. In the top left corner of it's display an activated charge indicator can be seen. The lighting is also activated. The handy device is upon a bobbin-inductor, part of it can be seen on the right. It is obvious that a diode connected with the receiving winding bobbin, has been plugged in the supply socket.

Both windings which carry out the necessary magnetic link for the charge process can be seen in the picture – fig.4, when the handy device is turned over. You can also see the stuck receiving winding on the handy device back. The connection of the bobbin by a diode to the handy device charge box can be seen as well.

A professional decision can build the receiving winding in the handy device body like it has been done in the following development.

C. Portable boring machine

A portable boring machine with accumulator batteries – 3,6V, 1,2Ah. Supply parameters: U_{ch} = 5,1V, $I_{ch\ eff.}$ = 80mA. Working frequency f_p = 26,3 kHz. The time diagrams of the voltage are shown in fig.5.

In the picture the boring machine handle is put in the inductor. The necessary magnetic link is achieved and the instrument is in a charge regime. The activated red-lamp indicator can be seen which means that a charge process is available. The receiving winding is built in the handle where the accumulators are as well.

A professional development can constructively design the inductor and the operative invertor class E in a comfortable support nest in which the charged instrument will lie.

D. High frequency supply of lighting lamps with high voltage.

A sodium lighting lamp of 50W power has been supplied. The invertor working frequency is 20 kHz In the picture of fig. 6 the lighted lamp can be seen and fig. 7 – it's reflexion on the oscilloscope display. The time diagrams of the working and control voltage can also be seen in the pictures.

Fig. 4

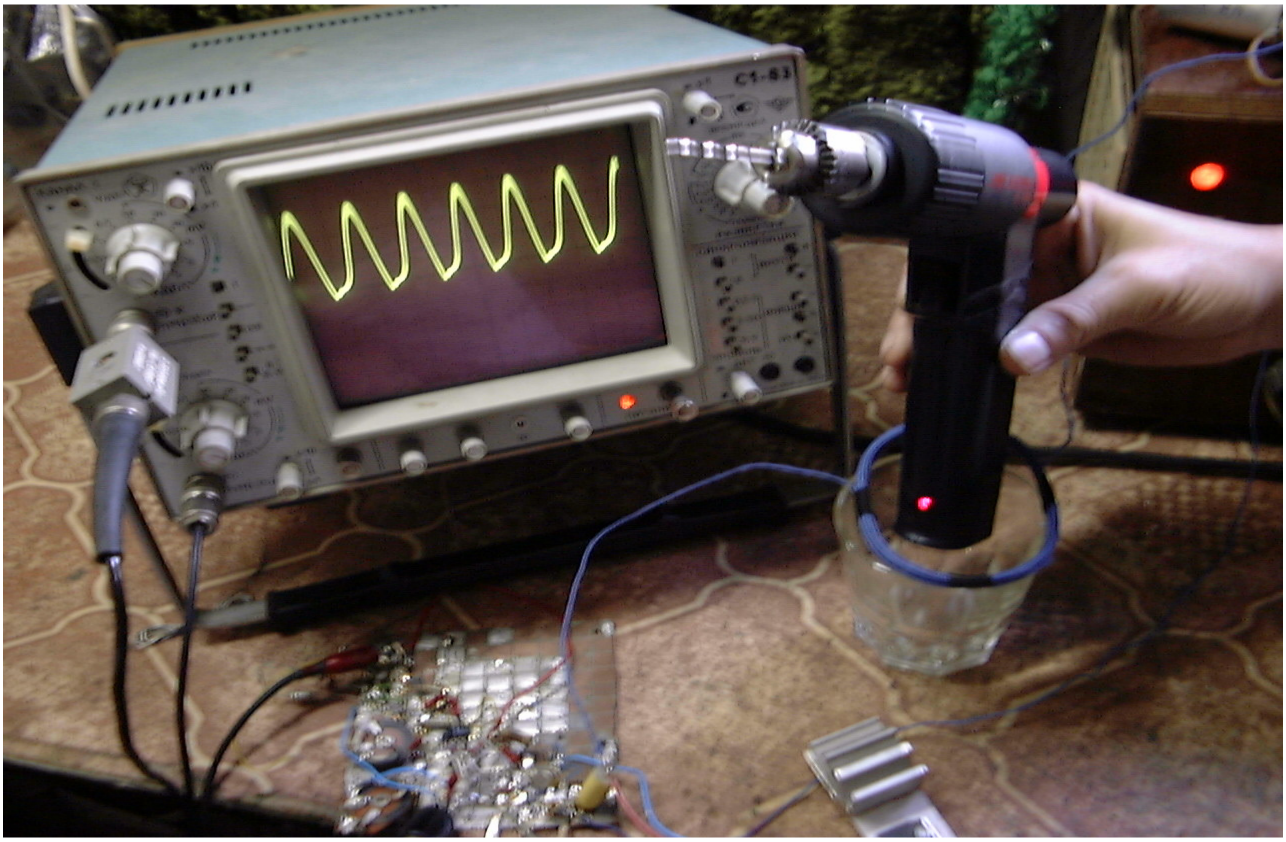


Fig. 5

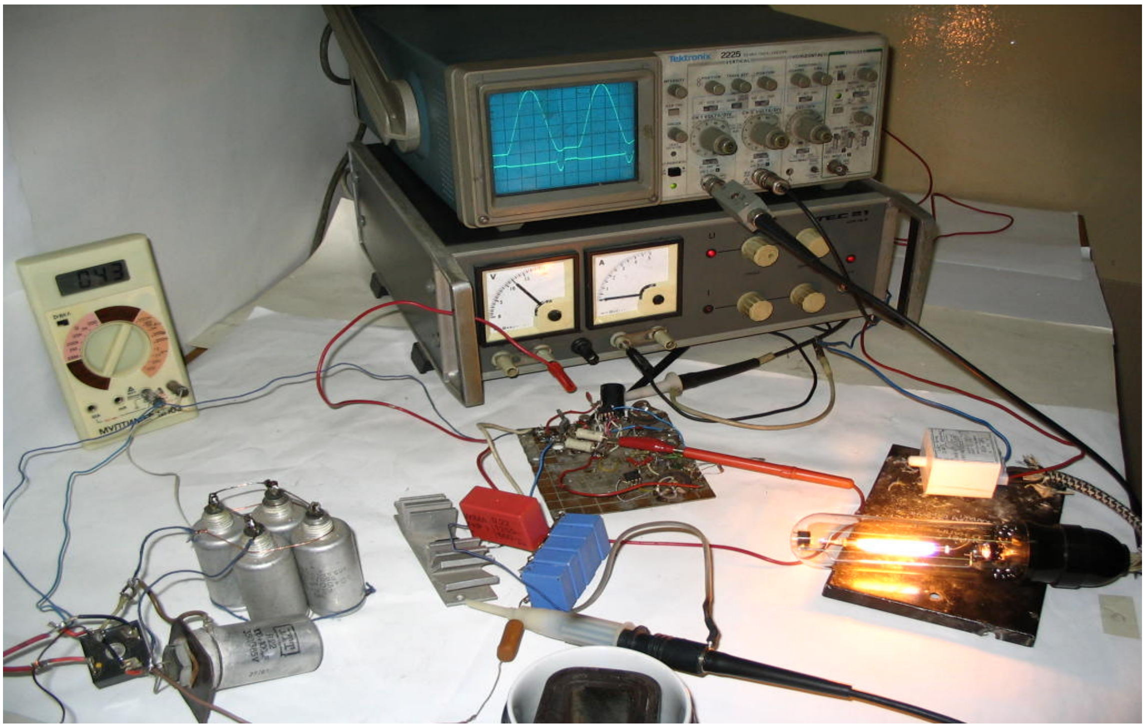


Fig. 6

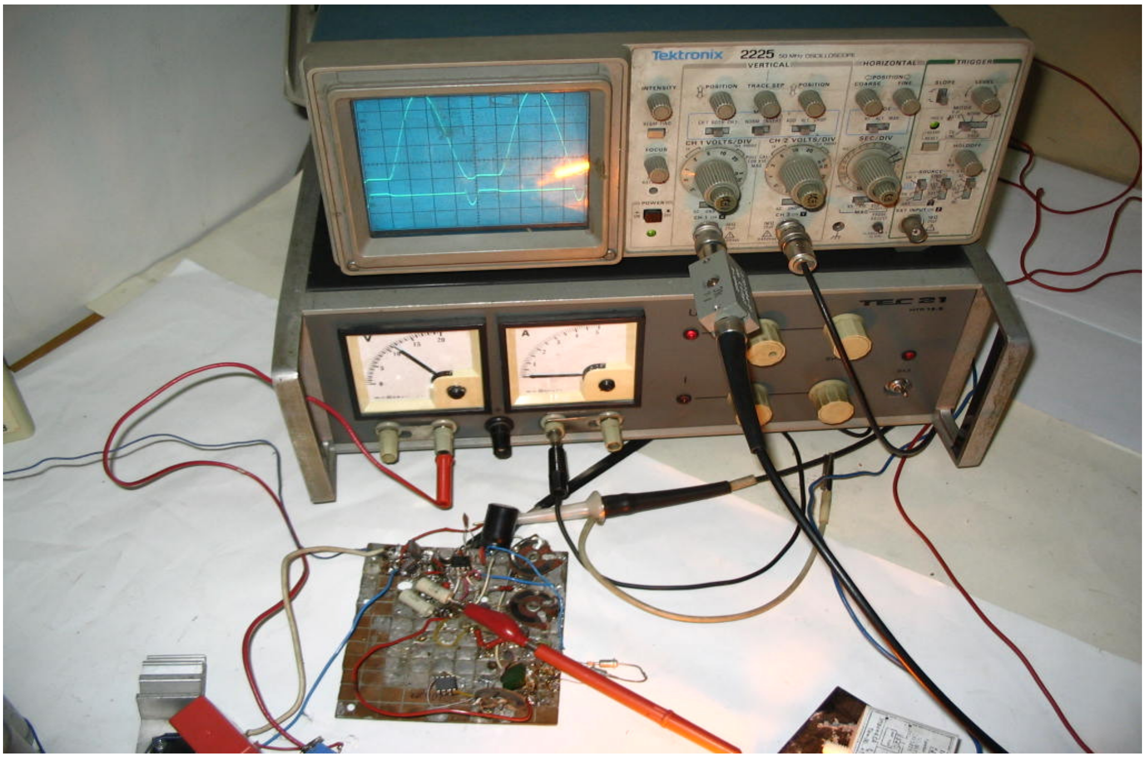


Fig. 7

A professional development can constructively design the inductor and the operative invertor class E in a compact device for assembly next to lighting lamp.

IV. Conclusions

During the carried out tests of steady processes took place at all tests with a reading of the charge by the indicators of the respective devices. The achieved results are on a very good level.

Development and evaluation of GPS aided strapdown INS for land vehicles by means of a Kalman filter

Yasen Angelov[1]

Abstract – **A real-time land vehicle navigation system has been developed. It consists of a GPS receiver and a low-cost inertial block with three axis accelerometer and a yaw rate gyroscope. A time synchronization algorithm is presented. A Kalman filter has been designed and its performance has been evaluated and tested in MATLAB. Two different models of inertial sensors are implemented and tested with real-world experimental data. The aim of the paper is to present a hardware structure of low-cost GPS/INS system for land purposes and outline the main principles and algorithms of the systems' integration and data fusion process.**

Keywords – **Land vehicle navigation, GPS, inertial sensors, Kalman filter**

I. INTRODUCTION

The Global Positioning System (GPS) has made navigation systems practical for a lot of land-vehicle navigation application. Although the purpose of GPS is to provide location in 3D space, in general, a land-vehicle navigation system cannot continuously position using a GPS alone due to the lost of GPS signal in urban environment or in tunnels. Inertial navigation system (INS) is independent and autonomous system. This fact makes is very suitable to complement the GPS shortcomings.

The main problem of the INS is the accumulation of error in velocity, position and angle with time because of the double integration of its data [1].

The fusion of GPS and inertial sensors has been used in various applications [2-4] and the most often used method for the fusion process is the Kalman filter. Integrated GPS/INS can be implemented in different modes: loosely, tightly and ultra-tightly coupled [5].

The integration mode used in this paper is from the first type, where the GPS data is used for calibration of the inertial sensors while the GPS signal is available. The main idea is to implement and evaluate a GPS/INS with very low-cost sensors - inertial devices used [6, 7] have a total price of about $50.

Some real-world experiments were carried out and presented in the paper. The Kalman filter was implemented in MATLAB and two models of the inertial sensors have been implemented and tested.

II. HARDWARE DESIGN

The developed integrated GPS/INS system consists of two main blocks: inertial (Fig. 1) and navigational (Fig. 3).

The inertial block measures the linear accelerations on the three axes and the angular rate in the ground plane.

For full 3D navigation a single gyroscope is not enough but because of the purpose of the system – land navigation, only one gyroscope is implemented.

A very precise analog to digital converter AD7739 [8] was used in this design to minimize the errors from conversion of analog outputs of the sensors to digital data. The MCU takes care of the synchronization of measurements and transfer of gathered inertial data to the other block.

The inertial block's algorithm has been developed to make measurements with frequency of 10Hz which is quite enough for land-vehicle applications [2]. For such an integrated system it is very important to keep it synchronized while working.

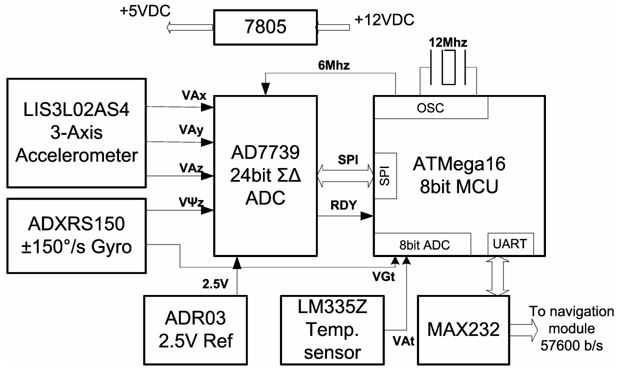


Fig. 1. Block diagram of the inertial sensors sub module

Two types of synchronization have been experimented: single synchronization at the begging and "per-PPS" synchronization on every PPS pulse from the GPS receiver. The first one has insufficient preciseness because of the limited accuracy of the quartz crystals used in both blocks. The difference in frequencies of the crystals causes unbounded increase of the time difference between the measurements of both systems which is unacceptable.

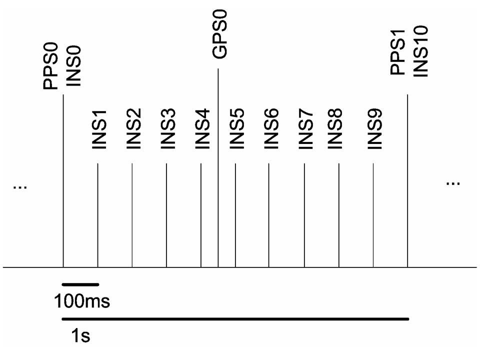


Fig. 2. "perPPS" time synchronization between modules

[1]Yasen Angelov is with the Technical University of Sofia, 8 Kl. Ohridski Blvd, 1000 Sofia, Bulgaria, e-mail: yasen.angelov@dir.bg

The synchronization scheme, proposed on Fig. 2, is used in the system. The navigational block initiates measurement cycle in the inertial block at every PPS pulse. The initiation is implemented via sending a special "start" character from the navigation module to the inertial one. The time offset introduced by the time for sending the "start" character depends only from the baud rate of the RS-232 link. The selected speed of 57600bps gives a time offset of:

$$\Delta T_{SYNC} \approx 0.174 ms .$$

The measurement cycle consists of 10 consecutive measurements of the inertial data set (ax, ay, ax, ψz). A special continuous conversion mode of AD7739 is used which gives a total time for conversion of all the channels of 4ms with channel bandwidth of 500Hz. Every next block of 10 measurements overlaps the previous thus a continuous series is obtained and it is synchronized with the GPS data. Obviously, the ΔT_{SYNC} value is quite small in relation to the ADC conversion time and any error introduced by this parameter can be neglected.

Fig. 3 shows the structure of the navigational module. It is based on powerful ARM7 MCU which handles the tasks for GPS coordinates processing and communications with the GPS receiver and the INS module. Although the MCU do not have floating point unit it is also capable of processing a simple Kalman filter.

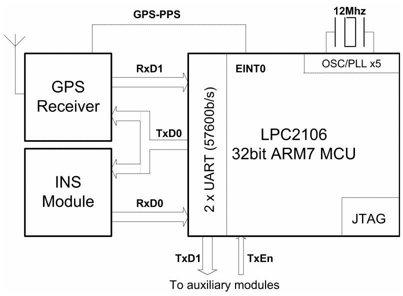


Fig. 3. Block diagram of the navigational sub module

III. DATA FUSION

The first step to integrate both systems is to select a common coordinate system. The assumed plane motion makes the NED frame the most suitable, because inertial data is directly integrated in it. System state in NED is defined as follows (the notation used is from [11]):

$$x_k = [x, y, v, \theta]^T , \quad (1)$$

where *x, y* are the vehicle coordinates in the north and east directions, *v* is the ground speed, θ is the course over ground. Again from the plane motion, the *z* coordinate (down direction in NED) is assumed unchanging while short-term navigating.

The measurement vector (from GPS) has the same form and it is:

$$z_k = [x_{GPS}, y_{GPS}, v_{GPS}, \theta_{GPS}]^T \quad (2)$$

The relation between consecutive time steps in the system state has the following general form:

$$x_{k+1} = x_k + \begin{bmatrix} v_k \cos\theta_k \Delta t \\ v_k \sin\theta_k \Delta t \\ a_k \Delta t \\ \psi_k \Delta t \end{bmatrix}, \quad (3)$$

where $u_k = [a_k, \psi_k]^T$ is the system's driving function built from the linear acceleration and the rotation rate of the vehicle in the plane of motion. $\Delta t = 100ms$ is the time interval between measurements.

An Extended Kalman Filter, based on the algorithm described in [11] and equations from (1) to (4) is built. The initial value x_0 is determined with the help of the GPS receiver:

$$x_0 = [0,0, GSPD, COG]^T , \quad (4)$$

where GSPD is the ground speed reported from GPS and COG is the corresponding course over ground at the starting point P_0 of the integration of the systems.

Equation (3) presents only the general form of the navigation equations when inertial sensors are assumed ideal. Two different models of the inertial sensors have been tested with the filter. The first model is a simple one [12], concerning only the bias and the scale factor of the sensors as constant values. For the accelerometer (the same form is also used for the gyroscope sensor) it has the following form:

$$a_k = (a_1 + a_2.\hat{a}_k) , \quad (5)$$

where a_1 is the bias, a_2 is the scale factor, $\hat{a}_k$ is sensor's output, a_k is the "true" value of the acceleration.

The second model evaluated is proposed in [3] and it is based on exponential functions. It has the following form for the accelerometer:

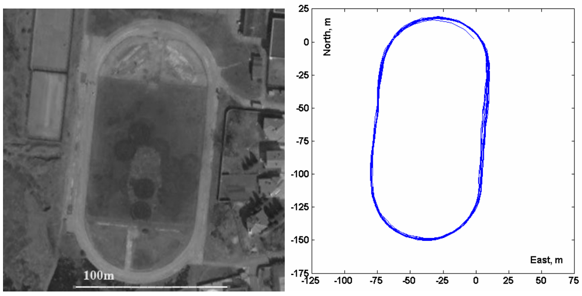
$$a_k = (a_1 + a_2 e^{a_3 t}) + a_4.\hat{a}_k , \quad (6)$$

Experiments with the sensors, used in this design, show that such an exponential processes are also observed in their output.

IV. PERFORMANCE OF THE MODELS IN REAL-WORLD EXPERIMENT

To validate the performance of both models a real time experiments have been carried out. The circular shape of trajectory is known [4] to provide the best performance in determining the horizontal accelerometer biases. Thus a stadium (Fig. 4) is selected as reference trajectory because its shape is very close to the circular.

The selected place for experiments is also suitable because stadiums have very low level of slope so the assumption of plane motion (without giving an account of the *z* coordinate in NED) is more trustworthy. The beginning of the NED (P_0) coordinate system is defined near the center of the northern arc of the stadium (Fig. 4b). The presented experiment consists of 600s continuous driving along the stadium.



a b

Fig. 4. Google Earth's image of the reference trajectory (a) and the corresponding GPS location data in NED (b)

To simulate GPS outage in specified moment, the "measurement update" process is stopped and the EKF continues to perform only the "time update".

The model from Eq. 5 has been evaluated with eight state EKF (the system states from Eq. 3 plus four parameter of the two sensors). To illustrate the estimation of the filter of the sensor's parameters (a_1 and a_2) the simulation is first started with uninterrupted GPS signal. Fig. 4 shows how the filter works on the estimation of the parameter trough the whole driving process of 600 seconds. It is obvious that some of the parameters do not tend towards some constant value and changes all the time. This change is due to the very simple nature of the model and hints that the assumption of constant bias and/or scale factor is not absolutely true. The same conclusion applies also from investigating the diagonal elements of the error covariance matrix: the corresponding error covariance of the changing parameters does not tend (or tend slowly) to a constant value.

A simulation of GPS signal outage has been carried out and the result of the calculated route is shown on Fig. 6. The GPS signal is "stopped" on the 300th second and the navigation remains only inertial using the last estimated sensors' parameters. Even very simple, the model shows good results for the low-cost sensors used. It is well noticeable that the gyroscope is more badly modeled and gives a constantly increasing error in the course over ground. At the end of the inertial navigation the amplitude of the position error in north and east directions reaches about 50m.

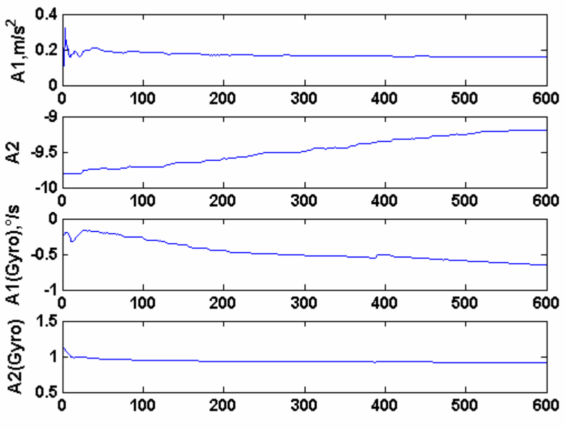


Fig. 5. The estimation process of the sensors' parameters with the aid of GPS data and EKF

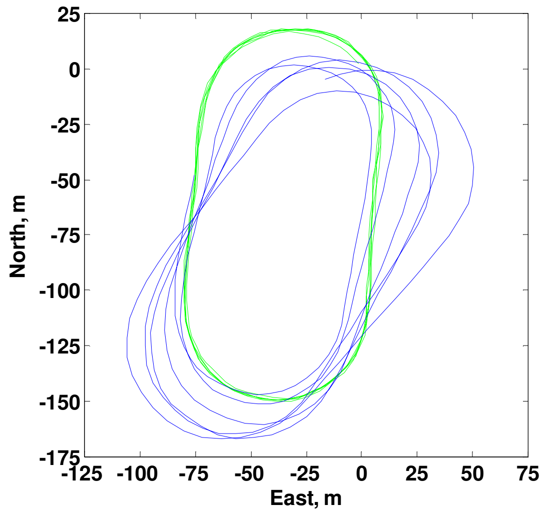


Fig. 6. Performance of the first model when only inertial navigation remains

The second model from Eq. 6 has been tested in the same circumstances. The composed EKF in this case comprises twelve states. The inertial only navigation (Fig. 7) shows better performance in relation to the previous model. Especially the gyroscope performance is improved and course over ground diverges slightly. At the end of the inertial navigation the amplitude of the position error in north and east directions reaches about 25m.

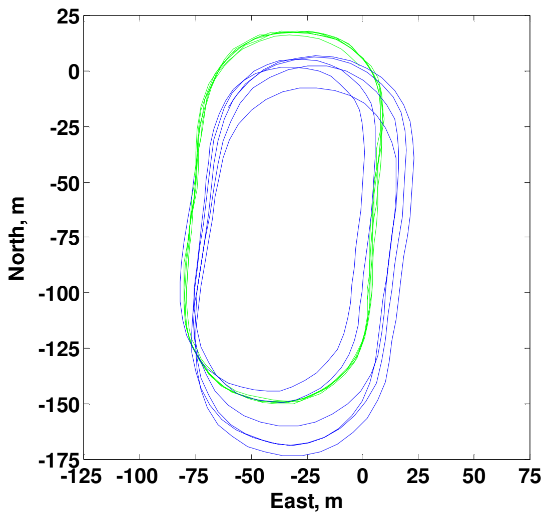


Fig. 7. Performance of the second model when only inertial navigation remains

The exponential model typically shows better performance because it is closer to the real-world sensor's performance but it will require more processing power. For filtering outside the real hardware system the size of the system state is not a concern, but when implementing different models they must be carefully judged to match the available calculating power of the system. Thus for a simpler systems, although it's lower performance, the first model is also applicable due to the smaller size of the system state.

V. CONCLUSION

A hardware design of an integrated low-cost INS/GPS system is proposed in this paper with a time synchronization scheme without time difference accumulation between two systems. An Extended Kalman filter is also synthesized for the data fusion and estimation of the sensors' parameters using GPS as a reference system. Real-world experiments have been carried out to validate the integration process and the performance of two different models for inertial sensors.

The investigated models both show that inertial only navigation can continue for more than 300s with error in position of 25-50m using very low cost inertial sensors and the standard GPS service. The implementation of such a system for a real urban route will require a lot of additional parameters to be considered, like temperature, the slope of the plane of motion which introduces static errors in accelerometers. Thus once has to look at this research more like theoretical approach although carried out with real-world experiments data.

The main contribution of this research work is in the founding of theoretically-experimental setting for evaluation of INS/GPS integrated systems, development and testing different models of the inertial sensors. The author's future effort will be focused namely on the research for more suitable and detailed inertial sensors' models, covering more factors of influence.

VI. REFERENCES

[1] R. Arnaudov, Y. Angelov, *Improvement in method for bias drift compensation in micromechanical gyroscopes*, RADIOENGINEERING Journal, VOL. 14, NO. 2, JUNE 2005, p. 7 -12

[2] G. Linsong, He Yong, Zhang Qin, Han Shufeng, *Real time tractor poition estimation system using a Kalman filter*, Transactions of CSAE, Sept. 2002, Vol. 18, No. 5, p. 96-101.

[3] B. Barshan, Hugh F. Durrant-Whyte, *Inertial Navigation Systems for Mobile Robots, IEE Transactions on robotics and automation*, vol. 11, no. 3, June 1995, p. 328-342

[4] J. Wang, H.K.Lee, S. Hewinston, Hyung-Keun Lee, *Influence of Dynamics and Trajecotry on integrated GPS/INS Navigation Performance*, Journal of Global positioning Systems, 2003, Vol. 2, No. 2, p. 109-116

[5] Alban, Santiago, Akos, D., Rock, S. and Gebre-Egziabher, D., *Performance Analysis and Architectures for INS-Aided GPS Tracking Loops*, Institute of Navigation's National Technical Meeting, January 2003, Anaheim, CA

[6] Analog Devices, *ADXRS150 ±150°/s Single Chip Yaw Rate Gyro with Signal Conditioning*, product datasheet, www.analog.com

[7] ST, *LIS3L02AS4: 3Axis - 2g/6g LINEAR ACCELEROMETER*, product datasheet, www.st.com

[8] Analog Devices, *AD7739 8-Channel, High Throughput, 24-Bit ∑-Δ ADC*, product datasheet, www.analog.com

[9] *Datum Transformations of GPS Positions*, Application Note, 5th July 1999, www.u-blox.ch

[10] J. A. Farrell and M. Barth, *The Global Positioning System & Inertial Navigation*, McGraw-Hill, 1999

[11] G. Welch and G. Bishop, *An Introduction to the Kalman Filter*, SIGGRAPH 2001 course 8. In Computer Graphics, Annual Conference on Computer Graphics & Interactive Techniques. ACM Press, Addison-Wesley, Los Angeles, CA, USA (August 12--17), SIGGRAPH 2001 course pack edition, 2001, http://www.cs.unc.edu/~welch

[12] K. Kimoto, C. Thrope, *Map building with radar and motion sensors for automated highway vehicle navigation,* In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS '97, Volume: 3 , 7-11

Performance Analysis of a Positioning Electric Drive System

Mikho R. Mikhov[1]

Abstract – **The performance of a DC electric drive system with positioning control has been discussed in this paper. Detailed investigation in the respective transient and steady state regimes has been carried out through modeling and computer simulation. The static position error has been determined and the drive system accuracy has been analyzed. The developed models and the results obtained can be used in optimization and final tuning of such types of position drive systems.**

Keywords – **Electric drive system, Positioning control, Position accuracy analysis.**

I. Introduction

Position is one of the main controlled variables in electric drive systems. Movement control of the driven mechanisms is required in many applications, such as:

- machine tools;
- lifting machines;
- woodworking machines;
- manipulators and robots;
- antenna orientation systems;
- radio telescopes, etc.

In accordance with their operation principle, position control systems can be classified into two groups:

- systems with positioning control;
- systems with tracking control.

Positioning is a regime of control providing movement of the respective driven mechanism into a position with required accuracy [1], [4], [5].

This paper considers the performance of a positioning DC drive system with non-linear position controller. The respective transient and steady state regimes of operation have been investigated. The static position error has been determined. The drive system accuracy has been analyzed and discussed.

II. Control Loops Optimization

The electric drive system block diagram is shown in Fig. 1, where the following notations have been used: $G_{pc}(s)$ - transfer function of the position controller; $G_{sc}(s)$ - transfer function of the speed controller; $G_{cc}(s)$ - transfer function of the armature current controller; K_p and τ_p - gain and time-constant of the respective power converter; $R_{a\Sigma}$ - armature circuit resistance; $\tau_{a\Sigma}$ - armature circuit time-constant; K_m - motor coefficient; $\tau_{m\Sigma}$ - summary electromechanical time-constant; K_{cf} - gain of the armature current feedback; K_{sf} - gain of the speed feedback; K_{pf} - gain of the position feedback; V_p - position reference signal; V_s - speed reference signal; V_c - armature current reference signal; V_v - armature voltage reference signal; V_{pf} - position feedback signal; V_{sf} - speed feedback signal; V_{cf} - armature current feedback signal; V_a - armature voltage; I_a - armature current; T_l - load torque applied to the motor shaft; I_l - static armature current; ω - motor speed; θ - angular position.

The position controller output voltage is the respective speed reference signal, i.e.:

$$V_s = K_{pc}(V_p - V_{pf}) = K_{pc}K_{pf}(\theta_r - \theta) = K_{sf}\omega_r , \qquad (1)$$

where K_{pc} is the position controller coefficient.

The reference deceleration for the respective position difference of $\Delta\theta = \theta_r - \theta$ can be determined from Eq. (1) as follows:

$$\varepsilon_{br} = \frac{d\omega_r}{dt} = -\frac{K_{pc}K_{pf}}{K_{sf}}\omega. \qquad (2)$$

In order to achieve a fast positioning process, it is necessary to provide a deceleration with $\varepsilon_{br} = \varepsilon_{b\max} = \text{const}$ at $I_a = I_{a\max} = \text{const}$ [1].

In compliance with Eq. (2), for the position controller coefficient the following expression is obtained:

$$K_{pc} = \frac{K_{sf}\varepsilon_{b\max}}{K_{pf}\omega} . \qquad (3)$$

Given deceleration $\varepsilon_{b\max} = \text{const}$, the position difference $\Delta\theta$ and the motor speed ω get related through the following equation:

$$\Delta\theta = \frac{\omega^2}{2\varepsilon_{b\max}} , \qquad (4)$$

from where the speed can be expressed as:

$$\omega = \sqrt{2\varepsilon_{b\max}\Delta\theta} . \qquad (5)$$

[1]Mikho R. Mikhov is with the Faculty of Automatics,
Technical University of Sofia, 8 Kliment Ohridski Str., 1797 Sofia, Bulgaria,
E-mail: mikhov@tu-sofia.bg

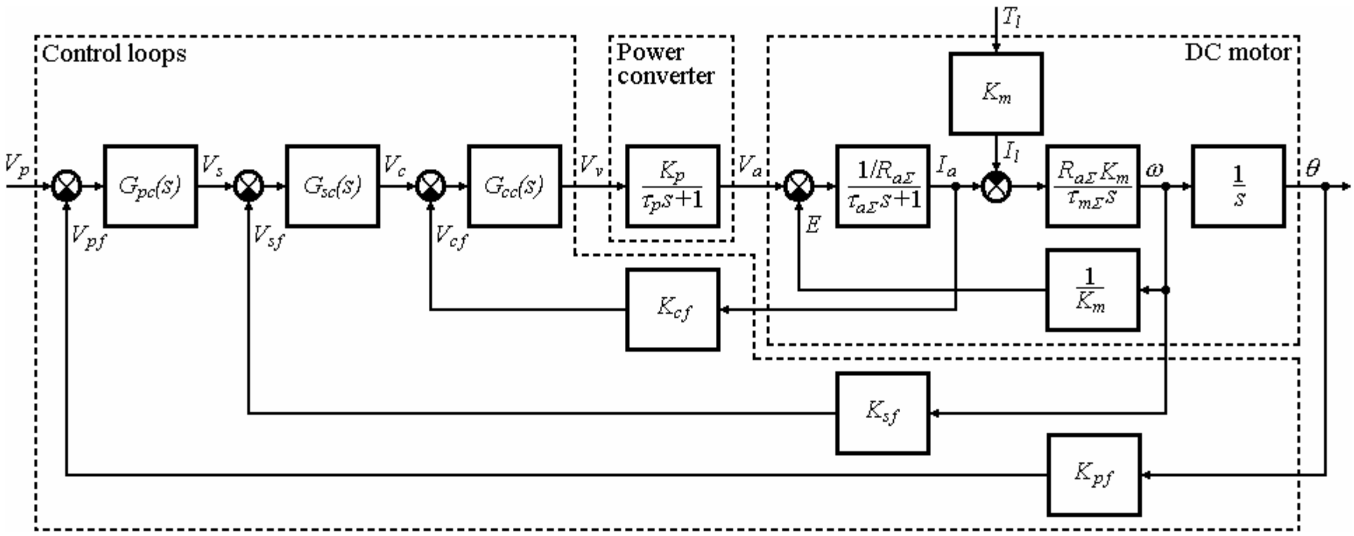


Fig. 1. Block diagram of the discussed electric drive system with positioning control

After substitution of Eq. (5) in Eq. (3), the position controller coefficient becomes:

$$K_{pc} = \frac{K_{sf}}{K_{pf}}\sqrt{\frac{\varepsilon_{b\max}}{2\Delta\theta}}\,. \tag{6}$$

In case of $\Delta\theta = \Delta\theta_{b\max}$ and $\omega = \omega_{\text{rat}}$, substituting Eq. (4) in Eq. (6), the following expression is obtained:

$$K_{pc\min} = \frac{K_{sf}\varepsilon_{b\max}}{K_{pf}\omega_{\text{rat}}}\,. \tag{7}$$

Using Eq. (6) from Eq. (1) the following nonlinear correlation $V_s = f(\Delta\theta)$ is received:

$$V_s = K_{sf}\sqrt{\frac{\varepsilon_{b\max}\Delta\theta}{2}}\,. \tag{8}$$

Reduction of $\Delta\theta$ to $\Delta\theta_s$ brings about an increase of K_{pc} up to a value equal to K_{pcs}, calculated according to the admissible overshoot:

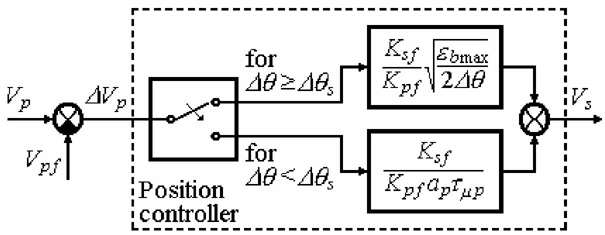
$$K_{pcs} = \frac{K_{sf}}{K_{pf}a_p\tau_{\mu p}}\,, \tag{9}$$

where: $\tau_{\mu p}$ is the respective small time-constant of the position loop, not subject to compensation.

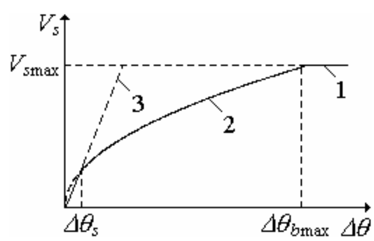
The positioning loop dynamic characteristics depend on the coefficient a_p value. For example, to provide for a transient process $\theta(t)$ without overshoot, the following condition should be fulfilled:

$$a_p \geq 4. \tag{10}$$

The block diagram of the synthesized position controller and its characteristic are shown in Fig. 2. The used notations are as follows: $\Delta\theta_{b\max}$ - maximum braking angle; $\Delta\theta_s$ - braking distance at which the controller coefficient value is switched over; $V_{s\max}$ - maximum speed reference signal. At $\Delta\theta \geq \Delta\theta_s$ the speed reference signal is $V_s = \text{const}$, while at $\Delta\theta < \Delta\theta_s$ the $V_s(\Delta\theta)$ characteristic coincides with the straight line 3.



a)



Speed reference signal

$V_s = f(\Delta\theta)$:

1 - $\Delta\theta \geq \Delta\theta_{b\max}$

2 - $\Delta\theta_{b\max} > \Delta\theta \geq \Delta\theta_s$

3 - $\Delta\theta < \Delta\theta_s$

b)

Fig. 2. Block diagram of the position controller and its characteristic

Using the MATLAB/SIMULINK software package a number of computer simulation models of electric drives with position control have been developed. A detailed study of the drive system under consideration has been carried out for the respective dynamic and static regimes.

III. Drive System Performance Analysis

Fig. 3 shows some simulation results illustrating the performance of the discussed electric drive system. Acceleration and deceleration are represented, as well as the operation in steady state regime. The reference position is 800 rad and the motor speed is limited to the rated value of $\omega_{\text{rat}} = 314\,\text{rad/s}$. The load torque acting upon the motor shaft is equal to the rated value of $T_{l\text{rat}}$. During the respective transient regimes the armature current is limited to the maximum admissible value of $I_{a\max} = 35.2\,\text{A}$, which provides good dynamics of the drive system.

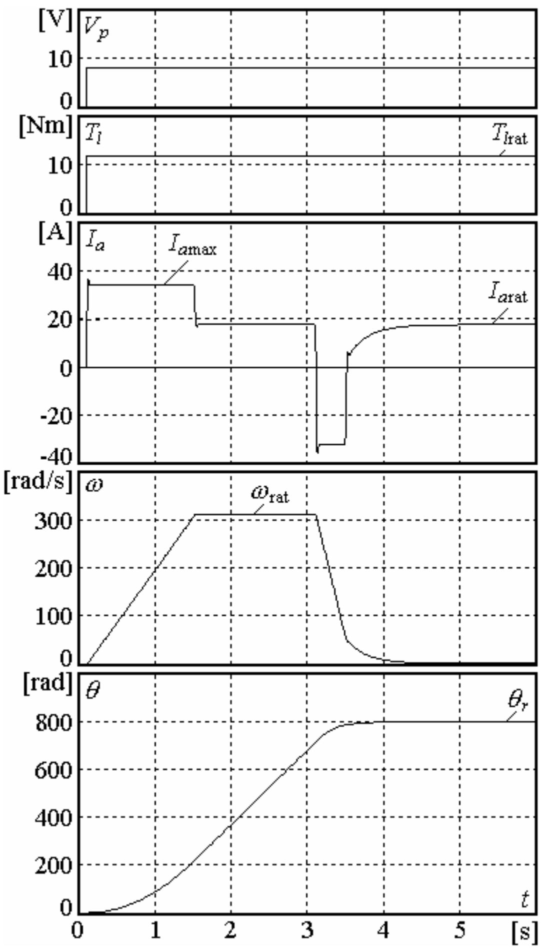


Fig. 3. Time-diagrams illustrating the drive system performance

Limitation of the static position error to the admissible value is a basic requirement in positioning systems as far as their regulation accuracy is concerned.

In order to determine the position error, the system block diagram has been transformed as shown in Fig. 4. This diagram has a control input of θ_r and a disturbance input of T_l. To make it convenient, the input and the controlled variable dimensions have been assumed to be equal, which means that the position feedback coefficient is $K_{pf} = 1$.

The optimized current control closed-loop is represented by the following transfer function:

$$M_c(s) = \frac{I_a(s)}{V_c(s)} = \frac{1/K_{cf}}{a_c\tau_{\mu c}^2 s^2 + a_c\tau_{\mu c}s + 1} \approx \frac{1/K_{cf}}{a_c\tau_{\mu c}s + 1}, \quad (11)$$

where: a_c is a coefficient influencing the current loop dynamic characteristics.

In the general case, the summary error of the drive system for the controlled variable θ can be determined using the following equation:

$$\Delta\theta_{\Sigma}(s) = \frac{\theta_r(s) + T_l(s)G_d(s)}{1 + G(s)}, \quad (12)$$

where: $G_d(s) = \theta(s)/-T_l(s)$,

$G(s) = \theta(s)/\theta_r(s)$.

The respective transfer function $G_d(s)$ can be represented as follows:

$$G_d(s) = \frac{K_{cf}K_m(a_c\tau_{\mu c}s + 1)}{G_{sc}(s)} M_s(s)\frac{1}{s}, \quad (13)$$

where: $G_{sc}(s)$ is the transfer function of the used speed controller;

$M_s(s)$ -the transfer function of the speed closed-loop.

After the respective tuning of the speed controller, the following expression is obtained for $M_s(s)$:

$$M_s(s) = \frac{\omega(s)}{V_s(s)} = \frac{1/K_{sf}}{a_s\tau_{\mu s}s(\tau_{\mu s}s + 1) + 1}. \quad (14)$$

Taking into account Eq. (14), the $G_d(s)$ transfer function is expressed as follows:

$$G_d(s) = \frac{K_{cf}K_m(a_c\tau_{\mu c}s + 1)}{\dfrac{K_{cf}\tau_{m\Sigma}}{K_mK_{sf}R_{a\Sigma}a_sa_c\tau_{\mu c}}}\left[\frac{1/K_{sf}}{a_sa_c\tau_{\mu c}s(a_c\tau_{\mu c}s + 1) + 1}\right]\frac{1}{s} =$$

$$= \frac{K_m^2R_{a\Sigma}a_sa_c\tau_{\mu c}(a_c\tau_{\mu c}s + 1)}{[a_sa_c\tau_{\mu c}s(a_c\tau_{\mu c}s + 1) + 1]\tau_{m\Sigma}s}, \quad (15)$$

where: a_s is a coefficient influencing the dynamic characteristics of the speed control loop;

$\tau_{\mu c} = \tau_p$ is the small time-constant of the current loop, not subject to compensation.

In compliance with Fig. 4, using Eq. (9), the open-loop transfer function becomes as follows:

$$G(s) = \frac{K_{sf}}{a_pa_sa_c\tau_{\mu c}}\left[\frac{1/K_{sf}}{a_sa_c\tau_{\mu c}s(a_c\tau_{\mu c}s + 1) + 1}\right]\frac{1}{s} = \frac{1}{a_pa_sa_c\tau_{\mu c}s[a_sa_c\tau_{\mu c}s(a_c\tau_{\mu c}s + 1) + 1]}. \quad (16)$$

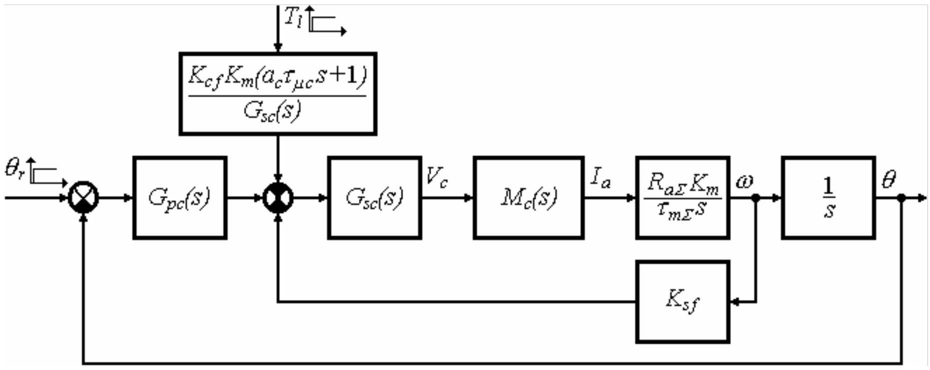


Fig. 4. Modified block diagram of the positioning control drive system

After substitution of Eqs. (15) and (16) in Eq. (12), the following expression is received for the summary error of the controlled variable:

$$\Delta\theta_\Sigma(s) = \frac{\theta_r(s)a_p a_s a_c \tau_{\mu c} s[a_s a_c \tau_{\mu c} s(a_c\tau_{\mu c}s+1)+1] + \dfrac{T_l(s)K_m^2 R_{a\Sigma} a_p a_s^2 a_c^2 \tau_{\mu c}^2 (a_c\tau_{\mu c}s+1)}{\tau_{m\Sigma}}}{a_p a_s a_c \tau_{\mu c} s[a_s a_c \tau_{\mu c} s(a_c\tau_{\mu c}s+1)+1]+1} . \quad (17)$$

The static error of the investigated drive system is obtained from Eq. (17) at $s = 0$:

$$\Delta\theta_s(s) = \frac{K_m^2 R_{a\Sigma} a_p a_s^2 a_c^2 \tau_{\mu c}^2}{\tau_{m\Sigma}} T_l(s) . \quad (18)$$

As seen from the last equation, in this case the position error of the system depends on the load torque, applied to the motor shaft.

The drive system parameters as well as the calculated position error for one of the investigated versions have been represented on Table I.

TABLE I
PARAMETERS AND STATIC ERROR OF THE DRIVE SYSTEM

Parameter	Dimension	Value
K_m	Nm/rad; rad/sV	1.484
$R_{a\Sigma}$	Ω	1.33
a_p	-	4
a_s	-	2
a_c	-	2
$\tau_{\mu c}$	s	0.004
$\tau_{m\Sigma}$	s	0.15
T_l	Nm	11.862
θ_r	rad	800
$\Delta\theta_s$	rad	0.1186
$\Delta\theta_s$	%	0.015

IV. CONCLUSION

The main features of a positioning drive system with nonlinear position controller have been described and discussed. The rated parameters of the used separately excited DC motor are as follows: $P_{\text{rat}} = 3.4\,\text{kW}$, $V_{1\text{rat}} = 220\,\text{V}$, $I_{1\text{rat}} = 17.6\,\text{A}$, $\omega_{\text{rat}} = 314\,\text{rad/s}$.

Detailed investigation has been carried out through modeling and computer simulation for the respective transient and steady state regimes of operation. The static position error has been determined and the positioning drive system accuracy has been analyzed.

The developed simulation models and the results obtained can be used in optimization and final tuning of such types of position drive systems.

REFERENCES

[1] V. I. Klyuchev, *Electric drive theory*, Moscow, Energoatomizdat, 1985 (in Russian).
[2] U. Keuchel, R. M. Stephan, *Microcomputer-based adaptive control applied to thyristor-driven DC motors*, London, Springer-Verlag, 1994.
[3] C. Ong, *Dynamic simulation of electric machinery,* New Jersey, Prentice Hall, 1998.
[4] I. Boldea, S. A. Nasar, *Electric drives*, Boca Raton, CRC Press, 1999.
[5] N. Mohan, *Electric drives – an integrative approach,* Minneapolis, MNPERE, 2003.
[6] M. R. Mikhov, “Modeling and investigation of a position control drive system”, *Proceedings of the International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST’04)*, pp. 461-464, Bitola, Macedonia, 2004.

Modeling and Analysis of a Switched Reluctance Motor Drive System

Mikho R. Mikhov[1]

Abstract – **The behavior of a switched reluctance motor drive system with speed control has been analyzed in this paper. The investigations in the respective dynamic and static regimes have been carried out through modeling and computer simulation. The developed models and the results obtained can be used as illustration in the process of teaching about such types of drive systems.**

Keywords – **Electric drive systems, Switched reluctance motor drives.**

I. Introduction

Switched reluctance motors (SRMs) have some features that make them an attractive alternative to existing DC and AC motors in adjustable speed drives. The SRM's main advantages can be summarized as follows [1], [2], [3]:

- simple and robust structure;
- each phase winding is independent and this makes the machine highly reliable;
- low power losses and high efficiency;
- high power-mass ratio;
- good heat dissipation characteristics;
- high torque-to-inertia ratio and good dynamics;
- motor torque is independent of the phase current polarity, thus the respective converter usually requires only one switch per phase;
- high speed capabilities.

In recent years SRM drive systems have been used in many applications such as electric vehicles, robotics, textiles, aerospace, office automation, machine tools, and many more [3].

II. Specific Features of SRM Drives

SRMs are doubly salient, singly excited electrical machines with passive (windingless) rotors. The motor phases are turned on sequentially through DC voltage pulses, which result in unipolar controlled current. Motion is produced because of the variable reluctance in the air gap between the stator and the rotor. When a stator winding is energized, producing a single magnetic field, reluctance torque is produced by the tendency of the rotor to move to its minimum reluctance position [1].

The cross-section of the SRM under consideration is shown in Fig. 1. The motor is made of laminated stator and rotor cores with 6 poles on the stator and 4 poles on the rotor. The phase number is $m = 3$ and each phase has concentrated coils placed on one pair of stator poles.

[1]Mikho R. Mikhov is with the Faculty of Automatics, Technical University of Sofia, 8 Kliment Ohridski Str., 1797 Sofia, Bulgaria, E-mail: mikhov@tu-sofia.bg

The magnetic circuit symmetry leads to almost zero mutual flux linkage in the SRM phases even under saturated conditions. This means that the motor may work with $m-1$ phases since no induced voltage or current will appear in the short-circuited phase.

The phase self-inductance L_i varies with the rotor position, and in presence of saturation the respective dependence is nonlinear (Fig. 2).

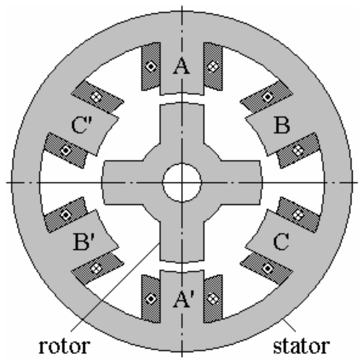


Fig. 1. Cross-section of a 6/4 switched reluctance motor.

The magnetic saturation influence is evident from Fig. 3, where a flux curve family is shown.

The motoring and regenerative regimes are illustrated with the simplified diagrams shown in Fig. 3. Positive (motoring) torque is produced when the inductance is rising as the shaft angle is increasing. A negative (breaking) torque is produced by supplying the motor winding with current while the inductance is decreasing.

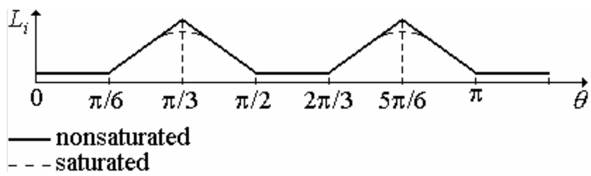


Fig. 2. The self-inductance profile for one motor phase.

The average torque T_i can be controlled by adjusting the winding current magnitude I_i or by varying the dwell angle θ_d . To reduce the torque ripples, it is advisable to keep the dwell angle constant and change the current magnitude.

The communication sequences are shown in Fig. 4.

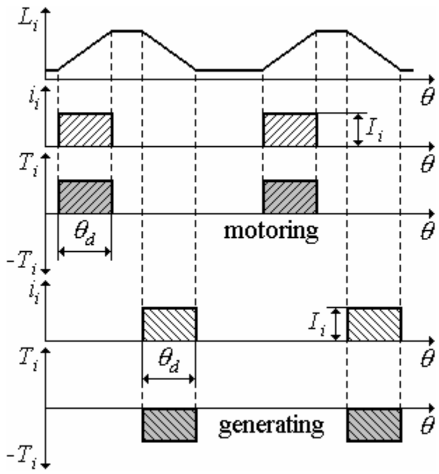


Fig. 3. Diagrams illustrating the operation modes of the motor.

The main problems of SRM drive practical implementation are as follows:

- the SRMs must always be electronically commutated and thus cannot run directly from a DC bus or an AC line;
- their salient structure causes strong non-linear magnetic characteristics, complicating its analysis and control;
- the pulsed nature of torque production leads to torque ripple and noisy effects.

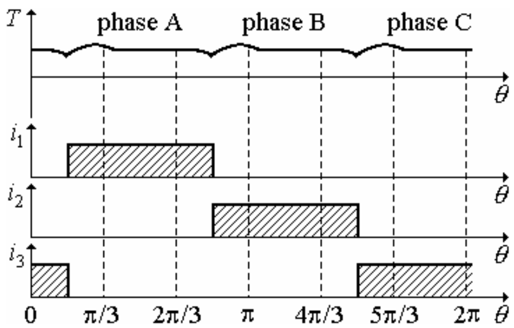


Fig. 4. Commutation sequences.

The major disadvantage of having a high torque ripple can be overcome by using suitable control methodologies [2], [10].

III. Modeling of the SRM Drive System

A simplified block diagram of the drive system under consideration is shown in Fig.5. The corresponding notations are as follows: SC - speed controller; RC - reference current block; CC - current controllers block; CB - converter control block; PC – power converter; PS - position sensor; SF - speed feedback block; CF – current feedback block; L - load of the electric drive; V_{sr} - speed reference signal; V_{cr} - three-phase current reference signals for the phases *a*, *b* and *c* respectively; $2\Delta i$ - reference hysteresis band; V_{sf} - speed feedback signal; V_{cf} - current feedback signals; V_d - DC link voltage; C - filter capacitor; ω - angular motor speed; θ - angular position; T_l - load torque applied to the motor shaft.

The electric equations used to represent the *m*-phase SRM are as follows [2], [3], [4]:

$$\left.\begin{aligned} v_1 &= R_s i_1 + \frac{d\Phi_1(\theta, i_1)}{dt}; \\ &\dots \\ v_i &= R_s i_i + \frac{d\Phi_i(\theta, i_i)}{dt}; \\ &\dots \\ v_m &= R_s i_m + \frac{d\Phi_m(\theta, i_m)}{dt}, \end{aligned}\right| \tag{1}$$

where:

$v_{i(i=1\div m)}$ is the voltage applied across the respective stator winding;

i_i - the phase current;

R_s the stator phase resistance;

$\Phi_i(\theta, i_i)$ - the phase flux linkages for a given rotor position θ and excitation current i_i.

The motion equations are as follows:

$$J\frac{d\omega}{dt} = T + T_l; \tag{2}$$

$$\frac{d\theta}{dt} = \omega, \tag{3}$$

where:

J is the total inertia referred to the motor shaft;

T - the motor torque.

The total instantaneous torque is given by:

$$T = \sum_{i=1}^{m} T_i, \tag{4}$$

where the respective component is:

$$T_i = \frac{\partial}{\partial \theta}\int_0^{i_i} \Phi_i(\theta, i_i)\, di_i. \tag{5}$$

The equation for each phase can be expressed as follows:

$$\begin{aligned} v_i &= R_s i_i + \frac{\partial \Phi_i(\theta, i_i)}{\partial i_i}\frac{di_i}{dt} + \frac{\partial \Phi_i(\theta, i_i)}{\partial \theta}\frac{d\theta}{dt} = \\ &= R_s i_i + L_i(\theta, i_i)\frac{di_i}{dt} + E_i(\omega, \theta, i_i), \end{aligned} \tag{6}$$

where:

$L_i(\theta, i_i) = \dfrac{\partial \Phi_i(\theta, i_i)}{\partial i_i}$ is the instantaneous inductance;

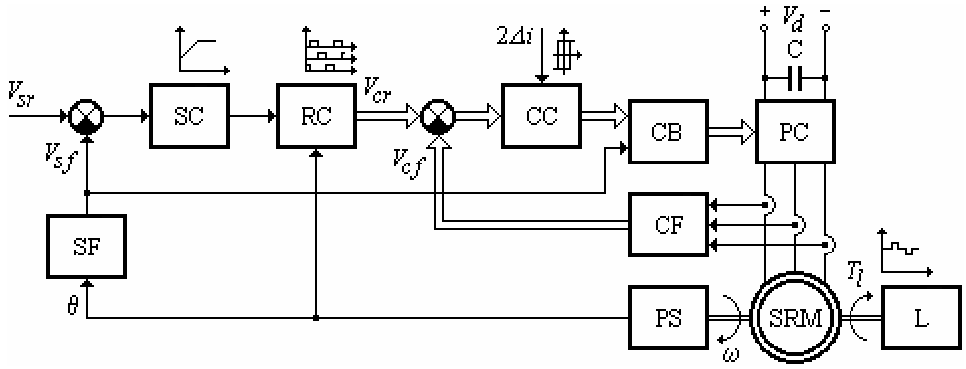


Fig. 5. Block diagram of the drive system under consideration

$$E_i(\omega,\theta,i_i) = \frac{\partial \Phi_i(\theta,i_i)}{\partial \theta}\omega$$ - the instantaneous back EMF.

The SRM mathematical model is highly nonlinear due to the magnetic saturation influence on the $\Phi_i(\theta,i_i)$ curve family (Fig. 6). Only in the absence of saturation, the instantaneous torque is:

$$T = \sum_{i=1}^{m} \frac{1}{2} i_i^2 \frac{\partial \Phi_i(\theta,i_i)}{\partial \theta}. \quad (7)$$

There are several possible strategies to energize a SRM. For this study hysteresis current control has been chosen. The respective phase current controllers have a programmable hysteresis band $2\Delta i$, which determines the modulation frequency of the power converter. The principle of current chopping is shown in Fig. 7, where i_{ir} is the respective reference phase current waveform.

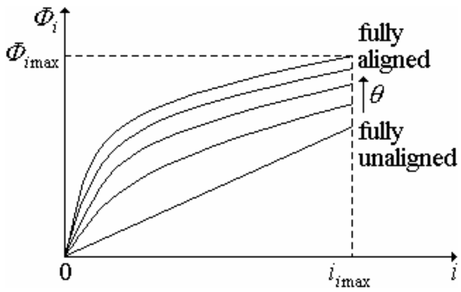


Fig. 6. Flux/current curves.

The current i_i cannot instantaneously rise or fall in the respective phase circuit. For that reason, the voltage to the stator winding is applied in advance by θ_a and the current turn-off is initiated in advance by θ_b. To optimize the drive system efficiency, it is necessary to choose appropriate values for θ_a and θ_b.

The following torque/speed zones can be distinguished (Fig.8):

1. $T = \text{const} \quad (0 < \omega \le \omega_1)$;
2. $T\omega = \text{const} \quad (\omega_1 < \omega \le \omega_2)$;
3. $T\omega^2 = \text{const} \quad (\omega_2 < \omega \le \omega_{\max})$.

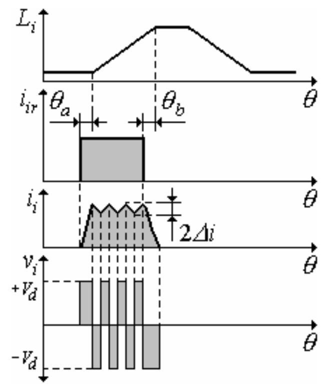


Fig. 7. Operational waveforms of the DC supply converter.

The angles θ_a and θ_b values should be corrected in compliance with both speed and operation modes.

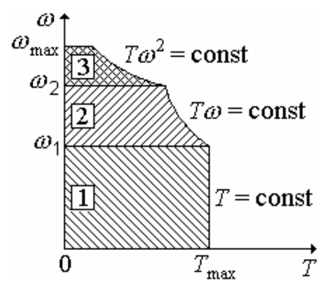


Fig. 8. Torque/speed zones of the SRM drive.

Using the MATLAB/SIMULINK software package some computer simulation models of electric drives with SRM have been developed. Detailed study of the drive system under consideration has been carried out for the dynamic and static regimes at various loading, disturbances and work conditions.

IV. Some Simulation Results

The current forming for one motor phase is illustrated in Fig. 9, where are represented the respective current and voltage waveforms.

Fig. 10 shows the phase current waveforms i_1, i_2, and i_3 obtained for reference current level $i_{ir} = 10\,\text{A}$. The load

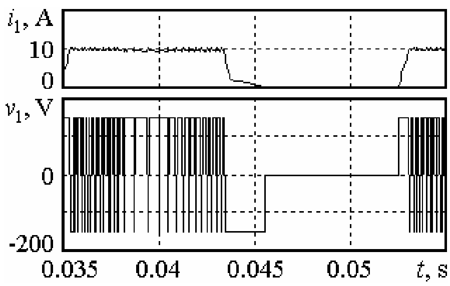


Fig. 9. Current and voltage waveforms for one motor phase.

torque applied to the motor shaft is $T_l = 8\,\text{Nm}$.

Fig. 11 shows current waveforms obtained in steady state regime at low and high motor speeds. In this case the load torque is $T_l = 6\,\text{Nm}$.

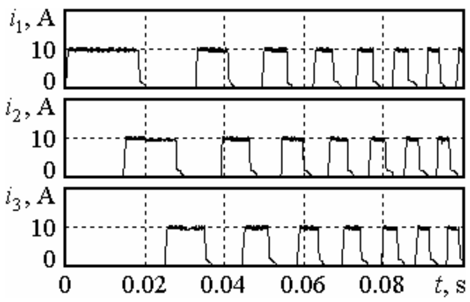
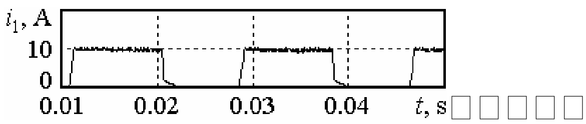
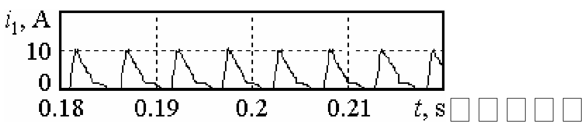


Fig. 10. Three-phase current waveforms.

The starting process and the motor speed stabilization is illustrated in Fig. 12, where the drive system reaction to dis-



a) at low motor speed



b) at high motor speed

Fig. 11. Current waveforms obtained for various motor speeds.

turbances is represented. The reference speed is $\omega_r = 314$ rad/s and the load changes applied sequentially to the shaft are $\pm\Delta T_l = \pm 0.25 T_l$. As evident, using a suitable speed controller, the drive static error is eliminated.

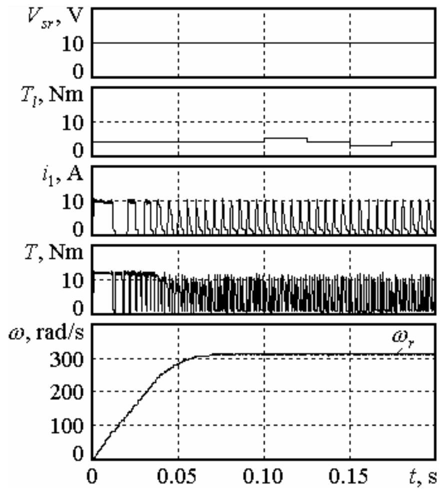


Fig. 12. Starting and motor speed stabilization.

V. Conclusion

The performance of a 6/4 three-phase SRM drive system with speed control is analyzed in this paper.

The investigations have been carried out through modeling and computer simulation for the respective transient and steady state regimes of operation.

The developed simulation models and the results obtained can be used for appropriately illustration in the process of teaching about such types of electric drive systems.

References

[1] T. J. E. Miller, *Switched reluctance motors and their control*, Oxford, Clarendon Press, 1993.

[2] T. J. E. Miller, Electronic control of switched reluctance machines, Oxford, "Newnes", 2001.

[3] R. Krishnan, *Switched reluctance motor drives*, Boca Raton, CRC Press, 2001.

[4] I. Boldea, S. A. Nasar, *Electric drives*, Boca Raton, CRC Press, 1999.

[5] A. Brosse, G. Henneberger, „Different models of the SRM in state space format for the sensorless control using a Kalman filter", *Power Electronics and Variable Speed Drives, Conference Proceedings,* pp. 269-274, London, UK, 1998.

[6] G. S. Buja, "Control characteristics of the SRM drives, Part II: Operation in the saturated region", *IEEE Transactions on Industrial Electronics,* vol. 41, no. 3, pp. 316-325, 1994.

[7] K. R. Tthompson, P. P. Acarnley, C. French, "Rotor position estimation in a switched reluctance drive using recursive least squares", *IEEE Transactions on Industrial Electronics,* vol. 47, no. 2, pp. 368-379, 2000.

[8] M. R. Mikhov, A. M. Avramov, R. N. Ognyanov, "Perspectives for implementation of switched reluctance motor drives". *Announcements of Union of Scientists*, vol. 3, pp. 160-162, Sliven, 2001.

[9] S. Cao, K. J. Tseng, „A new method for accurate analytical modeling of switched reluctance motor", *International Conference on Power Electronics Drives and Energy Systems for Industrial Growth*, *Conference Proceedings,* pp. 540-545, Perth, Australia, 1998.

[10] S. Mir, M. E. Elbuluk, I. Husain, Torque-ripple minimization in switched reluctance motors using adaptive fuzzy control, *IEEE Transactions on Industry Applications,* vol. 35, no. 2, pp. 461-468, 1999.

Slip Frequency Calculation Method in Speed Sensorless Induction Motor Drives

Nebojsa Mitrovic[1], Vojkan Kostic[2], Milutin Petronijevic[3]

Abstract – **Analysis of the PWM induction motor drive, where the regulation is made without the speed sensor, with the method of frequency compensation is presented in the paper. On the basis of the proposed method, compensation for the slip frequency and voltage drop compensation, being very important especially at the low speed range, is done. Slip frequency is determined on the basis of non-linear torque-speed dependence.**

Keyword – **Variable frequency, Speed control, Sensorless**

I. Introduction

Induction motor drive based on full digital control is being applied in a great number of industrial application, begining from low cost drive to high performance. Great part of scientific efforts, in the last few yers had a goal of eliminating the speed sensor, saving good static and dynamic performances like a solution with speed sensors. AC drives without speed sensor, according at methods applied for speed regulation can be clasified in two groups: 1) low cost drive of general purpose, 2) high performance drive.

In the first group of drive some of the following methodology foor speed calculation are used: slip frequency calculation method, constant volts-per-herts control, slot space harmonics, frequency compensation method [1].

In the second group following methods are applied: speed estimation, model reference adaptive method, speed observers, Kalman filter techniques, neural nerwork based estimator [1]. Most of these method for speed estimation is carried out from methods applied in vector control techniques which had a purpose of determination of "inaccessable" quantities (flux and torque).

All of these methods are based on procedure which comprehends measurement of electrical stator quantities, directly or in DC link of the inverter with phase current reconstruction performed switching function and knowing motor parameters. However, in practice, simple, cheap and reliable drive are often needed, where it is possible to control speed, with more modeset requests regarding dynamical features.

In this paper one of the slip compensation method is analysed. The proposed control algorithm is based on frequentcy calculation by using air gap power estimation and nonlinear relationship between slip frequency and air gap power. Besides frequency compensation, the stator resistance voltage drop compensation is realized.

[1]Nebojsa Mitrovic is with the Faculty of Electronic Engineering, A.Medvedeva 14, 18000 Nis,Serbia E-mail: nesa@elfak.ni.ac.yu

[2]Vojkan Kostic is with the Faculty of Electronic Engineering, A.Medvedeva 14, 18000 Nis, Serbia, E-mail: vkostic@elfak.ni.ac.yu

[3]Milutin Petronijevic is with the Faculty of Electronic Engineering, A.Medvedeva 14, 18000 Nis, Serbia, E-mail: milutin@elfak.ni.ac.yu

The proposed control scheme requires motor name-plate data, the stator resistance value and instataneous stator current in two phases. The basic characteristics of the drive with the applied algorithm are the simplicity of the practical realizetion, fair steady state and transient characteristics, and the price not higher than the open loop frequency controlled drive.

II. Frequency and Voltage Drop Compensation

The criterion for the stator frequency selection for a given load and the reference speed is illustrated in Fig. 1.

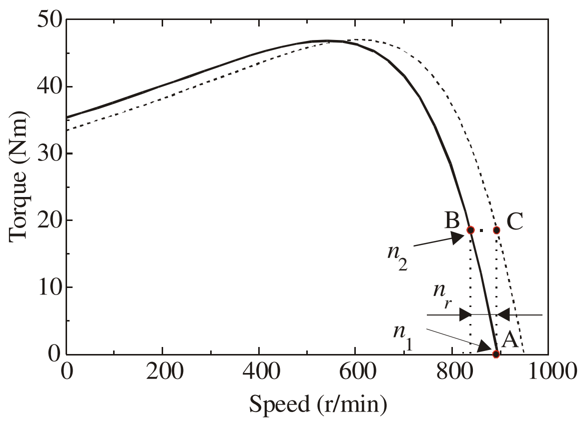


Fig. 1. The speed compensation principle

If the speed at no load was n_1 (point A), with the load increase operating point moves to the point B, with the speed "drop" of n_r. To compensate for the speed "drop", the cont–roller increases the speed to $n_2 + n_r = n_1$ by means of increasing the frequency, thus moving the operating point to C. In this way, the actual speed is again equal to the reference speed n_1 from the point A.

In order to implement this scheme it is necessary to know the relationship between load torque and slip. One of the possibilities is to assume a linear relationship between them [2]. Although this technique gives good results for high speeds its usefulness at low frequency and large torques is limited due to large steady state errors introduced by the linear approximation. To avoid above mentioned problem it is necessary to take account non-linear torque-speed characterristic of the motor [3,4].

The Electrical Machine Theory establishes the relationship for the electromagnetic torque (M_m) and break down torque (M_{pr}):

$$M_m = \frac{2M_{pr}}{s/s_{pr} + s_{pr}/s} \tag{1}$$

where s_{pr} is the slip at which the break down torque occurs. Eq. (1) is valid for any torque it is also valid for rated conditions. Defining M_{pr}=ν·M_n and using Eq. (1) yields:

$$s_{pr}/s_n = k = \nu + \sqrt{\nu^2 - 1} \qquad (2)$$

which is the break down slip in per unit of the rated slip (s_n).

Substituing Eq. (2) into Eq. (1), the slip (s) required to produce electromagnetic torque M_m can be writen as:

$$s = k \cdot \nu \cdot M_n \cdot \frac{s_n}{M_m}\left[1 - \sqrt{1 - \left(\frac{M_m}{\nu} \cdot M_n\right)^2}\right] \qquad (3)$$

For the practical reasons we need to eliminate the load torque from Eq. (3). The elimination can be done by using Eq. (4):

$$M_m = (p/4\pi)P_{ob}/(f_m + f_r) \qquad (4)$$

where p is the number of poles, P_{ob} is the air gap power, f_r is the slip frequency, and f_m is the frequency which corresponds ti the actual rotor speed.

Solving the Eqs. (3) and (4) for the slip frequency can be obtained:

$$f_r = \frac{1}{2 - a \cdot P_{ob}}\left(\sqrt{f_m^{\,2} + \frac{k \cdot s_l}{2 \cdot \text{ν}}P_{ob} - b \cdot P_{ob}^{\,2}} - f_m\right) \qquad (5)$$

$$a = p/(4\pi k \nu M_n s_n f_n); \qquad b = \left(p/(4\pi \nu M_n)\right)^2$$

When the ratio between the break down and rated torque is enough large, constants a and b become small and can be neglected. The physical explanation of this is that we have linear aproximation of mechanical motor characteristics, i.e the break down torque assume infinite value. For the linear dependence, in this case, for the slip frequency can be aproximately writen as:

$$f_r \approx 1/2\left(\sqrt{f_m^{\,2} + s_l \cdot P_{ob}} - f_m\right) \qquad (6)$$

where $s_l = p s_n f_n / \pi M_n$.

Air gap power (P_{ob}) can be obtained by using expresion:

$$P_{ob} = 3V_s I_s \cos\varphi - 3I_s^{\,2} r_s - P_{Fe} \qquad (7)$$

To determine power P_{ob} it is necessary to know rms value of induction motor stator currents and active component of stator current. Assuming the symetrical sysytem, the rms current value can be obtained by measuring instataneous phase current as:

$$I_s = \sqrt{2/3} \cdot \sqrt{i_{as}(i_{as} + i_{cs}) + i_{cs}^{\,2}} \qquad (8)$$

The active stator current components can be obtained by qd transformation in synchronous reference frame as follows:

$$I_{s(\mathrm{Re})} = \sqrt{3}\{i_{ac}\cos(\omega t - \pi/6) - i_{cs}\sin(\omega t)\} \qquad (9)$$

The last term in Eq. (7) under variable frequency operation is difficult to obtain but it can be approximated from the knowledge of rated values and constant flux operation. It can be easily shown that the core losses for nominal load:

$$P_{Fen} = P_{u\ln}\left(1 - \eta_n/(1 - s_n)\right) - 3I_{sn}^{\,2} r_s \qquad (10)$$

Core losses can be devided into two components [5]:

$$P_{Fen} = K_h B_n^{\,2} f_n + K_e B_n f_n^{\,2} \qquad (11)$$

where K_h i K_e are coeficient which depend of core type, B_n is rated flux density.

For constatnt flux operation these losses only vary with frequency. Assuming that at rated conditions both components are equal, after some manipulation, the total core loss at any frequency can be writen in terms of its rated value as:

$$P_{Fe} = \frac{1}{2}\left(\frac{1+s}{1+s_n}\left(\frac{f_s}{f_n}\right) + \frac{1+s^2}{1+s_n^{\,2}}\left(\frac{f_s}{f_n}\right)^2\right)P_{Fen} \qquad (12)$$

Substituting Eq. (12) into Eq. (7) gives air gap power as a function of reference frequency and measured variables. The slip measurement required in Eq. (12) is obtained from Eq. (5) or Eq. (6).

Voltage drop compensation can be obtained by keeping magnitude of the leakage stator flux at the constant and rated value. Based on the phasor diagram shown in Fig.2 can be writen:

$$V_s = I_s r_s \cos\varphi + \sqrt{(V_{s0} f_s / f_n)^2 - (I_s r_s \sin\varphi)^2} \qquad (13)$$

where with V_{s0} is marked amplitude of E_m at rated frequency f_n. The value for E_m needed for constant flux keeping (V/f=const.) at any another frequency f_s can be obtained as:

$$E_m = V_{s0} f_s / f_n.$$

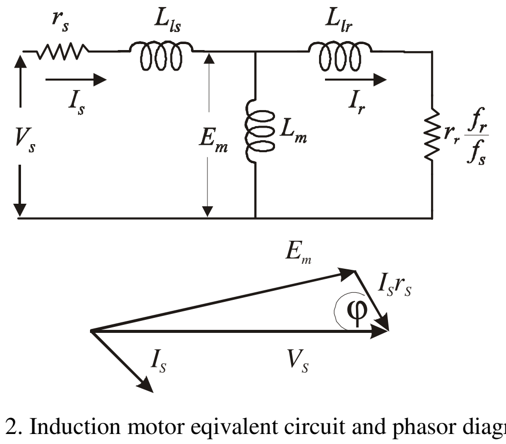


Fig. 2. Induction motor eqivalent circuit and phasor diagram

For implementation of equation (13) the value of rms stator current I_s is needed like its active components which can be obtained based on the equation (8) and (9). Final expression for stator voltage at any frequency and based of the instantaneous measurements of motor current is:

$$V_s = \frac{\sqrt{2}}{3} \cdot I_{s(\mathrm{Re})} r_s + \sqrt{\left(\frac{V_{s0} f}{f_n}\right)^2 + \frac{2}{9}\left(I_{s(\mathrm{Re})} r_s\right)^2 - \left(I_s r_s\right)^2} \quad (14)$$

III. Simulation and experimental results

In Fig.3 the block diagram of the drive is shown so the basic idea of applied speed regulation modus can be seen. Due to existing positive feedback at frequency compensation and voltage it is necessary to stabilize the system by using low-pass filter in the feedback loop.

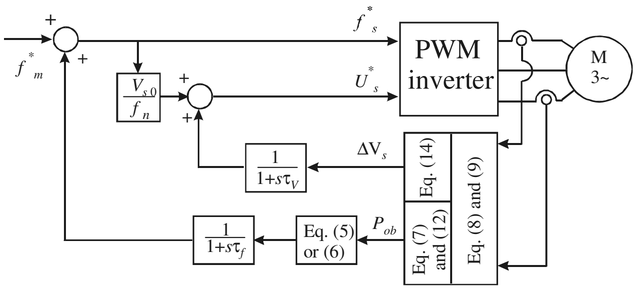


Fig. 3. Block diagram of the drive

The system is complex and nonlinear, therefore the dynamics were investigated on the corresponding model based on the fundamental component of the variables (large model) and the adequate linearized model [6,7]. The model that considers only the fundamental components of the varible was formed in Matlab/Simulink, after that the linearized model was developed by using incorporated linearization function.

Some results for different low-pass filter cut-off frequencies and for reference speed n_{ref} =400 r/min are shown on Fig. 4 and Fig. 5. The regions of instability are, in Fig.4, at low load and for filter cut-off frequency f_{cutoff}=150 Hz.

The analyses performed pointed out that the existence of the low-pass filter in the voltage and slip frequency feedback path has the crucial influence on the stability of the system. The filter cut-off frequency was chosen to be as low as the characteristic frequencies in the mechanical subsystem transients. This ensures a stable operation and good dynamic performance. The proposed speed control method is after that analysed by computer simulation on a detailed drive model, and than on the experimental results obtained on laboratory model.

Simulation results are obtained by using Matlab program package. In Fig. 6 simulation results obtained for reference speeds of 200, 600 and 1000 r/min are shown. First, drive system operated at the no-load while load torque in stedy state was 50% of rated torque. The speed regulation effect with and without speed regulation is illustrated in Fig.6. In it is clearly marked the moment of switching off regulation resulting in speed reduction. Dependence of the load torque versus time is also shown in figure.

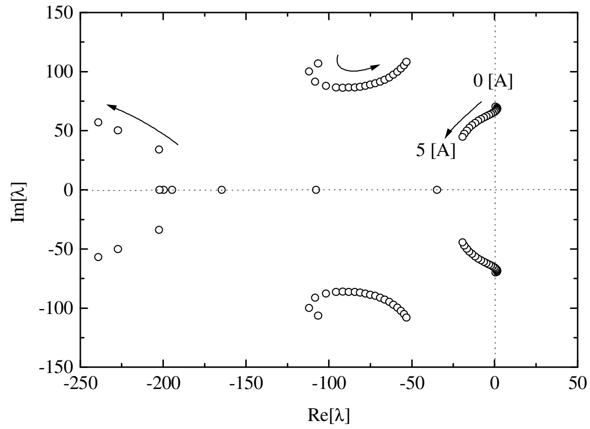


Fig. 4. Drift of eigenvalues for n_{ref}=400 r/min, f_{cutoff}=150 Hz

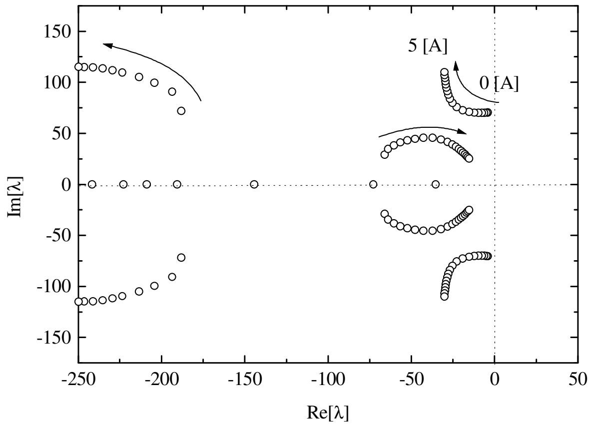


Fig. 5. Drift of eigenvalues for n_{ref}= 400 r/min, f_{cutoff}= 50 Hz

Experimental verification of the proposed algorithm were made on the induction motor drive; its parameters are given in appendix A.

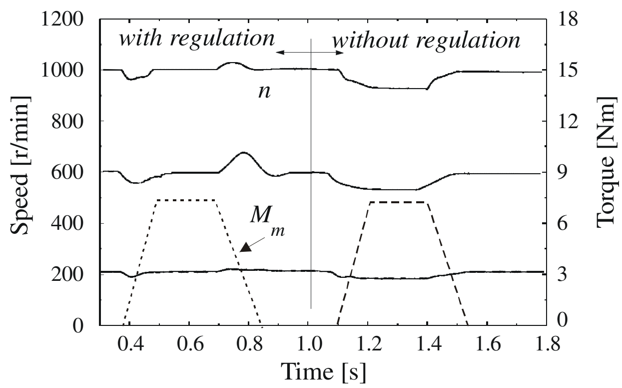


Fig. 6. Speed response. Simulation results.
(n_{ref}=200; 600; 1000 r/min)

Motor load was performed by the separately excited DC motor in dynamic braking regime. In Fig. 7 and Fig. 8 steady state speed-torque characteristics for different reference speed values in range of 100-1800 r/min are shown. The drive ability for keeping reference speed even at a very low speed is obvious.

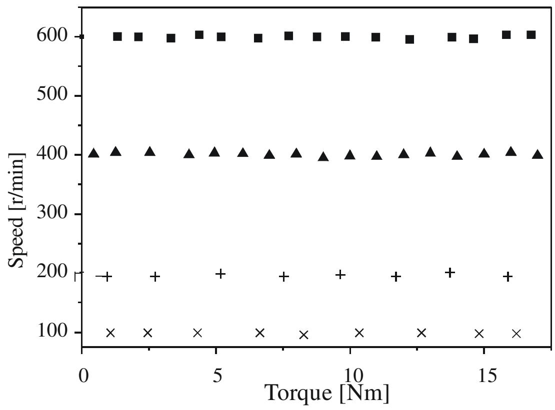


Fig. 7. Steady state speed response. Experimental results (n_{ref}=100 ÷ 600 r/min)

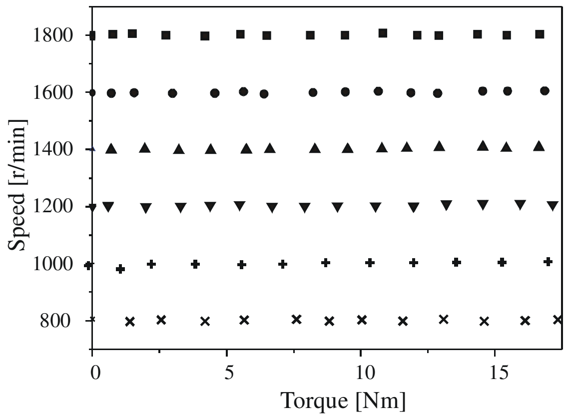


Fig. 8. Steady state speed response. Experimental results (n_{ref}=800 ÷ 1800 r/min)

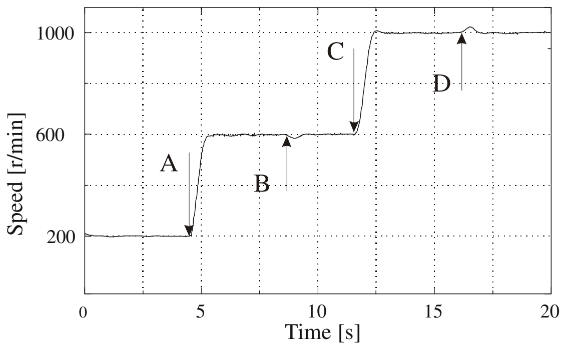


Fig. 9. Speed response. step change of reference speed and load. Experimentla results

The dynamic characteristics of the drive are presented in Fig.9 showing the response of the system to the change of reference speed and load change. First, the drive operated at no load at the speed of 200 r/min. At the moment related to point A the change of reference speed at 600 r/min had been done, so that the drive at the moment related to the point B was loaded. Afterwords at the point C the reference speed of 1000 r/min was changed, and at point D no-load resulted.

IV. Conclusion

Induction motor drive frequency compensation method based on *V/f* control has been presented. The proposed compensation method requires the knowledge motor nominal data, stator resistance and motor stator currents measurements. The control algorithm can be simply implemented in the existing drives which use the clasic *V/f* control. The theoretical and experimental analysis, the applicability of the frequency compensation method for drives with modest dynamic characteristics has been confirmed.

Appendix A

Induction machine data:

1,5 kW; 50Hz; 930 r/min; 220V; 4A; cosφ=0,8.

r_s=4Ω; r_r=3,4Ω; L_{ls}=0,01383H; L_{lr}=0,01383H; L_m=0,245 H

References

[1] K. Rajashekara, A. Kawamara, K. Matsuse, *Sensorless Control of AC Drives*, IEEE Press, 1996.

[2] A. Garcia, T. Lipo, D. Novotny “A New Induction Motor Open-Loop Speed Control Capable of Low Frequency Operation” *Annual Meeting*, New Orleans, Oct.1997. pp. 121-128.

[3] Nebojša Mitrović, “Regulacija brzine asinhronog pogona sa PWM invertorom kompenzacijom učestanosti bez senzora brzine”, *Doktorska Disertacija*, Elektronski fakultet, Niš 1998,

[4] B. Jeftenić, N. Mitrović, M. Bebić, “A Simple Speed Sensorless Control For Variable Frequency Induction Motor Drive”, *IEEE Trans. on Energy Conversion*, Vol.14, No.3, Sept.1999.

[5] D. Novotny, S. Nasar, B. Jeftenic, “Frequency Dependence of Time Harmonic Losses in Induction Machines”, *ICEM*, Boston, July 1990.

[6] Paul C. Karause, Thomas Lipo, “Analysis and Simplified Represenatations of a Rectifier-Inverter Induction Motor Drive” *IEEE Trans. App. Syst.*, vol. PAS-88, No.5, May 1969, pp. 588-596.

[7] M. M. Ahmed, J. A. Taufiq, C. J. Goodman, M. Lockwood, “Electrical Instability in a Volatage Source Inverter-Fed Induction Motor Drive at Constant Speed”, *IEE Proc.*, Vol. 133, Pt. B, No. 4, July 1986

Pseudorandom Encoder's Development Using Virtual Instrumentation

Dragan B. Denić[1], Miodrag Z. Arsić[2], Ivana S. Ranđelović[3], Goran S. Miljković[4]

Abstract – **A pseudorandom encoder as a new type of the absolute encoder with one code track coded by applying pseudorandom binary sequences together with new possibilities in its development which offers virtual instrumentation is considered in this paper. Here, a method for parallel reading of pseudorandom code using photodetector array is proposed. A problem of zero position adjustment at encoder installation is also considered and a concrete solution in accordance with requests of high technologies encoders is proposed. Then, two programs realized with software package LabVIEW 7.1 are presented and their front panels are shown. The first program simulates a code reading in encoder using sensor heads, while the second program simulates an electronic block of encoder operating. In the paper is pointed out advantage of computer using as auxiliary means in developing of optimal pseudorandom encoder solution.**

Keywords – **Position measurement, Pseudorandom encoder, Pseudorandom binary sequence, Virtual instrument, Zero position adjustment, LabVIEW.**

I. Introduction

The development of information technologies caused that virtual instruments are important part of measurement instruments. The possibility to customize the appearance and functional range of virtual instrument enables designers to apply complex procedures of their testing, [11].

The term Virtual Instrument (VI) is often used, but there is no official definition of its meaning. The truly virtual instruments (the software model of instrument) are designed and realized relatively seldom. They are usually used for an educational purpose only, or for modeling of some properties of real instruments, [11].

Generally, VI is usually a real measuring system based on a host PC, where the PC is used for set – up and control of measurement, data acquisition, data processing and displaying results. Both systems based on measuring separated modules and systems based on stand – alone measuring instruments controlled by PC or their conjunction are such designated. Advantages and disadvantages of such virtual instruments should be judged according to their application [11].

[1] Dragan B. Denić is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia and Montenegro, E-mail: ddenic@elfak.ni.ac.yu

[2] Miodrag Z. Arsić is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia and Montenegro, E-mail: marsic@elfak.ni.ac.yu

[3] Ivana S. Ranđelović is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia and Montenegro, E-mail: rivana@elfak.ni.ac.yu

[4] Goran S. Miljković is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia and Montenegro, E-mail: goranm@elfak.ni.ac.yu

The developments in encoder technology are consequently based on the permanent technical progress. As a main part of all control systems for positioning, absolute encoders provide measured information about sensor head position related to the measurement scale. Because each position is coded, the current position is defined apart from the previous position. This is the basic quality of the absolute encoders and hence proceeds their main feature that after the power is turned on, an information about the current position of movable system is instantly obtained.

The pseudorandom positional encoders are latest development trend in position measurement new methods at industrial movable systems. Pseudorandom encoders' development requests producing of code tracks which is very unpractical because during research in developing series of solutions different code reading methods are applied, [3, 4, 5, 6]. Each of applied methods uses a different number of reading heads, so computer have a great part in analysis of signals from sensor heads, by which expensive realizations of experimental systems would be evaded. In the paper computer is applied as auxiliary means for movable system simulation that is for measurement signals getting from sensor heads.

II. Parallel Code Reading of Pseudorandom Code and Its Usage in Position Encoders

The method of pseudorandom coding, which requires only one code track for absolute position determination, represents an attractive alternative to the classic measurement method. Its advantages are significant in the case of high-resolution position encoders and linear position encoders with very long code tracks. Coding is based on the "window property" of pseudorandom binary sequence (PRBS) $\{S(p) / p = 0, 1, ..., 2^n-2\}$. According to this, any n-bit code word $\{S(p+n-k) / k = n,...,1\}$ obtained by a window of width n $\{x(k) / k = n,...,1\}$ scanning the PRBS, is unique and may fully identify window's absolute position p relative to the beginning of sequence, [1, 2, 3, 4].

Pseudorandom binary sequences (PRBS) are long known, and in the field of telecommunication theory are used for finding the scope, scrambling, error detection, modulation, synchronization, etc. They are generated with a shift register of length n and a corresponding feedback. With the right choice of that feedback, the PRBS of maximal length $2^n -1$ are obtained, which are also known in literature as PN sequences or M-sequences, [8].

For realizing of true pseudorandom absolute encoder, it is needed to apply any kind of parallel code reading method and in order to avoid a need for initial movement. A usage of n particular detectors is unacceptable to encoders with high measurement resolution. It is problem to dispose those n

detectors on such little space [7, 8, 9, 10]. One of the possibilities is applying of integrated photodetector arrays for pseudorandom code reading. Integrated photodetector arrays are available on the market with different intervals between photodetectors. Those intervals are 13μm, 10μm, 7μm and smaller. It should use a large number of photodetectors in order to increase absolute position measurement precision [7, 8, 9, 10].

III. NEW ALGORITHM WITH REDUCED PROCESSING TIME

For entirely clearing of operation principle of pseudorandom positional encoder with parallel code reading, an algorithm of pseudorandom encoder electronic block is shown in Fig. 1a, which is detailed described in [8, 10].

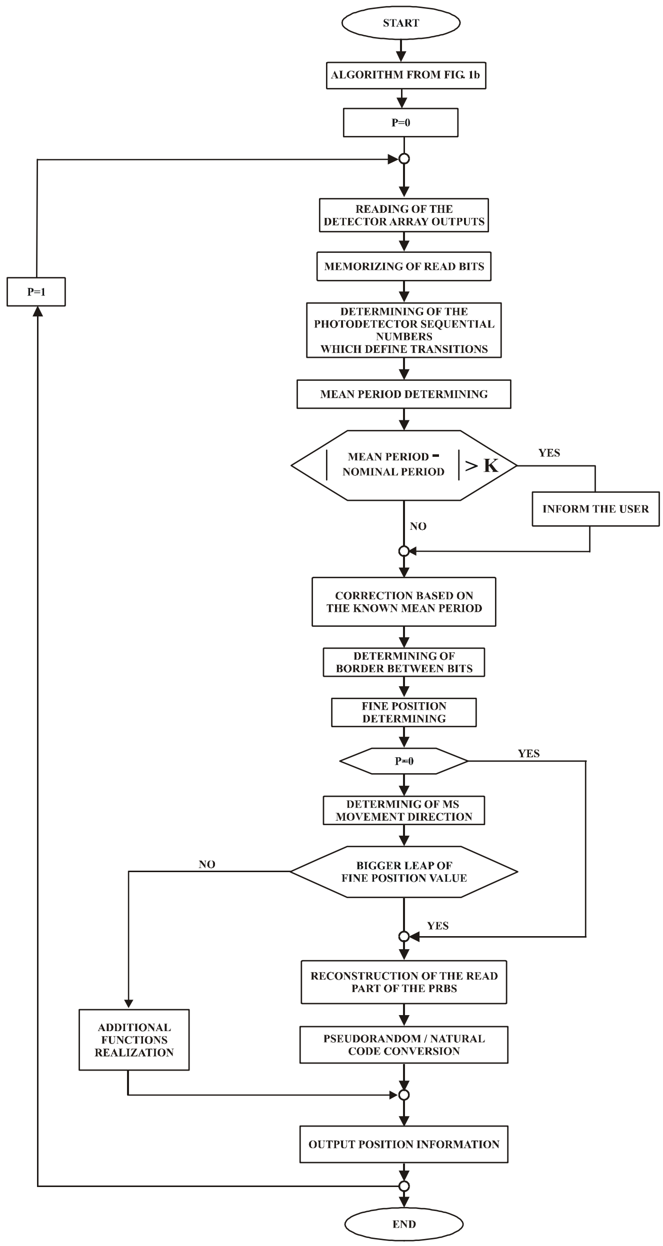


Fig. 1a. Algorithm of the proposed solution

This algorithm is product of many testing algorithms presented in references [8, 11]. For momentary position determining of movable system for case of code reading with linear photodetector array, first have to define method of rough and fine position determining. The algorithm shown in Fig. 1a defines fine position determining based on knowing sequential numbers of photodetectors which define transitions, then defines a moment of rough position determining so as method of movable system direction determining based on fine positions changes. It is easy come to information about movable system momentary position if is information about rough and fine position known. Because errors are occurred in code reading, algorithm defines method of error correction based on mean period calculating. When error is too much big, user is informed about that. In algorithm one pseudorandom to natural code converter is applied.

IV. DETERMINING OF MOMENTARY POSITION WITH DIRECT ADJUSTED ZERO POSITION

It comes to the unmatchings in the starting positions when installs an absolute encoder on movable system shaft. One of the leader factory in the world for production of new encoder generation, Stegmann, presented on the market a great choice of the incremental and absolute encoders with different mechanical interface, resolution and with the new electronic features which are in the scope of the regular standards, [11,12]. If at installation of encoder CA6S on shaft of the movable system exists an umatching in the starting positions, directly adjusting of the starting position is done by pressing a taster. More precisely, encoder installation is done and then movable system is placed in own zero position. After this by simple pressing on taster, output position information of encoder is put at value zero. With respect to mode of classic encoder functioning, this solution is probably based on correction parameters memorizing which figurate in main algorithm.

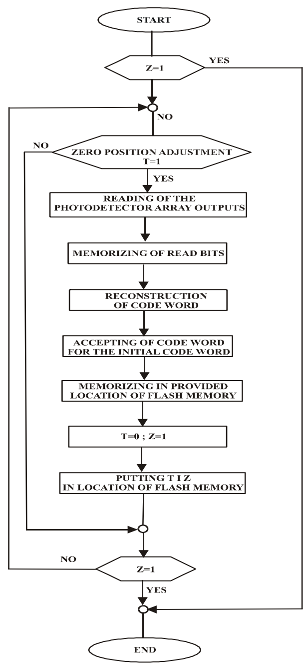


Fig. 1b. Algorithm of the proposed solution

Considering that direct zero adjustment will be soon a general request for any encoder, here is discussed a possibility of realizing such function at absolute pseudorandom encoder which is suggested here. By pressing the pushbutton at zero position of movable system, current output position information is accepted as a correction parameter and automatically a value of output position information becomes zero. Algorithm from Fig. 1a would be modified so as directly behind block “current position information“ would be a block “position information correction“. Of course, instead a first block (“algorithm from Fig.1b”) would be block which represents accomplishing of defining function above mentioned correction parameter and it’s storing in determined location of flash memory. It should know that a solution of this function requests execution of some arithmetical operations, for which unfortunately significant time is needed. Of course, this is in opposite with basic concept in this paper, and that is maximal possible decreasing of execution time of whole algorithm, in order to commercial real absolute pseudorandom encoder realization, [11].

The direct zero adjustment procedure itself is shown in algorithm, Fig. 1b. Proposed solution is based on fact that direct zero position adjustment is performed only one time at encoder installing. In algorithm from Fig. 1b, a direct zero position adjustment is done for parameter values “Z=0“, while “T=1“means that pushbutton is pressed. Parameters T and Z are put in flash memory of encoder itself, and then permanently storing of these parameters is done. That means at encoder restarting parameters T and Z stay memorized in flash memory of encoder itself so in all situations skip algorithm from Fig.1b. Therewith measurement time doesn’t increase. If direct zero position adjustment isn’t performed then pushbutton is pressed (“T=1“), reading of the detector array outputs is performed, adoption of code word for initial code word. This code word is stored in flash memory and also parameters “Z=1“, and “T=0“. When direct zero position adjustment is done then proceeds to steps executing which are presented with algorithm in Fig. 1a. The additional operation is eliminated by this, that is correction of each read value, [11]. After storing of code word in flash memory, which is stored as initial code word, code conversion pseudorandom/ natural is performed, according to the algorithm which is shown in Fig. 1a and 1b. There aren’t modifications in algorithm functioning. The first block whose execution function is presented in Fig. 1b is performed only one time during direct zero adjustment procedure. In followed permanent algorithm functioning this block doesn’t more in function and according to this doesn’t influence to algorithm functioning. Such solution which, according to this description, looks as ideal approach is only one more proof of great possibilities of pseudorandom encoders in relation to many requests which will become standard in future, [11].

At the end, two programs are presented realized by using software package LabVIEW 7.1, [13], which simulate a reading from sensor heads at pseudorandom positional encoder, so as electronic block operating of such encoder, and their front panels are shown in Fig. 2 and 3, respectively. The simulation begins with program starting which simulates code reading on first computer. Within this program a route of system moving is proposed, PSBS type that is its length, number of photodetectors which read a code, so as a moment by which is initial state defined, and at the end output information is stored in database which is later used by other program. An electronic block software realization of tested pseudorandom positional encoder represents the second program which accepts information from database and gives series of positions on output. This program determines rough and fine positions and then based on that determines output positions so as information about correctness of positions and about moving direction. A selection of zero position from which code conversion would be done later is achieved by pressing the ‘STOP’ taster.

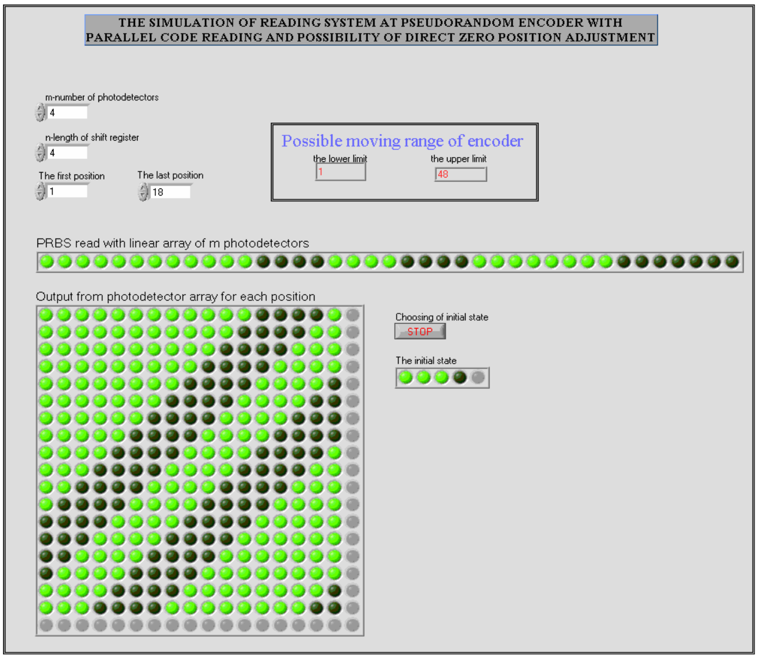


Fig. 2. Front panel of simulated reading system at pseudorandom encoder with parallel code reading and possibility of direct zero adjustment

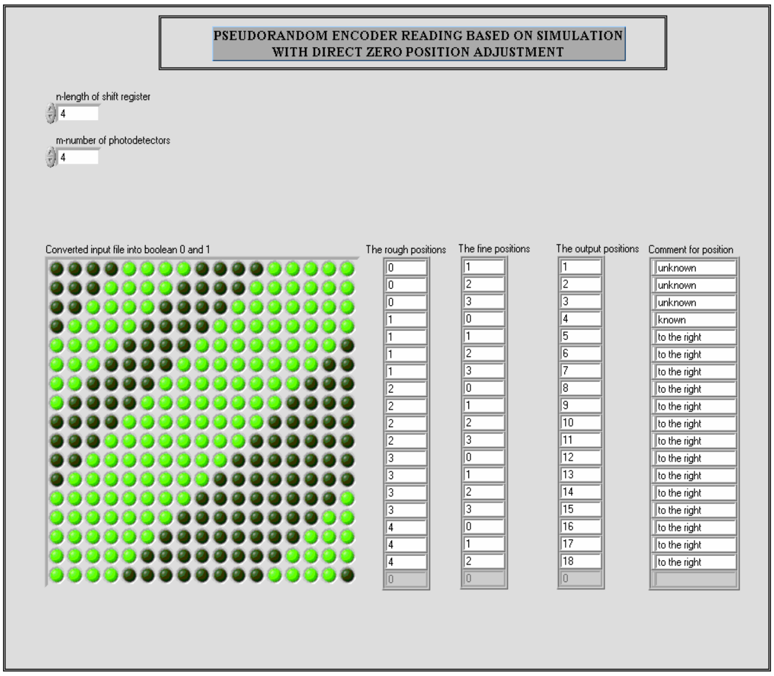


Fig. 3. Front panel of pseudorandom encoder with direct zero adjustment electronic block simulation

The output positions are decimal values which are obtained by summing of rough and fine positions binary values. Within developed solution one example for pseudorandom sequence is given with needed number of bits $n = 4$, number of photodetectors per one code bit $m = 4$, and proposed moving route is $(1-18)$, and read array of movable system momentary output positions is given in Fig. 3, in output list form, which of course can be printed.

V. Conclusion

Computer integrating within industry and other fields is caused by growth of electronics and microcomputer technology. In paper is shown advantage of computer using as auxiliary device in developing of optimal pseudorandom encoder solution. In the first part of the paper pseudorandom encoder with parallel code reading operating principles are explained, and then is presented a simulation of the particular parts of that encoder realized by software package LabVIEW 7.1. Simultaneously, a solution is suggested which on simple way enables direct zero position adjusting, at encoder mounting. After taster is pressed, written code word itself is adopted for started, or zero position and store in flash memory of encoder itself. Read code word is stored as initial code word and in relation to it code conversion pseudorandom/ natural is performed. On this way, anyone additional operation is eliminated, i.e. correction of each value of measured position, this is only possible way at classical absolute encoders.

References

[1] M. Arsić, D. Denić, "Position measurements based upon the application of pseudorandom encoding", *Facta Universitatis, Ser. Electronics and Energetics*, no. 6, pp. 13-23, 1993.

[2] H. Khalfallah, E.M. Petriu, F.C.A. Groen, "Visual position recovery for an automated guided vehicle", *IEEE Trans. Instrum. and Meas.*, vol. 41, no. 6, pp. 906-910, 1992.

[3] M. Arsić, D. Denić, "Code convertor pseudorandom/natural applied at position encoders", *ETRAN, Ser. Elektronika*, pp. 164-167, 1995.

[4] D. Denić, M. Pešić, M. Arsić, "New algorithm of code conversion pseudorandom/natural", *Proceedings ETAN,* volume II, pp. 39-44, Belgrade, 1993.

[5] E.M. Petriu, J.S. Basran, "On the position measurement of automated guided vehicles using pseudorandom encoding", *IEEE Trans. Instrum. and Meas.*, vol. 38, no. 3, pp. 799-803, 1989.

[6] M. Arsić, D. Denić, "New pseudorandom code reading method applied to position encoders", *Electronic letters* no. 10, vol. 29, pp. 893-894, 1993.

[7] D. Denić, I. Ranđelović, M. Arsić, "Absolute position measurement using the method of pseudorandom code parallel reading", *ICEST 2003 Conference Proceedings,* pp. 394-397, Sofia, Bulgaria, 2003.

[8] D. Denić, I. Ranđelović, M. Rančić, "High-resolution pseudorandom encoder with parallel code reading", *ISSN 1392-1215 Electronics and electrical engineering* Nr 7(56), pp. 9-14, Kaunas, Litvania, 2004.

[9] D. Denić, I. Ranđelović, J. Đorđević, G. Miljković, "New type of linear and angular displacement transducer based on pseudorandom encoding", *ICEST 2005 Conference Proceedings,* pp. 394-397, Niš, Serbia and Montenegro, 2005.

[10] D. Denić, I. Ranđelović, "New type of position encoder with possibility of direct zero position adjustment", *IDAACS 2005 Conference Proceedings,* pp. 299-305, Sofia, Bulgaria, 2005.

[11] M. Florezyk, W. Winiecki, "The parametric method for functional testing of virtual instruments", *IDAACS 2005 Conference Proceedings,* pp. 310-315, Sofia, Bulgaria, 2005.

[12] http://www.stegmann.com

[13] LabVIEW, User's Manual, National Corporation, USA, 2004

Method for Determining Coordinates of a Solid Body with Six Degrees of Freedom

Milena M. Stefanova[1]

***Abstract* – A method for determining of the coordinates of moving solid body with six degrees of freedom is described. The method could be applied in examining of behavior of moving solid body, in exception of ship model in maneuvering pool.**

***Keywords* – models, method for determining, trajectory, measuring.**

I. INTRODUCTION

A lot of optical technologies that can provide information about the movement of a body have been developed. These technologies have common mathematics frame concerning the optic of Fourier and the theory of signal processing. Historically these technologies have been developed separately from each other, because of the reason that they do not have common elements that are tie-pieces.

The above mentioned optical methods have a wide application in different areas – in production processes, in determining of contours of surfaces, in non destructive testing, in solving dynamic problems (analysis of vibrational parts), analysis of random forms etc.

The following method for determining the coordinates of moving body with six degrees of freedom is the subject of our research in finding suitable method for exploring the behavior of ship models.

II. ESSENCE OF THE METHOD

The purpose of the described method is determining of the moving of ship radio-controlled model in maneuvering pool with the following dimensions 40x40m.

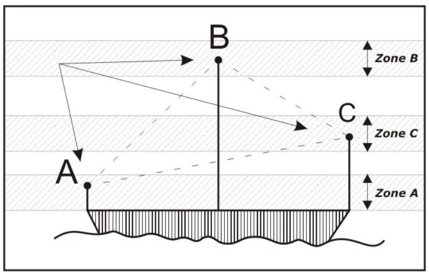


Fig. 1. Distribution of the light sources[1]

On two contiguous corners of the maneuvering pool are mounted cameras which trace the movement of three bright spot light sources (small electrical bulbs or light emitting diodes) attached to the model, see Fig. 2. The light sources that are used in the method are mounted on different heights, so that allows them to be presented in different areas on the camera screen in vertical direction. The transposition of the light sources on different levels, as shown on Fig. 1, allows exploration of the six degrees of freedom of a solid body. Each of the cameras is connected with tracking device, which automatically directs the optical axis of the camera to the object that is traced through the process of its movement.

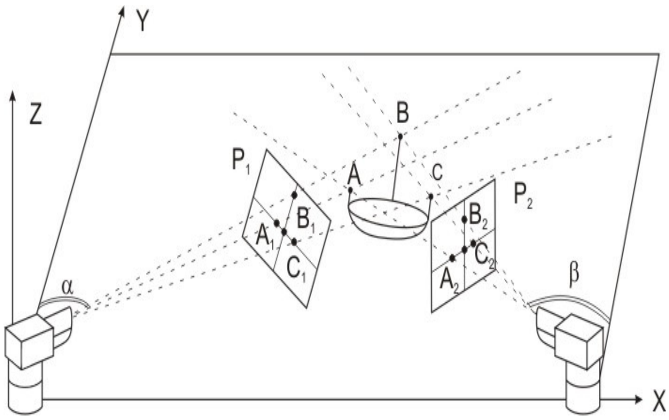


Fig. 2. Scheme of measuring system

Both of the cameras use pan-tilt mounts, as a part of a servo system containing direct current reversive engine, with rotary optical encoder between the camera and the platform.

At the time of working this method out, the revolving Moiré gauge is used for determining of the angle of rotation of the camera. It produces a number of electrical pulses when the camera is rotated in horizontal plain. These pulses in form of binary information are passed to a computer in order to make a database. Except this information the information passed to the computer contains data about the position of the three dot spot light sources in the field of vision of each of the cameras. Their position in vertical direction is given by counting the television lines considered from falling edge of vertical blanking pulse, Fig. 4 and Fig. 5. The horizontal position of the object in the field of vision of the camera is given by digitizing of the lines and taking into consideration the position of the light sources towards the falling edge of vertical blanking pulse.

The binary information packs incoming from both of the cameras, are stored into the computer memory with the purpose for further processing.

[1] Milena M. Stefanova – Faculty of Electronics,
Techincal University, 9010 Varna, Studentska str. 1
e-mail: mi1ena@abv.bg

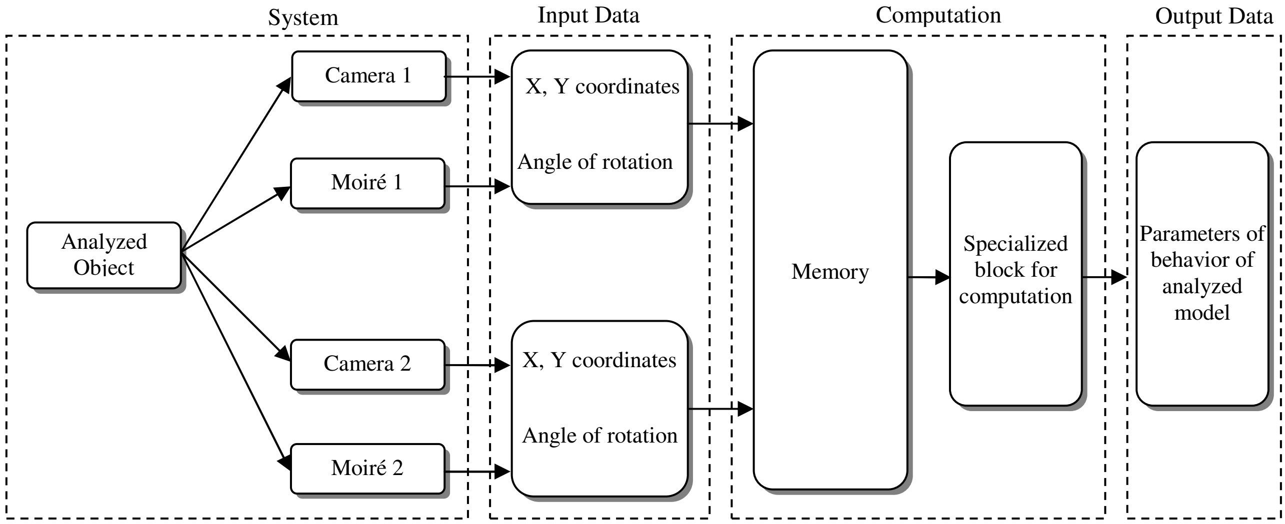


Fig. 3. Basic scheme of the method

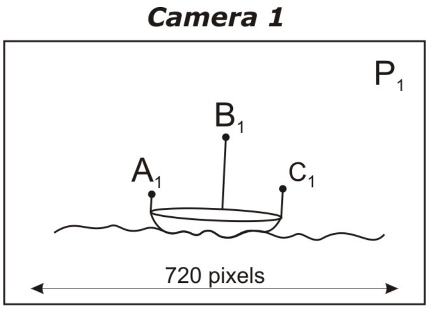


Fig. 4. Picture taken from camera 1

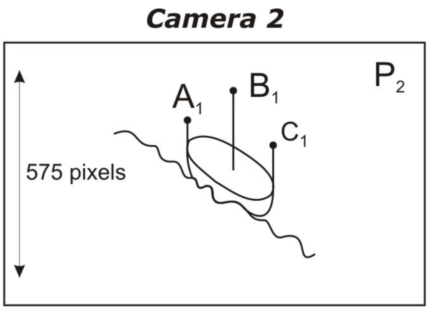


Fig. 5. Picture taken from camera 2

The main idea of the given method lies on a basic principle – if we know the coordinates of three points of a solid body, which means that we have an accurate definition of its exact position and orientation.

The position of the ship model is calculated using the two angles of rotation α and β. The orientation of the model is given by the coordinates of the three dot spot light sources upon the television screen.

On Fig. 3 is visualized the basic stages of the method. The System group provides the data needed for the input of computation stage, where are gathered all of the input data. The specialized block for computation makes all of the needed analysis and computation, all of the smooth filtering and interpolation works.

III. Determining of the key parameters

1. Determining of the angle of rotation of the Moiré gauges.

The angles of rotation α and β are determined by the rotation of the Moiré gauges shown on Fig. 6.

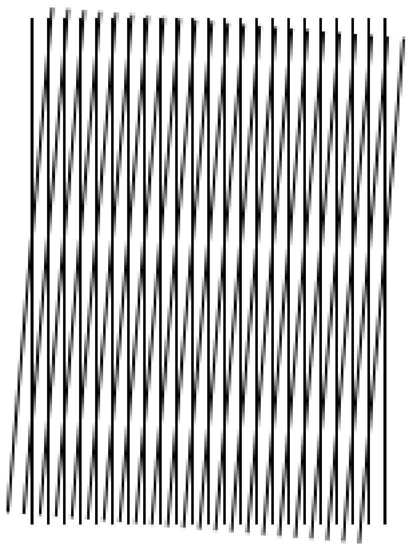


Fig. 6. Gratings of Moiré gauge

Moiré gauge consists of two identical gratings having the following parameters: l – constant of the gratings; γ – raster period angle between gratings. When the two gratings are absolutely parallel they no fringe is observed, only dark and light fields can be seen. When there is a small angle between the gratings there is a moiré pattern. The distance T between them is given by the following formula:

$$T = \frac{l}{2\sin\frac{\gamma}{2}} \tag{1}$$

The grating lines displacement on the width l cause perpendicular transposition of the moiré pattern at the distance T. For small angles $\gamma = 1'$ the transposition is given by the following equation:

$$T = \frac{l}{\gamma} \tag{2}$$

This principle is used in the pulse measurement of angle rotation.

In this method there are two cameras mounted on two neighboring corners of the pool. The optical axes of these two cameras using an automatic tracking system are always focu–sed on the tracing the object. The readings from both cameras are always time synchronized with each other by time code for obtaining correct data from them at the same time.

In this way the location of the object is determined by two angles of rotation. The two cameras have initial synchro–nization and run from 0˚ to 90˚ with pitch of 1'.

2. Determining the position of the object in the field of vision of the cameras.

The X and Y coordinates of the analyzed object are determined by using the formed raster image from the both cameras Fig. 7. Each camera is characterized by its parameters of resolution. In this case is used the standard with 625 lines (575 active).

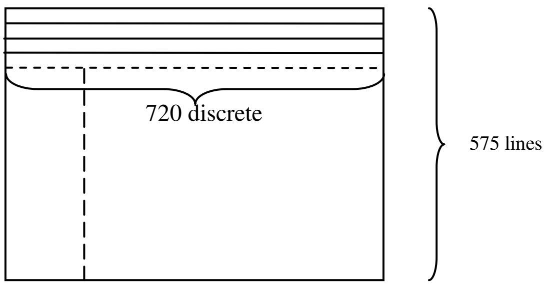


Fig. 7. Scheme of TV image

The Y coordinate is determined by the number of the line on which is situated the object (0 -575). The X coordinate is determined by the number of the discrete of the corresponding line on which is situated the object (0 - 720).

IV. Accuracy of the Method

The accuracy of the method exclusively depends on accuracy of Moiré gauges that have been used. To determine accuracy the length of the chord of maximal distance between the camera and the model is computed. With dimensions 40x40 m of the maneuver pool and rotation angle of 1' the length of the chord with maximal distance is 0,016 m, which satisfies the requirements of the practice Fig. 8.

The choice of the camera lens must be conformable to the requirement of the measured accuracy of X and Y coordinates. The camera optics is chosen so that it can comply with following conditions:

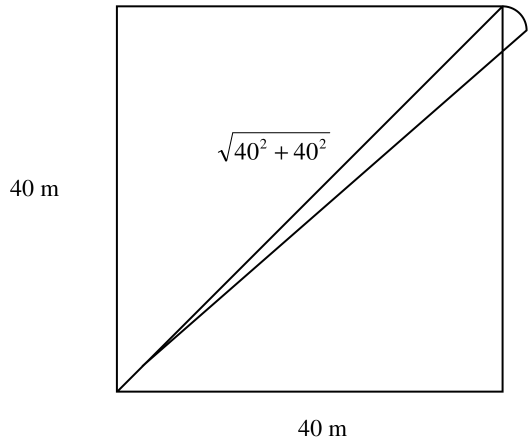


Fig. 8. Maximal distance scheme

- the model to be in range when it is too close;
- not to be so wide-angle, because from its side this will decrease the accuracy of measurement of the coordinates of the light sources.

The Moiré gauges should meet the technical requirements of at least 22 000 pulses/round. That will ensure the accuracy of the measurements as it is necessary 21 600 pulses/round.

V. Necessary Amount of Memory and Speed Performance

For the method is important to specify necessary amount of memory according to previously defined accuracy of measuring. In this case the position of the object in the maneuver pool is considered to be determined with accuracy of 0,016 m. The average speed of the model is considered 12km/h. The amount of memory is calculated for the following data components:

- Angle of rotation of the camera. The measurements of the rotation angle are with accuracy of 1'. The camera motion angle is 90˚ and that defines 5400 possible readings, witch on its side defines using of 13 bit binary number;
- X co-ordinate of the object. The co-ordinate has a value in rage 0 - 720, which necessitates using of 10 bit binary number;
- Y co-ordinate of the object. The co-ordinate has a value in rage 0 – 575, which necessitates using of 10 bit binary number.

The presence of three light sources, witch are defining the position of the object, necessitates for each one of them two readings for X and Y coordinates. The computations that have been made are pointing that for one obtained reading from each camera are necessary:

- 13 bits for rotation angle of the camera and 3 x 20 = 60 bits for X and Y coordinates of the three light sources;
- total of 73 bits for one surveying camera and 146 bits for two.

The necessitated requirements for accuracy of measurements up to 0,016 m at 12km/h average speed require data transfer speed of 17.82 Kbytes/s. For one minute object tracking is required the following amount of memory: amount of memory for 1 min = 60 x 17.82 Kbytes/s = 1069.2 Kbytes = 1.04 Mbytes.

VI. Conclusion

1. A method for determining of the position of moving object in continuous moments in time is described.
2. An assessment of necessary data transfer speed with an eye on obtaining necessary accuracy is accomplished.
3. So obtained values of the parameters empower for calculation and determining of the trajectory of movement, instant and average speed, and acceleration of analyzed object.
4. The necessary amount of memory for storing data obtained during analysis is defined.
5. Through database values and appropriate software the complete object's behavior could be visualized during the process of analysis.

References

[1] Craig, J.J., Introduction to Robotics – Mechanics and control, Addison, Wesley Publishing Company, 1991.
[2] Ján Bartl, Roman Fíra, Miroslav Hain, Inspection of surface by the moirè method, Measurement Science Review, vol. 1, number 1, 2001.
[3] Kafri O., Glatt I., The physics of moirè metrology, New York, Toronto, John Wiley & Sons.
[4] http://www.bshc.bg

Web Based Application for Distributed Remote Measurement Viewing

Ivan Stankov[1] and Grisha Spasov[2]

Abstract – **The paper discusses web based application for distributed automation. Realization is made over three-layer distributed model. XML table driven communication model is used for heterogeneous connection of different parts of the system. Functionality of the model is delegated and distributed among servers and embedded systems. Majority features of realization, concerning scalability, flexibility, distribution, collecting and delegating of functionality, reliability and security of the applications over heterogeneous entities are discussed.**

Keywords – **distribution, automation, web application, delegation, scalability, heterogeneous systems.**

I. INTRODUCTION

Recent years automation systems became more complex and wide spread with various applications due to ubiquitous using of communication technologies and especially Internet [3]. This led IT market to really huge growth and increasing familiarity with devices as pocket PCs, PDAs, Laptops, 3G Mobile phones, as well as PCs and PC compatible machines and concerned technologies. As a consequence of increased functionality, because of purposes laid in front of web based systems, is their complexity. For simplifying the process of monitoring and observation, new models for communication with system layer separation is assumed.

A realization of such kind of system takes into consideration with the particular aims of the end user. The most appropriate way for fast and accurate apprehension of environment's parameters is table driven and diagram approach. That way of data representation allows user for precisely observation of watch data.

In distributed information systems, based on World Wide Web, two technologies ASP (Active Server Pages) [11] and JSP (Java Server Pages) [10] are mostly used. These are current trends in web pages used over cyber space powered relevantly by *Microsoft* and *Sun*. Each of them has its features and advantages. The model of proposed realization is based on *ASP.NET*.

The entire automated process is distributed among embedded systems increasing in that way availability. This approach provides scalability, flexibility and avoids a crush of the whole system.

[1] Ivan Stankov – is with the Faculty of Electronics and Automatics in Technical University Sofia, branch Plovdiv, 4000 Plovdiv, Bulgaria, e-mail: istankov@tu-plovdiv.bg

[2] Grisha Spasov – is deputy dean of Faculty of Electronics and Automatics, Technical University Sofia, branch Plovdiv, 4000 Plovdiv, Bulgaria, e-mail: gvs@tu-plovdiv.bg

II. RELATED WORK

Distributed Automation Systems [1,2,4,6,7,8] exert requirements to increase performance, scalability, availability, security, reconfigurability, fault tolerance and graceful recovery from failures. These features of any application and discussed models are so called critical challenging factors [8].

Each realization owes some ways to manage and escape these weaknesses [8, 9]. Ubiquitous widespread automation models are based on hierarchical separation, while distributed models use delegation and separation of functions in one multi-layer model leading to Distributed Automation Systems (DAS). Each layer has a particular role in the overall functionality. Most commonly the system is build software and hardware heterogeneous, and communication model makes the whole system to work as an interconnected logic not separating individual parts working away from each other. Because of that reason communication is a crucial issue [2]. In our realization the employed model combines the advantages of both – distributed systems and automated systems [4, 7].

Communication model includes servers, which purpose is to connect one homogenous sub-system with others. The exchange of data and messages is based on classical client-server communication. Applications are working on each server to provide interconnections into entire system. Applications could be analyser [8] or transaction server, for data processing and getting complete distribution of information system. As a result of delegating functionality the model becomes light-weighted, with improved level of defeating DoS (Deny of Service). Even declining of any layer, or just a part of a sub-system, the rest functionality does not break down and still is working. If some self-testing algorithms are available, it is possible the information system to become self-reconfigurable and to continue working properly [6].

Separating the communication entry point to the distributed system by accounts or groups with equal privileges is crucial for requests analyzing and predicting data flows of each respond. Using data analyzer makes system capable to extract statistics and make point for improvement. It is possible to control data exchanges' route into data network.

Communication model is a significant part of each proposed concept closely connected with the model of system architecture. It is possible to avoid many of the problems by appropriate design of inner customized transaction messages. Well packed messages allow overcoming of heterogeneous character of the system, closely related with the design of communication model [14].

Recent years XML is demanding as a standard for customizable protocol for individual purposes [16]. This way for communication is very scalable and flexible, and easy for implementation.

Most of proposed models and realized solutions are based on distribution model, hierarchy organized, leading to simplicity in asynchronous interaction and information exchange into the inner network consisting of embedded systems and servers.

III. Discussed Model

This paper discusses realization of a distributed three-layer model, described in details in [14] (Figure 1). This model presents a simple distributed system for overcoming a lot of fore-mentioned problems. They are solved by realizing a few small applications, distributed on different machines with appropriate functionality and delegated privileges, to provide reliable system for data transaction with web interface. Distribution allows increasing availability of the system. Communication model is based on XML transaction messages. Two main points of the system are communication model and architecture model.

Proposed Communication Model overcomes more of the problems connected with web application; employing customizable XML transaction messages. The model is based on Local Area Network, which allows isolation and separation of different levels. That way increases security policy. Entire network is hidden behind a Web Server which connects the system to Internet. Internet supplies a connection to the system all over the world. The server accepts the client requests and sends to the lower level of the model a suitable transaction message. In three-layer-modelling, lower level applications are for data processing. Data processing includes parsing the request and working out suitable transactions to data layer for extracting the appropriate data. Data layer collects the data from embedded systems (in our realization these are values for measured temperature and humidity).

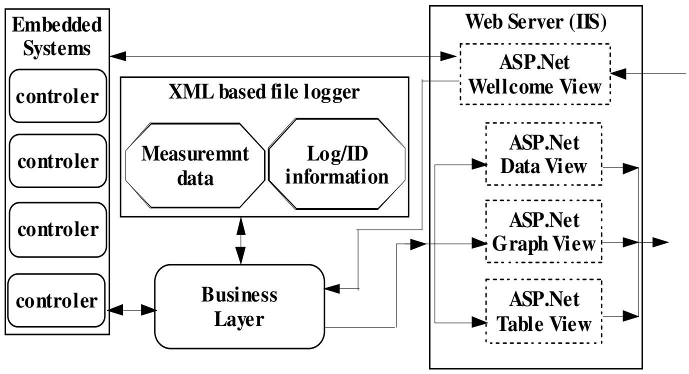


Fig. 1. Model Functional Scheme

In three-layer-model architecture the following layers are outlined:

- Client Processing Layer – this layer is used for interaction with end user of the system. The consumer could be a manager for controlling or monitoring of any parameters; this layer is represented by IIS (Internet Information Service), which is Microsoft technology for web server;
- Request/Response Processing and Data Management Layer are represented by application with one of the most important function in entire model. Its purpose is managing, collecting and distributing of entire data flow. This layer is centralized one. In most models this layer is called business layer. It makes the system homogenous because of its coordinating function. This part of communication is very flexible and could be expanded easily because of Object Oriented Thread programming.
- Data Processing Layer – its realization is the most distributed one among the system. Depending on its role in the model, the layer is separated in several tiers for data collecting, data storing, data logging and data extracting.

Each of these tiers interacts with upper layer of the model (Request/Response Processing and Data Management Layer). The aiming reason for distribution as much as possible is for increasing scalability and flexibility and decreasing the level of unavailability of entire system.

IV. Realization

The development tool used for the realization of the model is Visual Studio .NET. Web interface's realization is done with ASP.NET web pages. The interaction mechanism and code's connection is shown on Figure 2. So called Code Behind is doing all functionality, while View part visualize the web interface of the website. Cache is used for data transfer between Codes Behind and View's controls initializations fitted in web application.

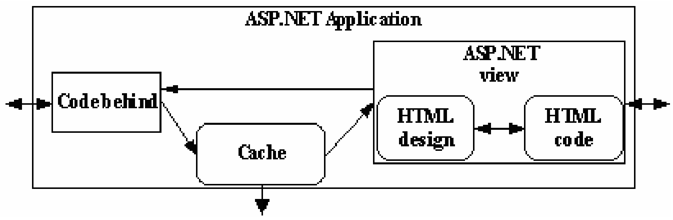


Fig. 2. Conceptual Scheme of ASP.NET

The realization of the model is three-layer: web page for acceptance of user's requests and displaying of the respond; Business Layer (Transaction Server) – it is for data processing layer; the last layer is data producer layer. Data processing layer is divided in two tiers: extract data and write it to Database (DB) and the second one for request acceptance from web server and read data from DB (Figure 3). Each tier is processed by an appropriate application. Its realization is done by building a separate thread for a single functionality. In this case Read and Write threads shares general resource – access to an XML filed driven Database. Synchronization is needed to be done over these two threads. It is a classic case of reader writer (producer/consumer) synchronization.

Read Thread accepts request from web server in a queue (FIFO). For each element in the queue the XML file is read from data server. If the file is not changed, since the last read, there is no need to re-read the same information, hence new reading is not initialized and the last read data is transferred. Write Thread works out a suitable transaction messages for extracting the necessary data from data producing layer (embedded systems). CNDEP protocol is used for data extraction [13]. When the data is extracted, it is written to an XML file.

Business layer's realization is done with daemon application, using TCP sockets for connection with web appli–cation, and UDP sockets for embedded systems connec–tion [13]. XML request/respond messages are transferred between business layer and web application. Business layer works out all necessary messages, packing them for producing an appropriate communication.

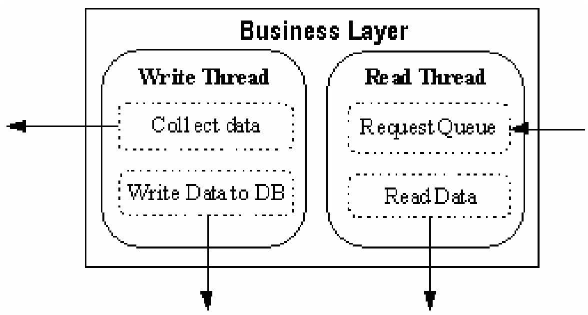


Fig. 3. Scheme of Business layer

This realization is very scalable, because a new function-nality could be very easily added to Business layer, and this part of the model reasonably is the most significant one. Business layer is the application which makes significant requirements to power calculation of the machine it is working on.

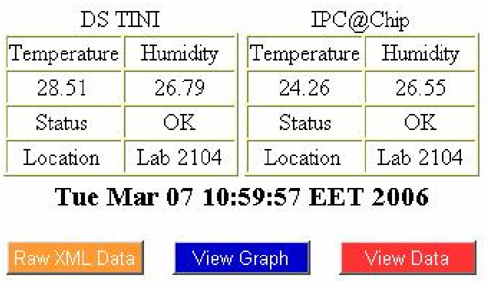


Fig. 4. General View (entry point for the system)

Web presentation part of the model (Figure 4) displays current values for measured temperature and humidity, as well as status of embedded system and the location of measu-rements. It also presents three functions; Raw XML data – no parsing of XML file is available except self-parsing capabi-lities of Internet Browser; View Graph displays in datagram view last ten measurements in the system (Figure 5). This representation of data is very suitable for fast data analysing. When elaboration is available in business layer the displaying data will be more accurate and adopted for the purpose of integration. The last function is for table view of already parsed data (Figure 6). The model represents the main approaches used for viewing and presenting data – in table and graphical view. The data, presented in such way, is possible to be analysed quickly, accurately and consistently with needed requests, done by the end user.

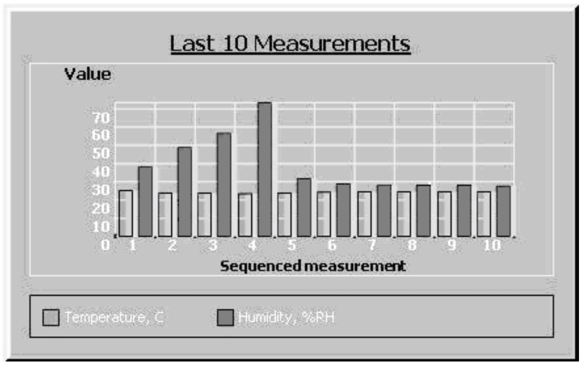


Fig. 5. Graph view representation

XML processing is the stage when data is received by an application and data is needed to be parsed. This stage is for data extracting. Many XML parsers exist and additional elaboration and modifications for each of them are available. Most often the parameters for comparison are reliability, data processing quickness, accuracy of parsed data and the scheme of parsing. In proposed realization a modification of SAX parser [12] is used, which is implemented in .NET. This parser is a "push" type. The other alternative is DOM parser. These two parsers are the most widely spread and commonly used ones.

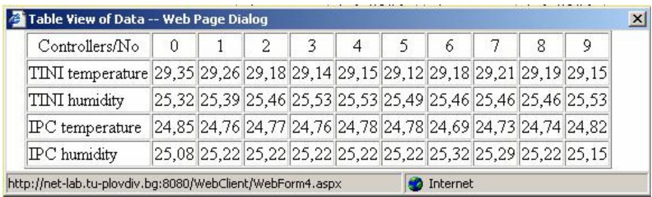


Table View of Data -- Web Page Dialog

Controllers/No	0	1	2	3	4	5	6	7	8	9
TINI temperature	29,35	29,26	29,18	29,14	29,15	29,12	29,18	29,21	29,19	29,15
TINI humidity	25,32	25,39	25,46	25,53	25,53	25,49	25,46	25,46	25,46	25,53
IPC temperature	24,85	24,76	24,77	24,76	24,78	24,78	24,69	24,73	24,74	24,82
IPC humidity	25,08	25,22	25,22	25,22	25,22	25,22	25,32	25,29	25,22	25,15

http://net-lab.tu-plovdiv.bg:8080/WebClient/WebForm4.aspx Internet

Fig. 6. Viewing Parsed XML data in table

XML files represent storing function of the DB. XML Reader class, used for XML parsing, is elaborated version of SAX parser, validation is skipped and no DTD files are required.

VI. Conclusion

Development and improvement of models for distributed embedded systems aim to overcome significant problems connected with scalability, flexibility, availability, easily expansion and modifications. All consequences from deve-loping embedded systems and implementing such models are oriented to significant decrease the level of DoS. For distributed automation this factor is really important. If self diagnostic tools are available it would be one really completed automation system. The realization of the proposed model conforms that most of the problems could be decided by delegating and separating the functionality on different levels and the whole system to be organized combining hierarchical and distributed approaches. Realization of hierarchical or distributed models is employed in systems, in which ideology is fitted multiple-point entries for requests and responds data flows, organized with data producing layers working independently from each other.

V. FUTURE WORK

A self-diagnostic system as well as developing of selective driven realization, based on web services, is required for accurate comparison of two possible approaches of system architecture. This approach allows comparing the proposed application to other realizations that uses HTTP as a transport protocol for communication.

VI. ACKNOWLEDGEMENTS

The work in this paper is supported by National Science Fund of Bulgaria project – "ВУ-966"/2005, entitled "Web Services and Data Integration in Distributed Automation and Information Systems in Internet Environment".

REFERENCE

[1] Britton, C., *IT Architectures And Middleware*, Addison-Wesley, 2001, 296 p.

[2] Burmakin, M., Krassi, B. "Distributed Automation and Control System"

[3] Duo Li, Serizawa, Y., Mai Kiuchi.. Concept design for a Web-based supervisory control and data-acquisition (SCADA) system. Proceeding of Transmission and Distribution Conference and Exhibition 2002: Asia Pacific. IEEE/PES , Volume: 1 , 6-10 Oct. 2002 Pages:32 – 36 vol.1

[4] Kakanakov, N., G. Spasov, "Adaptation of Web service architecture in distributed embedded systems", Proceedings on the International Conference – CompSysTech'05, 16-17 June 2005, Varna, Bulgaria, pp IIIB.10-1 – IIIB.10-6.

[5] Kakanakov,N., "Web Based Models For Distributed Automation", Proceeding Automatic and Informatic;

[6] Li, G., Sun, L., Li, H., "A Study of the self-organization Modeling of Distributed Manufacturing Network", Proceeding of the IEEE International Conference on E-Commerce Technology for Dynamic E-Business, 0-7695-2206-8/04

[7] Spasov, G. , N. Kakanakov, "CGI-based applications for distributed embedded systems for monitoring temperature and humidity", Proceedings on the International Conference – CompSysTech'2004,, Rousse, Bulgaria, pp. I.6-1–I.6-6.

[8] http://www.foundrynet.com/products/webswitches/serveriron/PDFs/XML%20Switching%20on%20ServerIron.pdf

[9] http://hum-molgen.org/meetings/courses/0689.html

[10] http://java.sun.com/products/jsp/

[11] http://msdn.microsoft.com/asp.net/

[12] http://msdn.microsoft.com/library/default.asp?url=/library/en-us/cpguide/html/cpconcomparingxmlreadertosaxreader.asp

[13] http://net-lab.tu-plovdiv.bg/CNDEP/

[14] http://net-lab.tu-plovdiv.bg/ASPThreeLayerModel.html

[15] http://www.microsoft.com/mind/1196/dcom.asp

[16] http://www.w3.org/XML/

An Approach for Development of Measurement Laboratories for Remote Experiments

Jelena Djordjević[1], Milan Jović[2], Dragan Živanović[3] and Miodrag Arsić[4]

Abstract **– The advance of the information and telecommunication technologies, based on multimedia and global networks, causes innovations in the field of measurement. The model of measurement laboratory for remote experiments is presented in this paper. Clients can access measuring system, i.e. instruments, via Internet and directly carry out real experiments.**

Keywords **– Virtual laboratory, client/server architecture, client application, remote experiments**

I. Introduction

Nowadays, in order to control and run technological processes, a number of process' parameters should be measured. Measuring is carrying out at different measurement places, with bigger or smaller distance between them. All measurement resources, which are used in that purpose, are connected in one functional entirety – measurement system. Low price of microprocessor's components and systems, made possible the realisation of the systems with distributed data processing.

The development of information technologies has opened new possibilities in realisation of measurement data aquisition systems. The idea of remotly accessible laboratories via Internet became true [1]. The laboratory heart consists of a group of specialised and/or general instruments, conected to the Internet through the PC. Within a remote measurement laboratory, clients can cooperate to each other, even they are on geographicaly distant places.

II. Distributed Architecture

As it's known, the distributed measurement systems are the systems where it is possible to realise the network by linking a great number of remote subsystems, and where it is possible to changes their's software. If access to these systems is alowed via Internet, we have remote laboratory, i.e. distributed measurement system based on client–server architecture. This systems are also called virtual laboratories.

A virtual laboratory for measurement and instrumentation must aim to realise an integrated environment for creation and distribution of virtual instruments and systems through a computer network [2].

[1]Jelena Djordjević is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia and Montenegro, E-mail: jelenadj@elfak.ni.ac.yu

[2]Milan Jović is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia and Montenegro, E-mail: milan.jovic@elfak.ni.ac.yu

[3]Dragan Živanović is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia and Montenegro, E-mail: dzile@elfak.ni.ac.yu

[4]Miodrag Arsić is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia and Montenegro, E-mail: marsic@elfak.ni.ac.yu

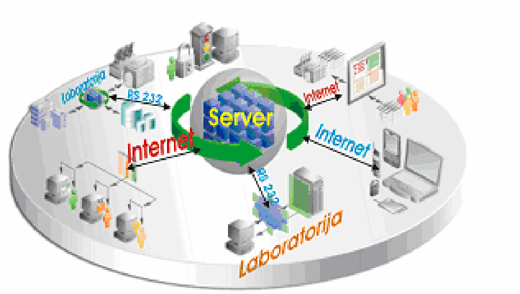


Fig. 1. Virtual laboratory.

The system architecture comprises a hierarchical structure on few levels (Fig. 1) [1]. Clients can access only to the first level via a simple Web browser. The base of this system is measurement system, which is conected to a computer (server) via standard communication interface (RS-232 protocol). It (allows) makes possible distribution of data and measured results via Internet to a client. System is based on standard protocols and features, and it can be easily expand.

Basic parts of the system are:

- main server,
- connection with measurement systems (laboratories),
- measurement instruments which perform the laboratory.

Main server manages the whole system. It allows access to the laboratory and the experiments, on one side, and forwards the tasks (assignment) to the features which are in the environment, on the other side. The instruments in the laboratory are connected to the main server over standard interface RS-232, which routes the instructions and results of the experiment.

III. Measurement Laboratory Realisation – Software Solution

Virtual instruments for measurement characteristics test of sensor module ADAM 4011 [3] are realised in order to develop the remote laboratory. These instruments are realised by program language LabVIEW [4]. They allows complete setup and control of the inteligent measurement module ADAM 4011, and measurement in all input bands.

Fig. 2 shows the hardware conection scheme of the virtual instrument. ADAM modules are connected to PC through RS-232/RS-485 convertor, ADAM 4520. Standard values of electric signals, generated by calibrator METRAtop 53, are attached to inputs of intelligent measurement module [5].

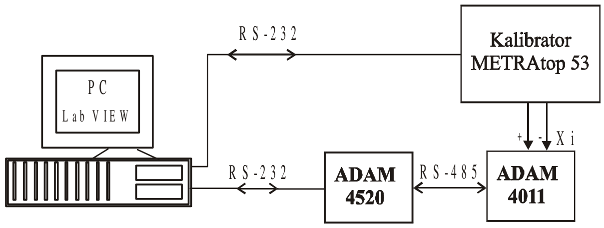


Fig. 2. Hardware connection scheme of the virtual instrument.

Tools and technologies, which are used for development of this virtual laboratory, are IIS, ASP, SQL Server, Visual C^{++}, Java, VB and CGI scripts (Fig. 3). Every mentioned technology and tool has a specific role in the laboratory work.

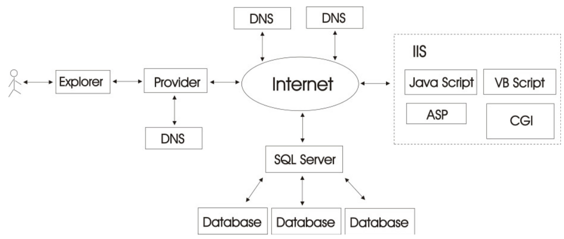


Fig. 3. Client application architecture.

In concrete system architecture, the Internet Information Server (IIS) is used as a web server. Client accesses the laboratory by standard browser and web address of the virtual laboratory. Browser forwards symbolic address to a local DNS server, where the pares of symbolic addresses and IP (Internet Protocol) addresses are hosted. If there are no errors, the page will be opened in the browser. Client requires the laboratory access by identifyng itself through Login and Password. After that, browser forwards the reqiurement to the virtual laboratory server. Web pages are written in ASP (Active Server Pages) program language.

The confirmation of the client is implemented at the begining of the next web page. The server checks added personal data and allows the client to access the system if the client exists in data base. If the client doen't exist in data base, server refuses that client with a note to try again.

After registration, new page appeares. The client has to choose a measuring experiment and requires the access. At the next page, the client has to enter specific set of data (range, sample rate, number of measurements...), which is necessary for laboratory setup. Based on this set of data, the input file of virtual instrument has been created.

By pressing the OK button, the proper LabVIEW application runs up. LabVIEW application carries out measurements with given pharameters, and the acheived results puts into output data file. Based on these data file, web page with the measurement results has been generated and the results are shown to the client.

The client receives report in the form of HTML page. In this case, the measurement results are given tabelary, thus the client can perform the required calculations. Calculations can be also performed using the equations given at the current page. Also, the client can graphicaly present results of the measurement.

Of course, the concurrence of more measuring requirements on the same measuring system at the same time has appeared. This problem can be solved in two ways. The first is not to allow any new client login while the laboratory is occupied by the client who is already logged and performs measuring. The second way is to put all logged clients in a queue. When a single measuring is done, the next client from the queue gets the laboratory for use (first-in first-out concept). The second solution is more practical because the laboratory occupation is splited between all logged clients. In this way, increase of clients number increases the waiting time in the queue, but it is equal for everyone who is in order to login.

IV. CONCLUSION

The information technologies advancement causes new possibilities for the realisation of data aquisition measurement systems. Suggested system is based on client-server architecture, it is easy to expand and it makes possibilities for distant clients to access the instruments. System expanding can be shown with the increasing number of laboratories, which are connected to the system. It will be necessary to develop a virtual instrument in the program language LabVIEW for each laboratory. Also, it will be necessary to make the program, which will enable the selection of available laboratory and prevent of client's concurrence, if the number of laboratories incresases. Equipment needed for the very next laboratory is certainly specific for every task and overload of a PC with a connected equipment is the only limit. It can be overcome by adding a new PC and connecting it to the established laboratory network. At the same time, the application has good bases for further development, which is related to easy handling and to the possibility of easy upgrading and interface changing. Thus, this lab shown possibility to make standards in comunication between modules, which role with devices are connected to the server, and internet applications which serve the users. If they would be created, widening of the system with new modules would become very simple and fast. It would bring up ability to develop system distributive by many developers from long distance places.

REFERENCES

[1] Jelena Đorđević, Miroljub Pešić i Miodrag Arsić, “An Approach for Distributed Measurement Systems Development”, *Metrological Congress 2003*, *Conference Proceedings,* Belgrade, Serbia and Montenegro, 2003.

[2] Alessandro Ferrero and Vincenzo Piuri, “A Simulation Tool for Virtual Laboratory Experimentation in WWW Environment”, *IEEE Transactions on Instrumentation and Measurement*, Vol. 48, No. 3, pp. 741-746, June 1999.

[3] *ADAM 4000 Series, Data Aquisition Modules*, User's manuel, Advantech, Taiwan, 1994.

[4] *Distance-Learning Remote Laboratories using LabVIEW*, User's manuel, National Instruments Corporation, USA, 2002.

[5] *METRAtop 51 ... 53 Benchtop Multimeter and Calibrator*, User's manuel, GOSSEN-METRAWATT GMBH, Germany.

Self-dependent Internet module for control and measurement

Ivo N. Dochev[1], Rumen I. Arnaudov[2] and Aleksander B. Bekiarski[3]

Abstract — **This paper presents a self-dependent Internet module for control and measurement. We describe a computer-system architecture, a block diagram of functional generator, data acquisition system and working algorithm.**

The remote control is realized by computer networks and using the TCP/IP protocols. For that purpose is used "Customer-Server" architecture.

Keywords **— Remote data acquisition system, Internet**

I. INTRODUCTION

Some of methods for control and automatic diagnostic of electronic systems require different calibration signals [1-8]. This problem is successfully solved by using the functional generators. These generators produce signals with calibration parameters. Some of them are: waveform, frequency, amplitude and phase.

The remote control and automatic diagnostic require communication area to be build. An opportunity is the existing computer networks to be used. This enables realizing remote control and automatic diagnostic without building new communication networks. Internet connects points from all over the world. We may use this property for carrying out our plan.

II. ARCHITECTURE

The computer-system architecture of self-dependent Internet module for control and measurement is represented on Fig. 1 [9-17]. The system includes users, Internet, local area network (LAN), functional generator board (FG), data acquisition board (DAB), the analyzed object, microcontroller (μC) and network controller (NC).

Communications between "users and Internet", "users and local area network", "Internet and network controller", "local area network and network controller " are based on TCP/IP protocols. The world wide web pages is based on HTML and Java script languages. The HTML language utilizes building web pages. The Java script language utilizes online data processing. The users need Internet browser software. Some of them are: Netscape communicator, Internet explorer, Opera, Konqueror web browser and Mozilla.

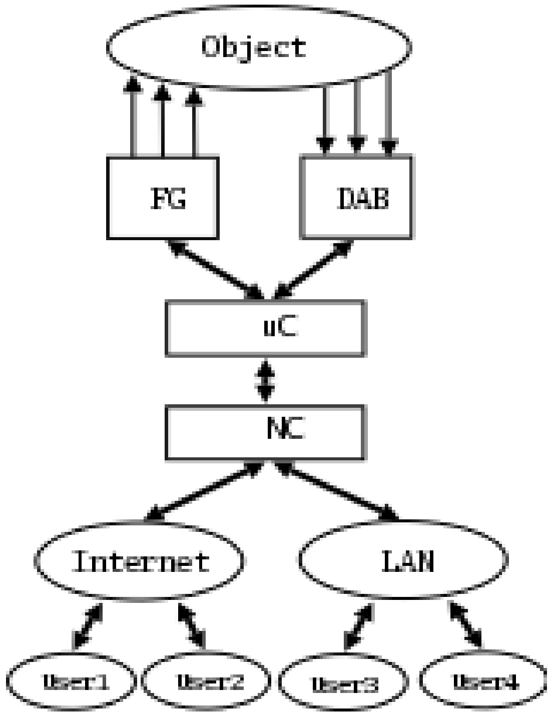


Fig. 1. The computer-system architecture of self-dependent Internet module for control and measurement.

III. USER ALGORITHM

The algorithm, which describes the processes in the user algorithm, is represented on Fig. 2. It includes login and control process. The user starts with connection to the self-dependent Internet module for control and measurement. After login the user chooses the signal type. The signal types are as follows: sinusoidal waveform, puls waveform, sawtooth waveform, triangle waveform and square waveform. The next step is selection of the signals calibration parameters. Some of them are: frequency, amplitude, phase and duty-cycle. Then the user chooses DAB channel, makes setup of the parameters and starting the measurement process. The software program makes data processing of the selected parameters. Then the data receive and display. The algorithm finishes with two options. The first option is the program returns to the main menu, so that, making a new choice of the signal type to be possible. The second option is the program goes to the end.

[1] Ivo N. Dochev is with the Faculty of Communications and Communications Technologies, Technical University, Kliment Ohridski 8, 1000 Sofia, Bulgaria, E-mail: idochev@tu-sofia.bg

[2] Rumen I. Arnaudov is with the Faculty of Communications and Communications Technologies, Technical University, Kliment Ohridski 8, 1000 Sofia, Bulgaria, E-mail: ra@tu-sofia.bg

[3] Aleksanser B. Bekiarski is with the Faculty of Communications and Communications Technologies, Technical University, Kliment Ohridski 8, 1000 Sofia, Bulgaria, E-mail: aabbv@tu-sofia.bg

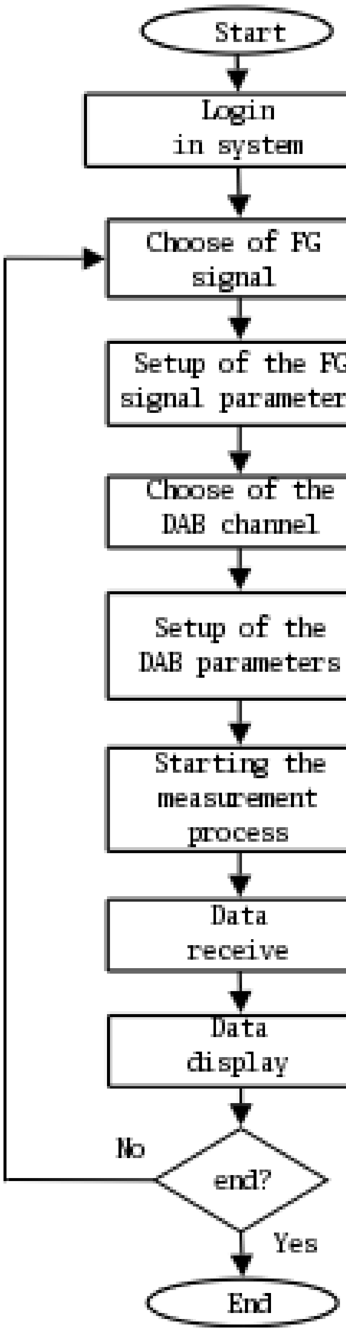


Fig. 2. Computer-system algorithm of of self-dependent Internet module for control and measurement.

IV. FUNCTIONAL GENERATOR BOARD

The functional generator board is represented on Fig. 3. The board consists of the following functional elements: microcontroller (μC), direct digital synthesizer (DDS), low-pass filter (LPF), amplifier (A) and programmable attenuator (PA).

The microcontroller receives the data from network controller (NC) and performs codes to the direct digital synthesizer and programmable attenuator. These codes define the work mode of the functional generator.

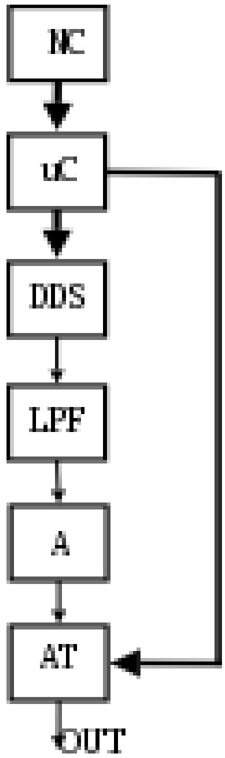


Fig. 3. Functional generator board controlled by Internet.

The parameters: waveform, frequency, phase, amplitude of the signals, generated by means of the direct digital synthesizer, may be controlled. All of them are defined by means of code sequence, produced from the microcontroller.

V. DATA ACQUISITION BOARD

The data acquisition board is represented on Fig. 4. The control board consists of the following functional elements: sensor (S), normalizing converter (NCon), analog multiplexer (Mux), programmable amplifier (PA), analog-to-digital converter (ADC), microcontroller (μC) and network controller (NC) .

It is imperative to be used a sensor for investigating the quantity if it is non-electrical. The normalizing converter is used for leveling the amplitudes of the signals in the dynamic range of the input signals. The normalizing converter consists of attenuators for signals with high amplitudes and amplifiers for low amplitude signals. One contains low-pass filters for eliminating the aliasing effect. The analog multiplexer selects only one of all input signals. The selected signal is amplified by means of the programmable amplifier. The amplified signal should be tend to the maximum value of the analog-to-digital converter range, but it must be weaker than the saturation value. The input signal mast be amplified on purpose of eliminating the quantization error. The measurement process, collecting of information and the connection with Internet is controlled by the microcontroller and network controller.

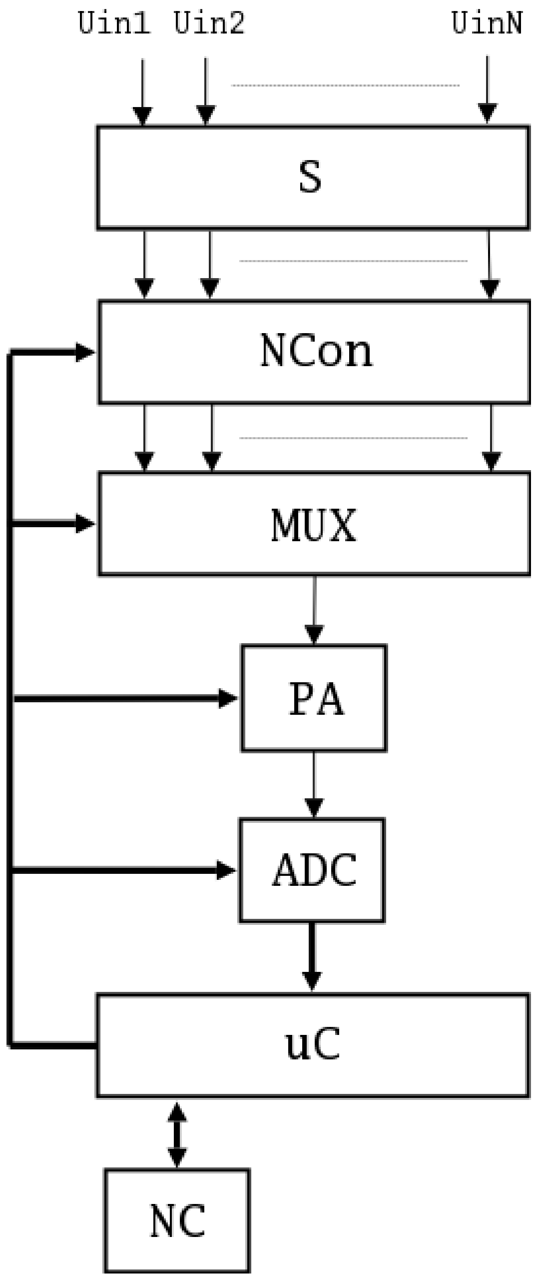


Fig. 4. Data acquisition board controlled by Internet..

IV. CONCLUSION

In this paper, we have presented a self-dependent Internet module for control and measurement. We described a computer-system architecture, block diagram of functional generator, data acquisition system and working algorithm. The remote control is realized by means of computer networks. The software algorithms are based on HTML and Java script languages.

The proposed self-dependent Internet module for control and measurement may be used for:

- remote control in industry;
- remote control in scientific research;
- remote automatic diagnostics of electronic systems, and
- distance learning education, which enable to analysis of different real objects investigation in laboratory work.

The advantages of the self-dependent Internet module for control and measurement, are:

- it is not necessary the setup of the system to be changed, when different objects are controlled,
- work flexibility,
- low cost,
- online data processing, and
- the browsing technologies enable asynchronous distance learning education with real active systems and devices.

REFERENCES

[1] A. Serdakov: Automatic control and technical diagnostics. Kiev, Technica, 1971.

[2] G. Franklin, F., J. D. Powell, M. L. Workman: Digital Control of Dynamic Systems. Stanford, Califonia.

[3] G. Mihov: Control and automatic diagnostics of micriprocesor system}. Sofia, Technical University, 1994.

[4] J. Marinov, E. Rangelova, V. Dimitrov: Technical control of radioelectronic system and device. Sofia, Technica, 1980.

[5] V. Kolriachko: Bild micriprocesor system control of REA. Moskva, Radio and sviaz, 1987.

[6] J. Marinov, E. Rangelova, V. Dimitrov: Technical control of radioelectronic system and device. Sofia, Technica, 1980.

[7] V. Karipskij, P. Parhomenko, E. Sogomonian: Technical diagnostics of control object. Moskva, Energia, 1967.

[8] V. Kolriachko: Bild micriprocesor system control of REA, Moskva, Radio and sviaz, 1987.

[9] V. Karipskij, P. Parhomenko, E. Sogomonian: Technical diagnostics of control object, Moskva, Energia, 1967.

[10] Zl. Stoilova: An access to database through Internet. In: Proc. Conf. Communication, Electronic and Computers Systems, vol. 2, Bulgaria, 2000, pp. 187-192.

[11] The Linux Serial Programming HOWTO: http://linuxdoc.org.

[12] The Linux Documentation: http://www.linux.org/docs/index.html.

[13] http://www.apache.org.

[14] http://www.php.net.

[15] http://www.linux.org.

[16] http://www.redhat.com.

[17] http://www.mysql.com.

Program System for Study of Surfaces' Micro Geometry

Vassil St. Donev[1], Slava M. Jordanova[2]

Abstract – **The paper presents program environment for study of surfaces' micro geometry. Proposed solution allows for substantial increase of monitored roughness parameters as well as for some new studies related to the functional features of the surfaces.**

Keywords – **roughness, surface, parameter of surface**

I. INTRODUCTION

The limited number of parameters defining roughness, submitted in the standards with numerical value, is related with the insufficiency of research of different parameters of roughness influence over the operational features of surfaces. The progress of microprocessor technique and the introduction of personal computers to the control – measuring technique allow wide ranges of possibilities for the devices, and more parameters to be monitored. [1,2,3,4]

The aim of this research is the creation of a program system for study of micro geometry of surfaces to provide determination of vastly larger number of parameters of roughness, and to allow studies related with the operational features of surfaces. The calculations are made on the base of data taken from system consisting of profile meter and analogue – digital converter ADC363. Parameters definition is based on going standards ISO 4287/1 and ISO 4287/2.

II. SYSTEM FOR TAKING THE INPUT DATA

The system for taking the input data is built on the base of the profile meter TAYLORHOBSON/Surtronic 3 (Great Britain).

Technical data:
Range of measurement - 0-9.99 and 0-25μm Ra;
Value of the piece - 0.25;0.8;2.5;8; 25 mm;
Feeler speed - 1mm/s per Ra; 0.25mm/s for a record;
Reading – Digital display;
Output - 10 mV/μm for range 99.9, 100 mV/μm for range 9.99;
Measuring precision - ±2% of the reading;
Environmental features :
temperature - +5 ÷ +40°C;
humidity -0÷99% relative humidity;

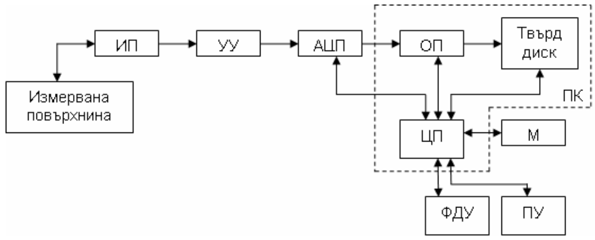


Fig. 1. Main blocks of the system

Symbols:
ИП – measuring device;
УУ – amplifier;
АЦП – analogue – digital converter;
ОП – random access memory;
ПК – personal computer;
М – monitor;
ЦП – central processing unit;
ФДУ – floppy disk drive;
ПУ – printing device;

The studied surface is tried by the feeler of the measuring device. After processing in the amplifier the obtained electrical analogue signal goes direct to the input of the analogue – digital converter. The analogue – digital converter transforms the information into digital and loads it into the random access memory of the personal computer. Right after that, the data is saved as a file on the hard disk drive of the personal computer. [1,2,3,4]

The developed program system reads the so created file and processes the information received. The calculated results for the parameters of the roughness are visualized on the monitor of the personal computer. The results might be printed on papr if needed or saved as a file on the hard disk drive or on a floppy diskette.

The program system processes the input data from measurements made with the system for obtaining input data. There are 25 parameters for roughness of the surface calculated for each base length as well as average for the five of them. The length of the base lengths is determined by ISO 4288 – 1985. The parameters of surfaces' roughness are determined by the standards of ISO 4287/1, ISO 4287/2.

III. SELECTION OF PROGRAM RESOURCES FOR PROGRAM REALIZATION

- *Visual Basic .NET 2003*

The environment for developing used in this work is Visual Studio .NET 2003 of Microsoft. It offers languages and tools for the creation of web – based, desktop and mobile applications.

[1] Vassil St. Donev - Techincal University, 9010 Varna, Studentska str. 1 ,
e-mail: donev@ms3.tu-varna.acad.bg
[2] Slava M. Jordanova - Techincal University, Computer Sciences and Technlogies , 9010 Varna, Studentska str. 1 ,
e-mail: slava_y@abv.bg

The environment of .NET (.NET Framework) is a multi-lingual environment, providing a rich platform for programming, developing of web based applications, data control, control of the resources access, processing of XML, program interface to the protocols HTTP, DNS, TCP, UDP. The language used for the design and programming of this application is Visual Basic .NET. It is also designed over the .NET environment. [5,6,7]

• *InnoSetup*

InnoSetup is a free installer for Windows applications. It supports all versions of Windows: Windows 95, Windows 98, Windows 2000, Windows 2003, Windows XP, Windows Me, Windows NT, Windows 4.0. This application makes it possible to create a single installation file .EXE for easier distribution.

InnoSetup provides a standard Windows 2000/XP style of the interface, full Uninstall opportunities and options for shortcuts. Support of several compression algorithms. In the present work InnoSetup is used for creating of the installer of the application.

• *Macromedia RoboHelp*

Macromedia RoboHelp is an application for creating of Help systems and documents or desktop and web based applications. RoboHelp gives the programmer the opportunity to choose the editor to work with. RoboHelp HTML editor can be used as well as any other HTML editor (Dreamweaver, Frontpage, Microsoft Word etc.) for generating of any popular help format (FlashHelp, WebHelp, Microsoft HTML Help, WinHelp, JavaHelp, Oracle Help for Java, and XML), creating of HELP systems working on every platform and browser. In this work RoboHelp is used for the Help part of the application.[5,6,7]

• *Macromedia Captivate*

Macromedia Captivate is a product used for interactive simulations and software demonstrations in Macromedia Flash format. Here Captivate is used for building of the Video manual for the application. [7]

IV. PROGRAM

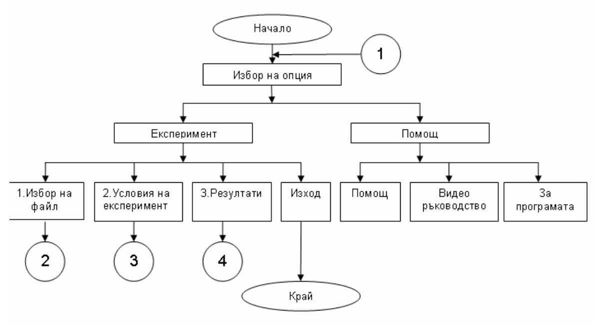


Fig. 2. Block scheme

The presented software system consists of:
frmMain.vb – main form;
ucPSelect.vb – control for tool selection;
ucResults.vb – control for visualization of the results;
mInit.vb – module containing functions loading data from the input file ;
mProcedures.vb – module containing procedures for mathematics realization.[5,6,7]

The main form *frmMain.vb* has operating functions. It watches the sequence of execution of the different stages of the experiment:

1. Selection of file;
2. Circumstances definition;
3. Results;

V. TESTING

Testing of the developed software system is made by external and structural tests, for finding and correcting of the errors.

• *External testing of program modules*

The external testing of modules is based on separating it from the main program and exploring its reactions resulted by the input data. [7]

• *Structural testing*

Structural tests are program oriented. For their realization are needed the following things:

a) Software separating of the module and defining its informational relation to the system. If the module refers to other modules they also need to be tested or on their place are put imitators;
b) Structure defining operators should be found and enumerated in the program source ;
c) Depending of type and semantics of the operators they should be presented as conditional IF structures;
d) Operators between two IF clauses are a linear section. IF operators are interpreted as nodes f the structural graph, an the linear sections as its arcs;
e) The structure graph is separated to a sequence of linear and nested loop cyclic sections;

For the execution of this test is used Visual Studio NET debugger and its Watch module. Breakpoints are placed and the values that parameters take are monitored. The passing through the individual braches of the graph is monitored. [6]

For the testing of the software is selected the top down strategy, where the main module is tested first. During the tests is accented the realization of the following functions:

• Impossibility for passing to step 2 of the experiment, without file loaded. This is provided by inactive options of the main menu or by message boxes;
• Impossibility for passing to step 3 of the experiment, without being loaded a file and defined circumstances for the experiment. This is provided by inactive options of the main menu or by message boxes;
• Visualization of the controls for selection of the experiment circumstances and output of results.

System requirements for software installation:

For the correct work of the program must be covered the following software and hardware requirements:

Hardware:

- X86 or compatible processor;
- At least 64 MB of RAM memory;
- For installation is required 3.4 MB free disk space;
- Graphic card with at least 1 MB of memory;
- CD-ROM or Floppy disk drive.
- Minimal Software requirements:
- Windows 98, 2000, XP;
- NetFramework 1.1;
- Microsoft Internet Explorer 5.01;
- Microsoft Windows Installer 2.0.[5,6,7]

VI. CONCLUSION

1. The developed software system for measuring and studying of the roughness of surfaces affords:

- Automated defining of 25 parameters for estimation of the roughness for one base length or average for the five base lengths;
- Representing of the results of the measurement in digital and graphical format;
- Implementing of wide program for studying of the micro geometry of technical surfaces and disclosure of the bindings between the different parameters for roughness.

2. The developed software system provides user customizable measurement of the parameters of roughness. The system can be used as base for developing specialized programs for quantity estimation of the bindings between separate parameters of roughness.

3. The results of the experimental measurements in conditions similar to the practice, shows high precision and reliability of the system.

4. In economical plan the system saves using of devices produced by the famous firms like Perthen (Germany) and Taylor-Hobson (Great Britain), which are having similar метрологични indices with these of that system and smaller count measured parameters.

5. So obtained values of the parameters empower for calculation and determining of the trajectory of movement, instant and average speed, and acceleration of analyzed object.

6. The necessary amount of memory for storing data obtained during analysis is defined.

7. Through database values and appropriate software the complete object’s behavior could be visualized during the process of analysis.

REFERENCES

[1] Русева Е., Русев Р. – Физическо металознание, ВМЕИ – Варна, 1988 г.
[2] Русева Е., Русев Р. – Физическо металознание – Ръковод-ство за лабораторни упражнения, ВМЕИ – Варна, 1986 г.
[3] „Програмиране на Microsoft Visual Basic .NET”, Джон Конел, Софтпрес, 2002 г.
[4] Microsoft Visual Basic NET – Професионални Проекти, Премиерпрес, 2003 г.
[5] Microsoft MSDN Library - http://www.microsoft.com/msdn/
[6] VB-Helper - http://vb-helper.com/
[7] Planet-Source-Code - http://planet-source-code.com/

Spectral and Statistical Analysis of Underwater Acoustic Signals

Yordan Sivkov[1], Ancho Draganov[2] and Mariya Nikolova[3]

Abstract – **Basic characteristrics of underwater acoustic signals are received through spectral and statistical analysis. Precision of this processing in sonar systems is very important for current requirements which are applied to them. The decision offered to solve the problem is consistent with these requirements and the current state of digital and computer technology.**

Keywords – **DSP, statistic and spectral analysis, underwater acoustic, noise**

I. Introduction

This paper studies the approaches in the analysis of underwater acoustic signals from passive sonar. They are in two domains - time and frequency. In hydro-acoustics we use three basic types of analysis that show main characteristics of signals [1, 2]:

– spectral analysis – present signals in frequency domain;

– correlation analysis – this analysis is part of statistical analysis but it is a very important characteristics of signals because noise doesn't have correlation and this helps us to divide a useful signal from an input mixture;

– statistical analysis.

A specific structure of signals radiated from different underwater objects is used in the process of detection, recognition and classification [2]. Receiving a description of signal characteristics after analyses the goal of the present paper is known as pre-processing. This signal processsing can be used in systems for automatic recognition in many different applications like machine monitoring, ship noise classification, traffic statistics, restricted areas surveillance, target detection, etc[4].

II. System Architecture

On Figure 1 is shown a block diagram of a real passive sonar system which realizes signal processing. It consists of sonar sensors, sonar system /filters, amplifiers and other/, analog to digital converter and FFT and statistical extractors.

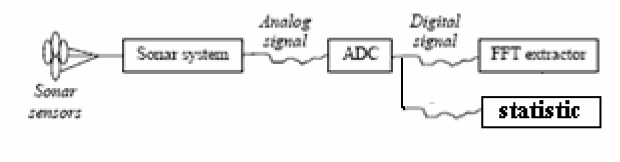


Fig. 1. System architecture

For the experimental character of the study, a system record of real signals is used through microphone input of the sound card of a portable computer, working in this case like an analog to a digital converter /ADC/. The output signal is taken from the hearing channel of a real sonar station. The files are recorded in a standard sound format .wav with discretion frequency 44100 Hz and 32 bits ADC. These parameters of input signal are completely enough for the needs of the study and the passive hydro location [3, 7].

III. Data Processing and Algorithm

Computed experiments are used to simulate the analysis of spectral and statistical analysis with recorded sound files. . For programming the realization of the experiment Matlab is used – a program environment for mathematical modulation. The reason for this choice is to reach libraries with a built in function which work with sound files, signal processing toolbox and toolboxes for the following signal processing. Another positive side of Matlab is input and output functions libraries [5].

The realization of the program on Matlab language follows the algorithm:

1. Reading wav-file and loading the values of signal, frequency /Fs/ and the number of bits per sample in array.
2. Drawing a graphic of signals.
3. Computing the frequency spectrum.
4. Drawing a graphic of frequency spectrum.
5. Verification of results with inverse FFT and drawing a graphic of signal.
6. Computing statistical characteristics /mean, standard deviation and co variation/ and drawing autocorrelation function.

For computing algorithm we use following mathematics:

To get the frequency spectrum, we calculate discrete Fourier transform, computed with a fast Fourier transform (FFT) algorithm[8]:

$$X(k)=\frac{1}{n}\sum_{i=o}^{n-1}X(n).W_N^{kn} \qquad k=0,1...,n-1 \qquad (1)$$

where:

X(k) - frequency spectrum, **N** - Time window size;

[1]Yordan Sivkov is with the Department of Electronics, Naval Academy "N. J. Vaptsarov", 73 Vassil Drumev St., 9026 Varna, Bulgaria, E-mail: j.sivkov@gmail.com

[2]Ancho Draganov is with the Department of Electronics, Naval Academy "N. J. Vaptsarov, 73 Vassil Drumev St., 9026 Varna, Bulgaria, E-mail: agdraganov@yahoo.com

[3]Mariya Nikolova is with the Department of "Mathematics and Informatics", Naval Academy "N. J. Vaptsarov, 73 Vassil Drumev St., 9026 Varna, Bulgaria, E-mail:mpn@abv.bg

X(n) - time function, $\mathbf{W_N} = \exp(-j2\pi/N)$.

These signal processing techniques were used in many previous works but precision is better than them. In this study are used FFT of 32384 points which leads to range discrimination less than 1.4 Hz of sample.

The approach used to calculate statistical characteristics of underwater acoustic signals is a standard Statistics formula for computing mean, standard deviation and autocorrelation function are shown below[6]:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2}$$

$$s = \left(\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{\frac{1}{2}} \tag{3}$$

$$\widehat{R}_{xx}(m) = \begin{cases} \sum_{n=0}^{N-m-1} x_{n+m} x_n^* & m \geq 0 \\ \widehat{R}_{xx}^*(-m) & m < 0 \end{cases} \tag{4}$$

IV. RESULTS

With the presented algorithm we examine five types of ships:

- Ship 1 – motor-launch, displacement – 150 BRT, propel - with high turnovers 900;
- Ship 2 – tanker, disp. – 90 000 BRT, propel – special diesel motor - 100 rpm;
- Ship 3 – tanker, disp. - 180 000 BTR, propel - diesel motor – 150 rpm;
- Ship 4 – merchant ship, disp. – 10 000 BTR, propel diesel with 800 rpm;
- Ships 5 – high speed passenger-ship, disp. – 5 000 t, propel – with high turnovers 500 rpm.

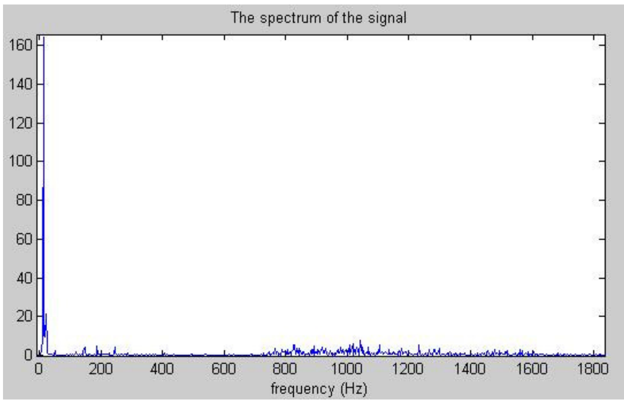
In table 1 is shown a parallel between statistical characteristics of these five types of ships. The second column shows the statistical mean of the signal and in the next column is calculated the standard deviation.

TABLE I
STATISTICAL CHARACTERISTICS OF SIGNALS

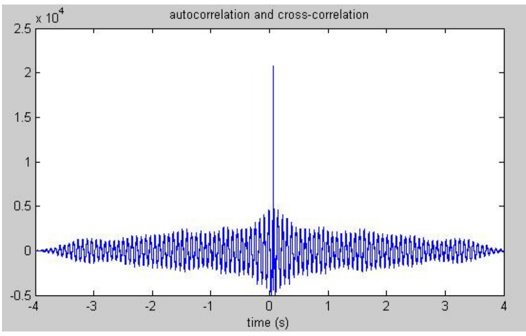
	mean	standard deviation
Ship 1	0.0012	0.3424
Ship 2	0.0013	0.3823
Ship 3	0.0015	0.0899
Ship 4	0.0209	0.1684
Ship 5.1	0.0034	0.3216
Ship 5.2	0.0012	0.3094

On Figure 2 – 7 are shown grafics of autocorrelation function and frequency spectrum of signals from the studied types of ships.

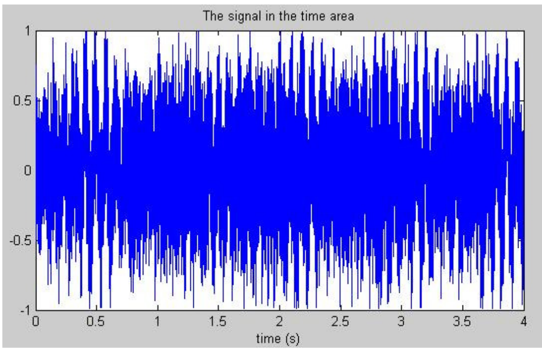
We see good visual distinguishing of all characteristics in different signatures which can be used to make other processing like classification and recognition.



a) FREQUENCY SPECTRUM



b) AUTOCORRELATION FUNCTION



c) TIME DOMAIN

Fig. 2. Ship 1 Autocorrelation and spectral analysis.

In figure 2 from above the signal is represented with three graphics: frequency spectrum (a), autocorrelation function (b) and signal in time domain (c). The goal of this presentation is to show the impossibility to detect a ship in time domain without a human operator.

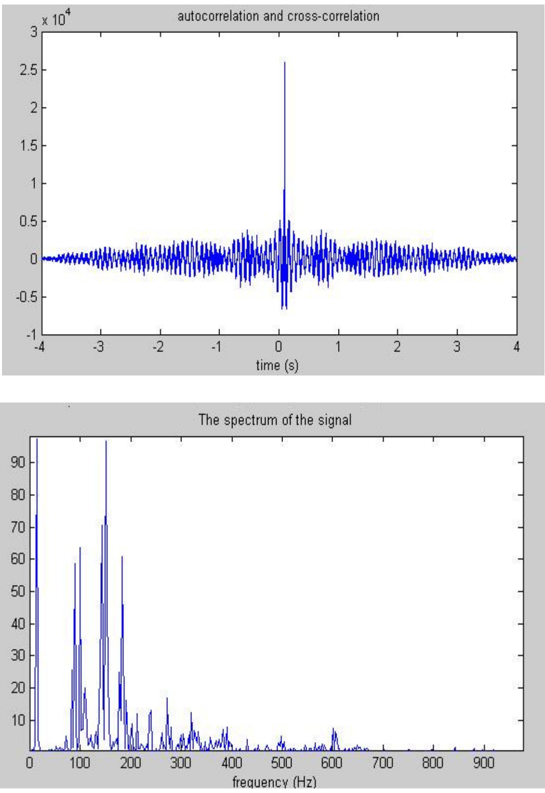
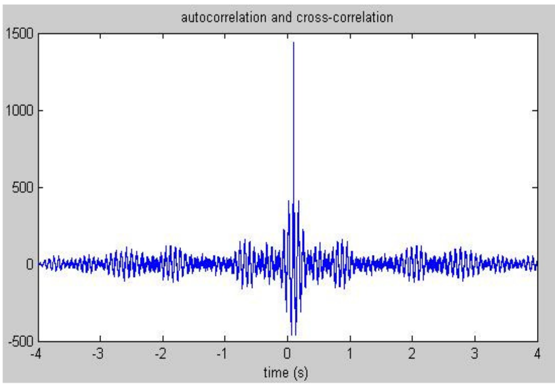
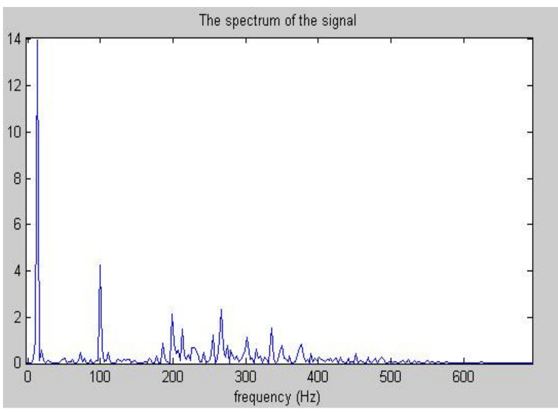


Fig. 3. Ship 2 Autocorrelation and Spectral Analysis.



a) autocorrelation function



b) frequency spectrum

Fig. 4. Ship 3 Autocorrelation and Spectral Analysis.

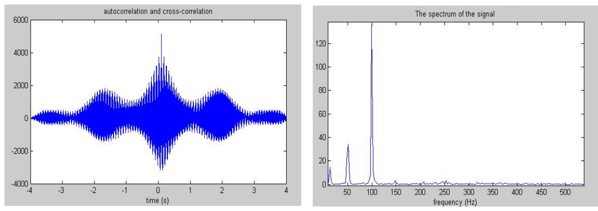


Fig. 5. Ship 4 Autocorrelation and Spectral Analysis.

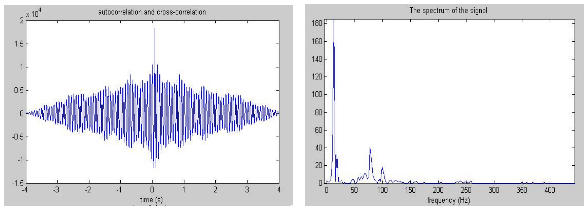


Fig. 6. Ship 5 Autocorrelation and Spectral Analysis.

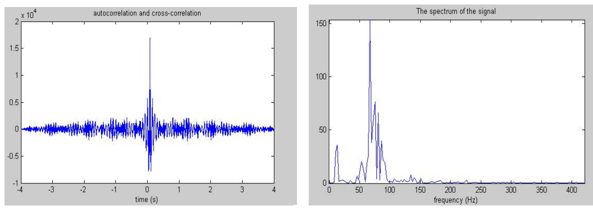


Fig. 7. Ship 5. Autocorrelation and Spectral Analysis.

In graphical presentation of spectral analysis results are shown only frequencies in range 0 – 1000 Hz because in passive hydro location the main noise signals radiated from ships are located there. The rest of the frequency spectrum can be ignored because other noises like wave, rain, animals, etc. have a high level [3].

On fig. 6 and 7 are shown spectral and autocorrelation graphics of ship 5 and difference between them are that first the ship move away and next they approach. Both of graphics have a same structure with different amplitude of signals.

V. CONCLUSION

Receiving results can be summarized in the following conclusions:

- from the received graphics and data in table 1 it is visible that different types of ships have different characteristics and they can easily be discriminated;
- the formation and record array with computed data can be used for the realization of following signal processing;
- in frequency domain we see pronounced signal of propeller in low frequencies (0-15 Hz) which give a turnovers and characteristics of propeller and information about ship (type of engine, size and turnovers and other);
- high frequencies (up to 1kHz) give to us additional information about ships – subclass, dimensions and other;
- the program area of Matlab allows the processing of signals in passive hydro location and it can be used to realize the entire system with reading data from PC inputs, pre-processing and detection, recognition and classification;

- receiving and collection of this characteristics aren't hard because realization is easy (just one portable computer);
- discrimination by computer program make decision accurate and powerful tool for following signal processing;

Apart from the conclusions above further trends can be outlined for the development of the survey:

- accumulation data base for large types of ships and statistic for every type;
- determination of the variation of characteristics of ships in different bearing and speed;
- realization of following signal processing for solving problems with recognition and classification.

REFERENCE

[1] A.P. Evtutov and other, "Hydro-acoustic handbook", Sankt Peterburg, "Sudostroenie", 1982(rus.).
[2] V.A. Zaraiskii, A.M. Turin, "Theory of hydro location", Sankt Peterburg, "Sudostroenie", 1975(rus.).
[3] A. Openhaim, "Application of DSP", Sofia, Tehnica, 1982 (bg).
[4] J.M. Fonseca, L. Correia, "A Real-Time Classifier for Identification of Acoustic Signatures", Computazione Evolutiva, Anno 1, v.1, pp. 5-13, Portugal, 1996.
[5] V. Diakonov, "Matlab 6 Educational course", Sankt Peterburg, Piter, pp.534-537, 2001 (rus.).
[6] D.C. Montgomery, G.C. Runger, "Applied statistics and probability for engineering", 3rd. ed., John Wiley & Sons Inc., 2002.
[7] A.G. Draganov, "Analog to digital converter for signals with PIC processor and recording in PC", Naval Academy Conference 2002, pp. 103-108, Varna, 2002.
[8] R.G. Lyons, "Understanding digital signal processing", Prentice Hall Ptr., 2001.

Input-output linearization control of induction motors with load torque compensation

Stanislav E. Enev

Abstract: **An input-output linearizing control for the third order induction motor model described in the fixed stator frame, along with a simple scheme for identifying and compensating the load torque is presented in the paper. The load torque identification algorithm is designed based on rotor speed information. A simulation study of the proposed control algorithm is presented.**

Keywords: **induction motor, input-output linearization control.**

I. Introduction

The induction motor is probably the most widely used electric machine in industrial applications due to its reliability, ruggedness and relatively low cost. Its control however presents an extreme challenge because of the highly complicated nonlinear dynamics of the machine. These two reasons make it very attractive for control researchers and practitioners. A lot of approaches to this problem can be found in the literature. The first solutions gave the so-called field-oriented control [5],[8], which consists of rewriting the equations of the motor through a nonlinear transformation in order to decouple the rotor flux and the rotor speed. The disadvantage of this method is that the decoupling is valid only after the flux is constant.

The control algorithm investigated here is based on the feedback linearization of the induction motor. A good introduction to the exact linearization by means of a state feedback can be found in [10],[11],[12]. A lot works based on this approach can be found in the literature. In [3] and [4], dynamic feedback linearizing transformations are presented, enabling the full state linearization of the induction motor. However the most commonly targeted for control variables i.e. the rotor speed and the rotor flux magnitude are well “hidden” in the new transformed states, so that the advantage of the linear behavior of the system is reduced. Input-output linearizing transformations for the induction motor, with system outputs the rotor speed and the rotor flux magnitude, are found in [1], [2], [5], [9]. Using this approach, the stability of the so-called zero dynamics (in the motor dynamics remains a nonlinear part, made unobservable by the introduced feedback) must be guaranteed. The advantage of the input-output linearization approach over the field-oriented control is the fact that, by applying the linearizing transformation, a complete decoupling of the rotor speed and flux is achieved, which enables the optimization of the power efficiency of the motor without degradation of the speed regulation. However, the exact cancellation of the nonlinear terms is possible only when perfect knowledge of the motor parameters and the load torque is available. Generalized algorithms for adaptive control of feedback linearizable systems are proposed in [11]. In [9], an adaptive input-output linearizing control is designed for a fifth-order model of the motor (voltage-command mode), including algorithms for identification of the load torque and the rotor resistance, which are assumed to be constant.

Here, an input-output linearizing control for the third order (current-fed) induction motor model described in the fixed stator frame, along with a simple scheme for identifying and tracking the load torque, which enters the system as a disturbance is presented and investigated.

Stanislav E. Enev, Ph.D. Student, French Language Department of Electrical Engineering,Technical University of Sofia, 8 Kl. Ohridski Str., 1000 Sofia, BULGARIA, e-mail: sta_enev@yahoo.com

II. Dynamic modeling of the induction motor

The induction motor considered here is a three-phase stator, three-phase short circuited rotor machine. Since a squirrel-cage rotor can be represented as a three-phase short-circuited one by means of a simple transformation, the following considerations are valid for this case too. The common assumptions are adopted i.e. symmetrical construction, linearity of the magnetic circuits, sinusoidal distribution of the field in the air-gap. After a series of transformations, the following two-phase equivalent model with all state variables, i.e. the stator currents $I_{S\alpha}$, $I_{S\beta}$, the rotor fluxes $\Psi_{R\alpha}$ $\Psi_{R\beta}$, expressed in the fixed α-β stator frame is obtained:

$$\begin{aligned}
&\dot{\omega} = \mu(\Psi_{R\alpha} I_{S\beta} - \Psi_{R\beta} I_{S\alpha}) - c/J\omega - \tau_L/J \\
&\dot{I}_{S\alpha} = -\gamma I_{S\alpha} + \eta\zeta\Psi_{R\alpha} + \zeta n_p \omega \Psi_{R\beta} + V_{S\alpha}/(\sigma \bar{l}_S) \\
&\dot{I}_{S\beta} = -\gamma I_{S\beta} + \eta\zeta\Psi_{R\beta} - \zeta n_p \omega \Psi_{R\alpha} + V_{S\beta}/(\sigma \bar{l}_S) \\
&\dot{\Psi}_{R\alpha} = -\eta \Psi_{R\alpha} - n_p \omega \Psi_{R\beta} + \eta m I_{S\alpha} \\
&\dot{\Psi}_{R\beta} = -\eta \Psi_{R\beta} + n_p \omega \Psi_{R\alpha} + \eta m I_{S\beta}
\end{aligned} \quad , (1)$$

with

$$\mu = n_p m/(\bar{l}_R J),\ \eta = r_R/\bar{l}_R,\ \zeta = m/(\sigma \bar{l}_R \bar{l}_S),$$
$$\sigma = (\bar{l}_R \bar{l}_S - m^2)/\bar{l}_R \bar{l}_S,\ \gamma = (\bar{l}_R^2 r_S + m^2 r_R)/(\sigma \bar{l}_R^2 \bar{l}_S).$$

where: $l_{S(R)}$ - the stator(rotor) windings inductances, $r_{S(R)}$ - the stator(rotor) windings resistances, $m_{S(R)}$ - mutual inductances between the stator (rotor) windings, m_0 - mutual inductance between stator and rotor windings, $m = 3/2m_0$,

$\bar{l}_S = l_S - m_S$, $\bar{l}_R = l_R - m_R$, J - rotor moment of inertia, τ_L - a load torque, c - friction coefficient, n_p - number of pole-pairs. The complete derivation of the model can be found in [2],[5],[6],[7].

By using fast control loops (generally implemented with high-gain PI controllers), the currents are forced to follow a reference trajectory. Thus, if the tracking is fast enough, the current dynamics equations can be neglected and one can achieve current command of the motor, with $I_{S\alpha}$ and $I_{S\beta}$ the new inputs. The system can be put in the general form $\dot{x} = f(x) + g_1(x)u_1 + g_2(x)u_2$:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -c/Jx_1 - \tau_L/J \\ -\eta x_2 - n_p x_1 x_3 \\ -\eta x_3 + n_p x_1 x_2 \end{bmatrix} + \begin{bmatrix} -\mu x_3 \\ \eta m \\ 0 \end{bmatrix} .u_1 + \begin{bmatrix} \mu x_2 \\ 0 \\ \eta m \end{bmatrix} .u_2 \quad (2)$$

with

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \omega \\ \Psi_{R\alpha} \\ \Psi_{R\beta} \end{bmatrix}, \quad u_1 = I_{S\alpha}, \quad u_2 = I_{S\beta}.$$

III. INPUT-OUTPUT LINEARIZATION

For a 2-by-2 system of the form:

$$\begin{aligned} \dot{x} &= f(x) + g_1(x)u_1 + g_2(x)u_2 \\ y_1 &= h_1(x) \\ y_1 &= h_2(x) \end{aligned}, \quad (3)$$

with y_1 and y_2 the outputs, the feedback linearizing control law is given by:

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = A(x)^{-1} \begin{bmatrix} v_1 - L_f^{\gamma_1} h_1 \\ v_2 - L_f^{\gamma_2} h_2 \end{bmatrix}, \quad (4)$$

where $A(x) = \begin{bmatrix} L_{g_1} L_f^{\gamma_1 - 1} h_1 & L_{g_2} L_f^{\gamma_1 - 1} h_1 \\ L_{g_1} L_f^{y_2 - 1} h_2 & L_{g_2} L_f^{\gamma_2 - 1} h_2 \end{bmatrix}$ - decoupling matrix.

$L_* h$ stands for the Lie derivative of the scalar function h with respect to the corresponding vector field.

As long as the decoupling matrix is non-singular, the feedback law transforms the system into two decoupled differential equations representing two chains of integrators, v_1 and v_2 being the new inputs

$$y_1^{(\gamma_1)} = v_1 \text{ and } y_2^{(\gamma_2)} = v_2. \quad (5)$$

The zero dynamics is given by:

$$\dot{x} = f(x) - [g_1 \quad g_2] A(x)^{-1} \left[L_f^{\gamma_1} h_1 \quad L_f^{\gamma_2} h_2 \right]^T. \quad (6)$$

For the case of the third order induction motor model, we define the outputs as follows:

$$h_1 = x_1, \; h_2 = x_2^2 + x_3^2. \quad (7)$$

We have the following:

$$L_f h_1 = -c/Jx_1 - \tau_L/J, \; L_{g_1} h_1 = -\mu x_3, \; L_{g_2} h_1 = \mu x_2$$
$$L_f h_2 = -2\eta(x_2^2 + x_3^2), \; L_{g_1} h_2 = 2\eta m x_2, \; L_{g_2} h_2 = 2\eta m x_3.$$

The decoupling matrix is given by:

$$A = \begin{bmatrix} -\mu x_3 & \mu x_2 \\ 2\eta m x_2 & 2\eta m x_3 \end{bmatrix}. \quad (8)$$

Since $\det(A) = -1/(2\eta m \mu (x_2^2 + x_3^2))$, the matrix is nonsingular as long as the rotor flux is not zero.

Applying the state feedback control law:

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} -\mu x_3 & \mu x_2 \\ 2\eta m x_2 & 2\eta m x_3 \end{bmatrix}^{-1} \begin{bmatrix} v_1 + c/Jx_1 + \tau_L/J \\ v_2 + 2\eta(x_2^2 + x_3^2) \end{bmatrix} \quad (8)$$

the system is transformed into the following two equations:

$$\frac{dx_1}{dt} = v_1, \frac{d}{dt}(x_2^2 + x_3^2) = v_2. \quad (9)$$

The zero dynamics is stable.

IV. LOAD TORQUE IDENTIFICATION

In the case, when the value of the load torque is not exactly known, in the presence of the linearizing control law, the rotor speed equation is given by the following:

$$\frac{dx_1}{dt} = \dot{y}_1 = \frac{(\tau_{LR} - \tau_L)}{J} + v_1, \quad (10)$$

where τ_{LR} - the load torque value fed to the linearizing controller. We can define the linearizing controller load torque as:

$$\tau_{LR} = k_L (y_{1ref} - y_1), \quad (11)$$

where y_{1ref} is the output of an ideal integrator given by $\dot{y}_{1ref} = v_1$. Substituting the expression for τ_{LR} in the output equation and using the Laplace transform, one can obtain the following expression:

$$y_1(s) = \frac{v_1(s)}{s} - \frac{1}{k_L}\frac{\tau_L(s)}{T_L s + 1}\ , \qquad (12)$$

where $T_L = \dfrac{J}{k_L}$. It can be seen that in the case of a load torque entering the system as a step function, the behavior of the first output is "returning" to that of an ideal integrator with a time constant T_L, i.e. the load torque is tracked asymptotically with the same time constant.

V. SIMULATION RESULTS

The control scheme is presented in Fig. 1. Speed measurement is assumed. A conventional open-loop observer, representing a simulation of the motor equations, is used to generate the flux signals.

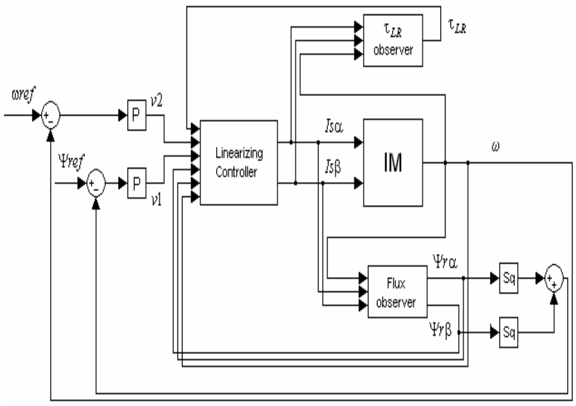


Fig. 1. Control scheme

For the simulation, the same motor as in [7] is selected. The motor parameters are as follows:

$m = 0.0813H$, $\bar{l}_S = 0.084H$, $\bar{l}_R = 0.0852H$,
$r_R = 0.842\Omega$, $r_S = 0.687\Omega$, $J = 0.03kgm^2$, $n_p = 2$,
$c = 0.0014kgm^2s^{-1}$.

In the outer control loops, P controllers are used for both control subsystems, as seen from Fig. 1. The gain values are set to 60 for the speed control loop, and 40 for the flux control loop. The reference signals represent the responses of a simple lag for the speed loop and a critically damped second-order system for the flux loop to step inputs. This is done to avoid jumps in the current values. In Fig. 2 are given some transient responses showing the evolution of the two outputs, the rotor fluxes and the stator currents for τ_L changing as in Fig.3 and when the exact values of the motor parameters are known. The response of the speed subsystem to load torque changes entering the system as step functions and the torque tracking are shown in Fig. 3. It can be easily found that the transfer function from the load torque τ_L to the output y_1 in the presence of the outer P control loop is given by:

$$\frac{y_1(s)}{\tau_L(s)} = -\frac{1}{kk_L}\frac{s}{(Ts+1)(T_L s+1)}\ , \qquad (13)$$

where k is the gain of the P controller (equal to 60 in this case) and $T = 1/k$. Thus, the effect of the disturbance (the load torque) is asymptotically reduced to zero, assuming that it is a step function. The speed of response can be tuned by the choice of k and k_L. The value of k_L is set to 5 for the simulation.

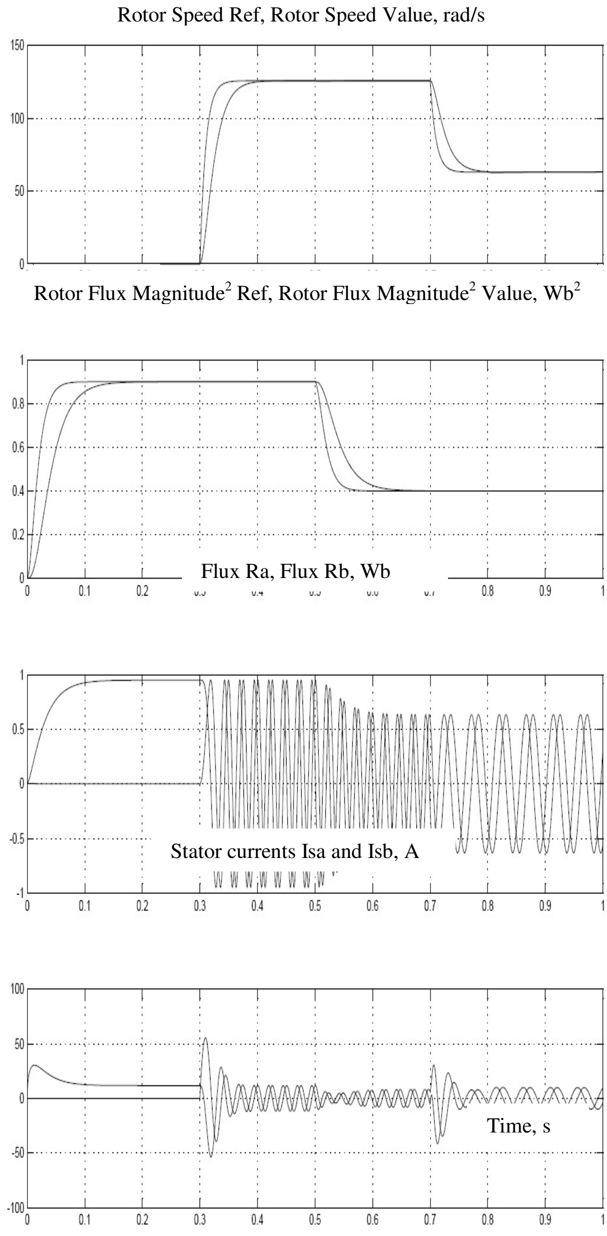


Fig. 2. Some transient responses

The rotor flux dynamics isn't affected by the load torque uncertainty as seen from the motor equations. This remains also in the presence of the linearizing control law.

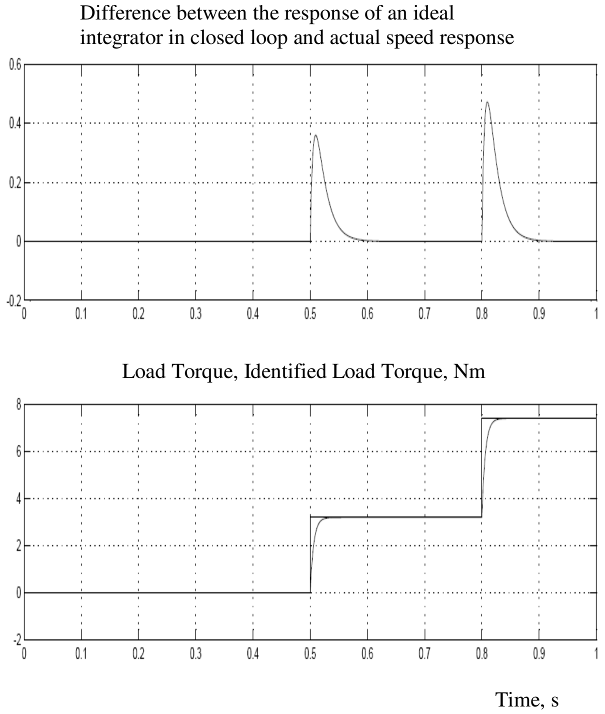


Fig. 3. Speed response to step load torque

VI. CONCLUSIONS

An input-output linearizing control for the third order (current-fed) induction motor model described in the fixed stator frame is presented in the paper. Assuming exact knowledge of the values of the motor parameters and the load torque, the use of this control scheme enables the complete decoupling of the dynamics of the system outputs i.e. the rotor speed and the square of the rotor flux. In order to compensate for changes in the load torque, a simple scheme for identification of the load torque is introduced in the control algorithm. The load torque identification algorithm is designed based on rotor speed information. The analysis of the identifier shows that the load torque is tracked asymptotically with time constant T_L in the case its changes take place as step functions. This enables a compensation for these changes in the linearizing control law, and thus the behavior of the speed output can "return" to that of an ideal integrator with the same time constant .

An algorithm for identification of the rotor resistance value, which may change significantly due to heating, can be introduced in the control scheme, in order to ensure the proper work of the conventional open-loop observer.

REFERENCES

[1] M. Bodson, J. Chiasson, *Differential-Geometric Methods for Control of Electric Motors*. Int. J. Robust Nonlinear Control, pp. 923-954, 1998.

[2] M. Bodson, J. Chiasson, R. Novotnak, *High-Performance Induction Motor Control Via Input-Output Linearization*, IEEE Control Systems, August 1994.

[3] J. Chiasson, *A New Approach to Dynamic Feedback Linearization Control of an Induction Motor*, IEEE Trans. On Automatic Control, vol. 43, №3, pp. 391 – 397, March 1998.

[4] J. Chiasson, *Dynamic Feedback Linearization of the induction motor*, IEEE Trans. On Automatic Control, vol. 38, pp. 1588 – 1594, Oct. 1993.

[5] J. Chiasson, *Modeling and High-Performance Control of Electric Machines*, John Wiley & Sons, Inc., Hoboken, New Jersey, 2005.

[6] E. Delaleau, J.P. Louis, R. Ortega, *Modeling and Control of Induction Motors*, Int. J. Appl. Math. Comput. Sci., 2001, Vol. 11, №1, pp. 105-129.

[7] J. Hu, D.M. Dawson, Z. Qu, *Adaptive tracking control of an induction motor with robustness to parametric uncertainty*, IEE Proc.-Electr. Power Appl, Vol. 141, № 2, March 1994.

[8] W. Leonhard, *Control of Electrical Drives*, 2nd Edition, Berlin, Springer, 1996.

[9] R. Marino, S. Peresada, P.Valigi, *Adaptive input-output linearizing control of induction motors*, IEEE Transactions on Automatic Control, Vol 38, Issue: 2 pp.208-221, Feb 1993.

[10] W. Respondek, "Geometry of Static and Dynamic Feedback". *Summer School on Mathematical Control Theory*, Trieste, Italy, Sept. 2001.

[11] S. Sastry, M. Bodson, *Adaptive Control: Stability, Convergence, and Robustness*, Prentice-Hall Advanced Reference Series (Engineering), 1989.

[12] J.J.E. Slotine, W. Li, *Applied Nonlinear Control*, Englewood Cliffs, NJ: Prentice Hall, 1991.

PSpice Simulation of Atmospheric Pressure Air Glow Discharge Current-Voltage Characteristic

Peter D. Dineff[1], Diliana N. Gospodinova[2] and Elisaveta D. Gadjeva[3]

Abstract – **The current-voltage characteristic of an Atmospheric Pressure Air Glow (*APAG*) discharge has been simulated by commercial circuit simulation software such as PSpice®. PSpice model has been developed for the plasma discharge in a *APAG* discharge applicator, which consists of two parallel electrode plates with a small gap between electrodes. At least one of the electrodes is covered by a dielectric barrier.**

An *APAG* discharge operating gap can be modeled as an electric capacitor without plasma, as an air gap containing plasma. The cold plasma itself has been modeled as a voltage-controlled current source that switches on when the voltage across the air gap exceeds the value of the discharge ignition voltage.

The simulation current-voltage behavior agrees with the experimental data from an actual parallel-electrode-plate plasma generator. It has been found that in different operating regimes, the discharge current of the *APAG* discharge plasma generator is described by a voltage linear law.

Keywords – **Atmospheric pressure air glow (*APAG*) discharge, cold plasma, one-atmosphere glow discharge, plasma ignition voltage, voltage-controlled current source, current-voltage behavior, current-voltage characteristic.**

I. Introduction

[1]The characteristics of many electrical systems can be simulated with proprietary computing tools such as *PSpice*®. Devices employing cold plasmas are embedded in electrical systems in many situations. It is advantageous to simulate the complete system, including the plasma as phenomena, with such commercial software. Previous paper regarding the computational simulation of high pressure plasma in air discharge were investigated, [2, 3, 4].

Normally the *APAG* discharge plasma system for plasma-chemical modification of low energy surfaces consists of a power supply, transformer, impedance matching network, and plasma applicator. This electrical system was an object of simulation with such computational tools [2, 4].

In an electrical discharge system, the inductors and capacitors, in the impedance matching network, the power supply, and the transformer are ordinary electrical components and have well-developed *PSpice* models. However, there is no available electrical model in *PSpice* for the plasma *APAG* discharge, so a principal simulation task is to develop such a model, [3, 4, 6].

J.-R. Roth introduced, on the base of the observed phenomenological characteristic of the normal glow discharge voltage-current behavior, the name "*one-atmosphere uniform glow discharge*" (OAUGD). Other cases of such behavior have been often reported in the contemporary literature. The normal glow discharges, as the dielectric-barrier discharges, ignite and burn at constant plasma ignition (or burning) voltage, [2].

R. Gadri shows that an atmospheric *RF* glow discharge in helium exhibits the same phenomenology as a current source, and its output current follows a power law of the applied voltage [2].

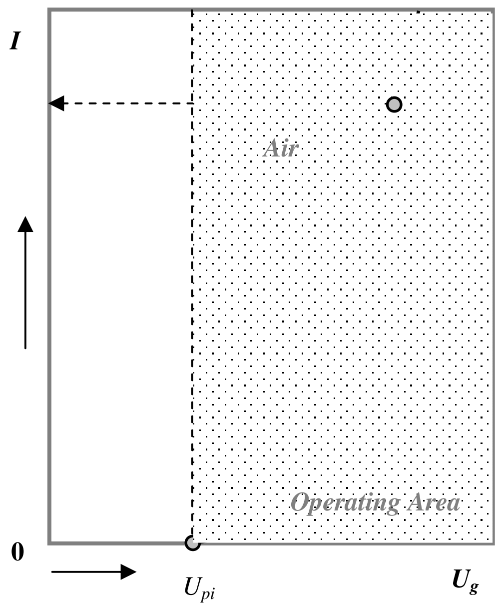


Fig. 1. Power law model of current-voltage behavior of an *APAG* discharge according to *J. Roth.*

J. Roth reported that the current-voltage (I-U) relationship of the high-power atmospheric pressure air glow discharge was $I \propto U^2$, and $I \propto U^3$, depending on the operating regime - the output current I is defined in Eqn. 1 by a power law function of the difference between the gap voltage U_g and the plasma initiation voltage U_{pi}, in order to simulate the current-voltage behavior in the operating range, Fig. 1 [2]:

$$(1) \qquad I \begin{cases} =0, \text{ for } U_g < U_{pi} \\ \alpha\left(U_g - U_{pi}\right)^n, \text{ for } U_g > U_{pi} \end{cases} ,$$

where n is an integer that ranges from 1 to 12 in different air glow discharge plasma devices.

[1]Peter D. Dineff is with the Faculty of Electrical Engineering, Technical University of Sofia, Blvd. St. Kliment Ohridski 8, 1000 Sofia, Bulgaria, E-mail: dineff_pd@abv.bg.

[2]Diliana N. Gospodinova is with the Faculty of Electrical Engineering, Technical University of Sofia, Blvd. St. Kliment Ohridski 8, 1000 Sofia, Bulgaria, E-mail: dilianang@abv.bg.

[3]Elisaveta D. Gadjeva is with the Faculty of Electronic Engineering and Technology, Technical University of Sofia, Blvd. St. Kliment Ohridski 8, 1000 Sofia, Bulgaria, E-mail: egadjeva@tu-sofia.bg.

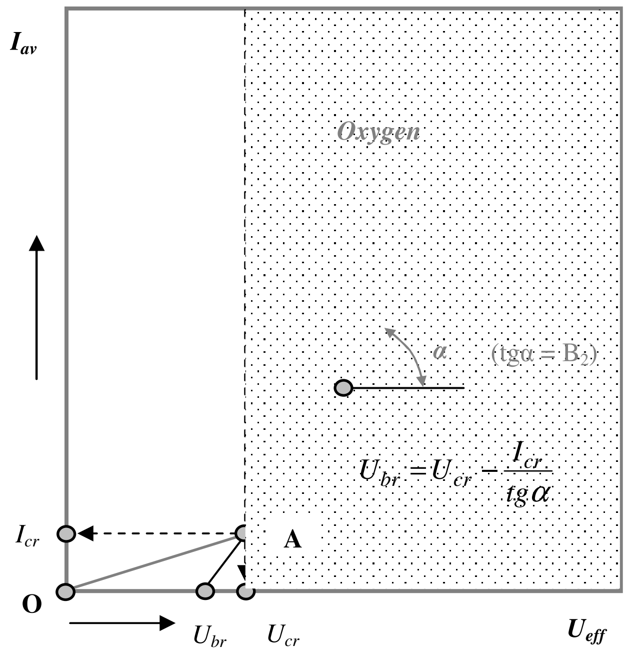


Fig. 2. Linear model of the current-voltage behavior of an *APAG* discharge in oxygen according to *I. Emelyanov*.

I. Emelyanov assumed that the current-voltage relationship of the atmospheric pressure glow discharge in oxygen was $I_{av} \propto U_{eff}$ - the output average current I_{av} is defined in Eqn. 2 by a linear law function of the gap effective voltage U_{eff} , Fig. 2 [3],:

$$(2) \qquad I_{av} \begin{cases} = B_1\, U_{eff} \;\; for\, U_{eff} < U_{cr}/\sqrt{2} \\ = B_2\, U_{eff} + A \;\; for\, U_{eff} > U_{cr}/\sqrt{2} \end{cases} .$$

P. Dineff and *D. Gospodinova* reported that the current-voltage relationship of the atmospheric pressure air glow discharge - *APGD*, was $I_{av} \propto U_{eff}$ for every of both operating areas of relationship - the first operating area being that of the ozone- and oxygen-containing non-equilibrium air plasma, and the second operating area being that of nitrogen oxides (NO_x)-containing non-equilibrium air plasma, Fig. 3 [3]:

$$(3) \qquad I_{av} \left\{ B_0\, U_{eff} \;\; for\, U_{eff} < U_{cr1}/\sqrt{2} \right.$$

$$I_{av} \begin{cases} = B_1\, U_{eff} + A_1 \;\; for\, U_{eff} > U_{cr1}/\sqrt{2} \\ = B_2\, U_{eff} + A_2 \;\; for\, U_{eff} > U_{cr2}/\sqrt{2} \end{cases}$$

In this paper, specific circuit *PSpice* models and *PSpice* simulation of the current-voltage relationship of *APAG* discharge plasma parallel-plate cold plasma generator have been obtained with the proprietary circuit simulation software *PSpice* and compared with experimental data.

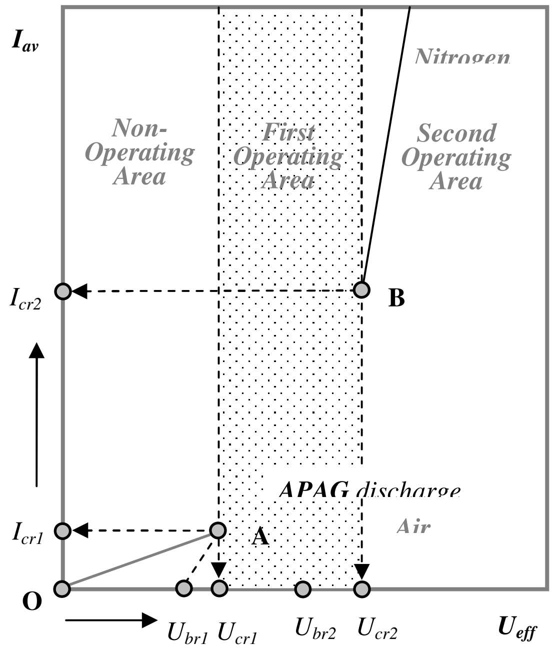


Fig. 3. Linear polynomial model of the current-voltage behavior of an *APAG* discharge in oxygen after *P. Dineff - D. Gospodinova.*

II. PSPICE MODEL FOR SIMULATION OF CURRENT-VOLTAGE RELATIONSHIP FOR PARALLEL-PLATE *APAG* DISCHARGE PLASMA GENERATOR

A parallel-plate *APAG* discharge plasma generator consists of two parallel metal electrode plates with a small gap between them. One of the electrodes is covered with a dielectric plate or coating.

The *APAG* discharge burns at constant drop of voltage U_{br} across the operating gap. *J. Roth* introduced the name "*one-atmosphere uniform glow discharge*" (OAUGD®). This fact is confirmed by multiple researchers [1, 2, 3].

At the same time, however, the current-voltage behavior of a parallel-plate *APAG* discharge plasma generator is governed by a power law in the Roth's model of current-voltage behavior. In order to satisfy the requirement for constancy of the voltage drop across the operating gap during the period of burning of the *APAG* discharge, the current-voltage relationship should vary linearly with the increase of the applied voltage.

This requirement is met in the *Emelyanov*'s linear model of current-voltage behavior of the *APAG* discharge in oxygen.

P. Dineff and *D. Gospodinova* have proposed a linear polynomial model of current-voltage behavior of the *APAG* discharge in air that contains two operating areas corresponding to the elementary processes (dissociation, ionization, and chemical processes) conducted with the participation of oxygen - non-equilibrium ozone- and oxygen-containing plasma, and to the elementary processes conducted with the participation of nitrogen - non-equilibrium nitrogen-oxides-containing plasma, Fig. 3 [3].

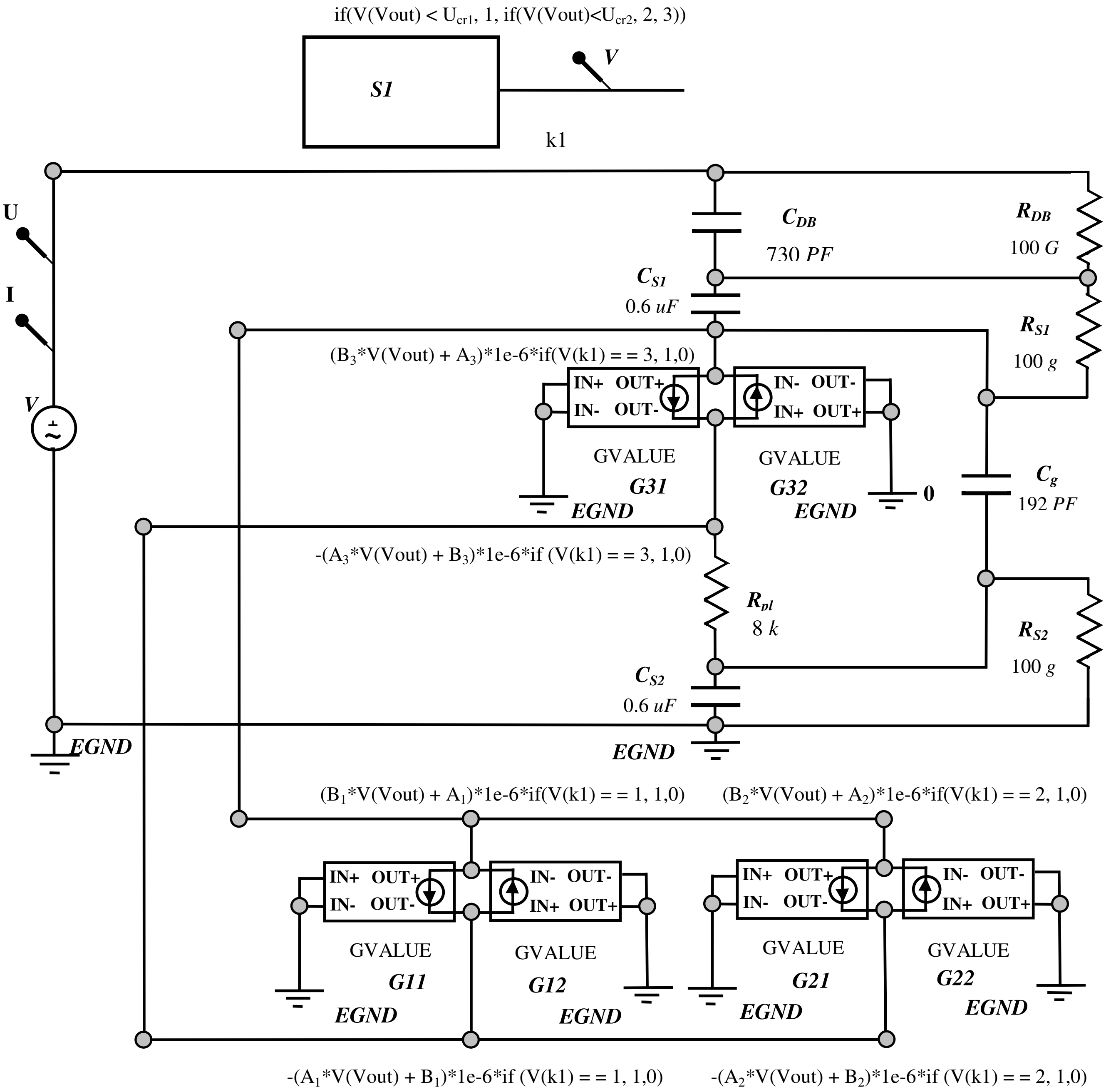


Fig. 4. *PSpice* circuits simulation model for an *APAG* discharge plasma generator current-voltage behavior.

This model reflects adequately the specificity of an *APAG* discharge current-voltage behavior and represents on the basis of equality all areas of the characteristic – the first non-operating area (before discharge ignition) and the two operating areas of burning *APAG* discharge.

All this imposed the creation of a new *PSpice* schematic model for an *APAG* discharge plasma generator, which is governed by the following general relationship reflecting the stationary regime of burning of the *APAG* discharge in both operating areas:

$$(2)\quad \begin{cases} \Delta I = I_{av} - I_{cr1} = B_1\left(U_{eff} - U_{cr1}\right) & for\; U_{cr1} < U_{eff} < U_{cr2} \\ \Delta I = I_{av} - I_{cr2} = B_2\left(U_{eff} - U_{cr2}\right) & for\; U_{eff} > U_{cr}. \end{cases},$$

Based on the phenomenology of *APAG* discharges, the *APAG* plasma discharge in the operating gap can be itself modeled as a voltage-controlled current source that is switched on as long as the voltage across the gap exceeds the value of the plasma ignition voltage U_{br} [3].

The simulation schematic model is constructed on using dependent current sources GVALUE (from the library abm.slb), for which the control is preset as a mathematical expression (linear relationship), Fig. 4.

The control of the three pairs of dependent current sources *G11* and *G12*, *G21* and *G22*, and *G31* and *G32*, one for each half-wave of the harmonically varying voltage of the ideal voltage source *V*, is realized by voltage-controlled switch *S1* (*operator for data statements*): the pair *G11* and *G12* operates when the voltage is below the critical voltage U_{cr1}, the pair *G21* and *G22* when the voltage is below the critical voltage U_{cr2}, and the pair *G31* and *G32* when the voltage is higher than U_{cr2}, Fig. 4.

Table 1. Results from accomplished simulations

Active Power p_A, W	**Active Power** p_A, W	**Active Power** p_A, W	
Experimental current-voltage relationship	*PSpice simulation polynomial linear law model*	*PSpice simulation power law model*	
		$n = 1.6$	$n = 12$
28.1	27.0	38.0	36.0
Relative error, %			
base	**- 3.9**	**+ 35.2**	**+ 28.1**

The values of the critical voltages of discharge ignition, U_{cr1} and U_{cr2}, are entered for each operating area; they have been calculated in a known manner from the current-voltage relationship [3].

The behavior of the model in time (*Transient Response*) as a result of the effect of the voltage across the electrodes that varies harmonically with determinate frequency (50 Hz) with time is investigated, Fig. 6.

An *AC analysis* is assigned additionally by defining the active power p_A by means of the selected format, Table 1.

In accordance with [4], more comparative investigations are performed on the PSpice simulation power-law model of a burning *APAG* discharge, developed by *P. Dineff* and *D. Gospodinova,* for two different values of the power exponent *n*, Table 1.

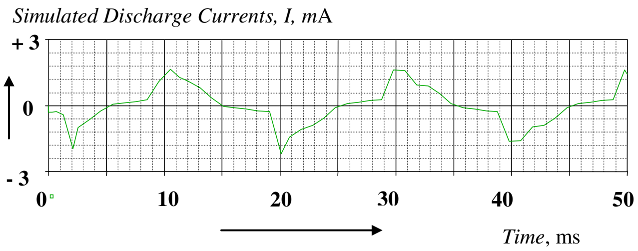


Fig. 6. Simulated discharge currents for an *APAG* discharge plasma parallel-plate generator according to *PSpice* simulation polynomial linear law model (U_{eff} = 12 kV; 50 Hz; δ_{bar} = 3 mm; δ_{gap} = 3 mm).

A voltage U_{eff} is assumed in such a way that simulation will be realized in the transition area between the two operating areas, namely 12 kV, because a maximal relative error in simulation is expected for this transition area.

Results presented in Table 1 demonstrate the improved accuracy (- 3.9 %) of the proposed simulation model of the *APAG* discharge current-voltage behavior. At the same time the lowest value of relative error (+ 35.2 %) is obtained for the highest values of the power exponent - $n = 12$.

III. Discussion and Conclusion

The simulation verifies that the amplitude of the simulated discharge current is determined by four independent variables - the values of the critical parameters - the voltages U_{cr1} and U_{cr2}, currents I_{cr1} and I_{cr2}**;** dielectric barrier capacitance C_{DB}, and applied voltage U_{eff} across the operating gap. It has been found, too, that the polynomial linear law simulation model has a minor relative error compared to the power law simulation model.

We have developed a satisfactory PSpice circuit model for the *APAG* discharge plasma generator current-voltage behavior simulation.

The variant calculations which allow following the effect of parameters of the electrode system or the electric circuit, e. g. those of the matching series capacitor, upon the *APAG* discharge current-voltage behavior for different frequencies may be especially valuable.

Acknowledgement

The National Science Fund, Ministry of Education and Science of Bulgaria, is gratefully acknowledged for the financial support of research project VUF 9/2005 Plasma Assisted Technologies and Devices for Fire Protection of Polymeric and Wood Materials.

References

[1] U. Kogelschatz. Dielectric-Barrier Discharges: Their History, Discharge Physics, and Industrial Application. *Plasma Chemistry and Plasma Processing*, Vol. 23, No 1, March, 2003, pp. 1÷46.

[2] Z. Chen. PSpice Simulation of One Atmosphere Uniform Glow Discharge Plasma (OAUGDP) Reactor Systems. *IEEE Transactions on Plasma Science*, Vol. 31, No 4, August, 2003, pp. 1÷9.

[3] D. Gospodinova. Technological Processing and Apparatus for Plasma Surface Treatment of Materials at Atmospheric Pressure. - *Dissertation Thesis*. Sofia, Technical University, 2006.

[4] E. Gadjeva, T. Kouyomdjiev, and S. Farhi. *Computer Modelling and Simulation of Electronic and Electrical Circuits with OrCAD PSpice*. Sofia, Heron Press, 2001.

[5] P. Dineff, D. Gospodinova, and E. Gadjeva. PSpice Simulation of Atmospheric Pressure Air Gow Discharge Plasma Applicator Systems. *XLI. International Scientific Conference on Information, Communication and Energy Systems and Technologies "ICEST 2006"*, Sofia, Bulgaria, June 29÷ July 01, 2006, Proceedings of Papers (in press).

PSpice Simulation of Atmospheric Pressure Air Glow Discharge Plasma Applicator Systems

Peter D. Dineff[1], Diliana N. Gospodinova[2] and Elisaveta D. Gadjeva[3]

Abstract – **Electrical characteristics of an Atmospheric Pressure Air Glow (*APAG*) discharge plasma applicator system have been simulated by using commercial circuit simulation software such as PSpice®. A plasma applicator system integrally includes a power supply, transformer, impedance matching network, and plasma applicator. A PSpice® model has been developed for the plasma discharge in an *APAG* discharge applicator, which consists of two parallel electrode plates with a small gap between electrodes. At least one of the electrodes is covered by a dielectric barrier.**

An *APAG* discharge plasma applicator can be modeled as an electric capacitor without plasma, and as an air gap containing plasma. The cold plasma itself has been modeled as a voltage-controlled current source that switches on when the voltage across the air gap exceeds the value of the discharge ignition or plasma initiation voltage.

The simulation results agree with experimental data from actual applicators. It has been found that in different operating regimes, the discharge current of the *APAG* discharge plasma applicator is described by a voltage linear law.

Keywords – **Atmospheric pressure air glow (*APAG*) discharge, cold plasma, one-atmosphere glow discharge, discharge ignition voltage, voltage-controlled current source, current-voltage behavior, parallel-plate and coplanar discharge plasma applicators**

I. INTRODUCTION

Atmospheric pressure air glow (*APAG*) discharges, or *dielectric-barrier* (simply *barrier* or *silent*) *discharges*, have been well known for more than a century. First experimental investigations concentrated on the generation of ozone were reported by *W. Siemens* (1857). A few years after Siemens' original publication, *T. Andrews* and *P. Tait* (1860) proposed the name "*silent discharge*" (*stille Entladung*, *décharge silentieuse*) which is still used frequently in the English, German, and French scientific literature, [1].

J.-R. Roth introduced, on the base of the observed phenomenological characteristic of the normal glow discharge voltage-current behavior, the name "*one-atmosphere uniform glow discharge*" (*OAUGD*). Other cases of such behavior have been often reported in the contemporary literature. The normal glow discharges, as the dielectric-barrier discharges, ignite and burn at constant discharge ignition (or burning) voltage, [2].

The characteristics of many electrical systems can be simulated with proprietary computing tools such as *PSpice*®. Devices employing cold plasmas are embedded in electrical systems in many situations. It is advantageous to simulate the complete system, including the plasma as phenomena, with such commercial software. Previous papers concerning the computational simulation of high pressure plasma in air discharge were investigated, [3, 4].

Normally, the *APAG* discharge plasma applicator (or reactor) system for plasma-chemical modification (or functionalization) of low energy surfaces consists of a power supply, transformer, impedance matching network, and plasma applicator. This electrical system was an object of simulation with such computational tools, Fig. 1.

In an electrical discharge system, the inductors and capacitors, in the impedance matching network, the power supply, and the transformer are ordinary electrical components and have well-developed *PSpice* models. However, as there is no available electrical model in *PSpice* for the plasma discharge in an *APAG* discharge plasma applicator, to develop such a model is a principal simulation task, [3, 4, 6].

In this paper, specific circuit *PSpice* models of *APAG* discharge plasma parallel-plate and coplanar applicator systems have been obtained with the proprietary circuit simulation software *PSpice* and compared with experimental data.

II. PSPICE MODELS FOR PARALLEL-PLATE AND COPLANAR *APAG* DISCHARGE PLASMA APPLICATORS

A parallel-plate *APAG* discharge plasma applicator consists of two parallel metal electrode plates, one or two dielectric plates (barriers) with one or two small air gaps between barriers and electrodes. At least one electrode is covered with a dielectric plate or coating, as shown in Fig. 2.

Coplanar *APAG* discharge plasma applicators consist of a flat panel with multiple coplanar plasma electrode strips alternating in polarity, and with each polarity connected in parallel. The flat panel consists of a thin dielectric plate either with electrode strips on one or both sides, or with electrode strips on one side and a metallic electrode sheet on the other. When a voltage is applied across a flat coplanar panel, a planar plate plasma layer is generated either on both sides of the dielectric or only on the side having the electrode strips, depending on the electrode configuration.

[1]Peter D. Dineff is with the Faculty of Electrical Engineering, Technical University of Sofia, Blvd. Saint Kliment Ohridski 8, 1000 Sofia, Bulgaria, E-mail: dineff_pd@abv.bg.

[2]Diliana N. Gospodinova is with the Faculty of Electrical Engineering, Blvd. Saint Kliment Ohridski 8, 1000 Sofia, Bulgaria, E-mail: dilianang@abv.bg

[2]Elisaveta D. Gadjeva is with the Faculty of Electronic Engineering and Technology, Blvd. Saint Kliment Ohridski 8, 1000 Sofia, Bulgaria, E-mail: egadjeva@tu-sofia.bg

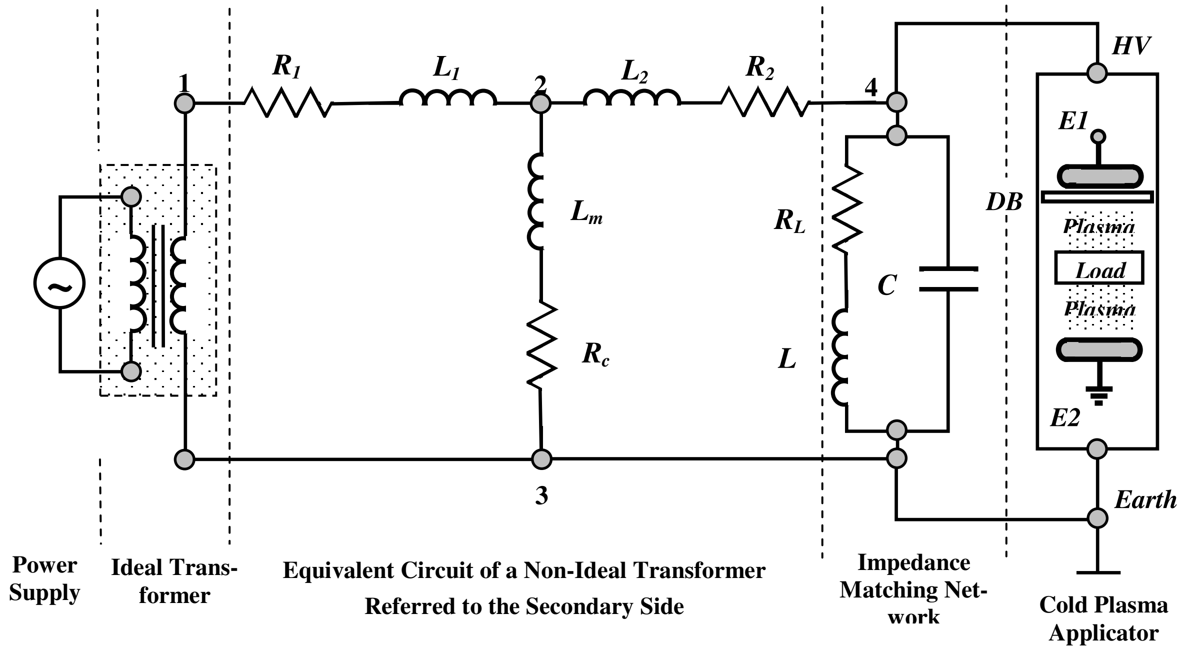


Fig. 1. Schematic presentation of an *APAG* discharge plasma applicator system for *PSpice* circuits simulation

Since the plasma generated on both sides of the coplanar plasma applicator systems, Fig. 3a, resembles the plasma in two-operating-gap parallel-plate applicators, Fig. 2c, we can construct a *PSpice* model for coplanar applicators based on the model for parallel-plate applicators.

Since the plasma energized on one of the sides of the coplanar plasma applicators, Fig. 3b, resembles the plasma in one-dielectric-barrier parallel-plate applicator systems, Fig. 2b, we can also construct a *PSpice* model for coplanar applicators based on the model for parallel-plate applicators.

Z. Chen provides a general plasma circuit *PSpice* simulation model – that of a one-dielectric-barrier discharge plasma applicator, which consists of one operating gap, capable of incorporation into circuit simulations, Fig. 2 [3].

The *PSpice* model for the dielectric plate or coating can be modeled as a capacitor with high parallel resistance, as can the gap containing the cold plasma. Based on the phenomenology of normal glow discharges and *APAG* discharges, Fig. 2, the *APAG* discharge plasma itself can be modeled as a voltage-controlled current source that is switched on as long as the voltage across the operating gap exceeds the value of the discharge ignition voltage [3].

The current source and its output current vary in accordance with a power law of the applied voltage. *R. Gadri* shows that an atmospheric *RF* glow discharge in helium exhibits the same phenomenology as low-pressure *DC* glow discharges.

J. Roth reported that the current-voltage (I-U) behavior of the high-power glow discharge was $I \propto U^2$, and $I \propto U^3$, depending on the operating regime:

$$I \begin{cases} =0,\ for\ V_g < V_{pi} \\ \alpha\left(U_g - U_{pi}\right)^n,\ for\ V_g > V_{pi} \end{cases} , \tag{1}$$

where n is an integer that ranges from 1 to 12 in different glow discharge plasma devices, [3].

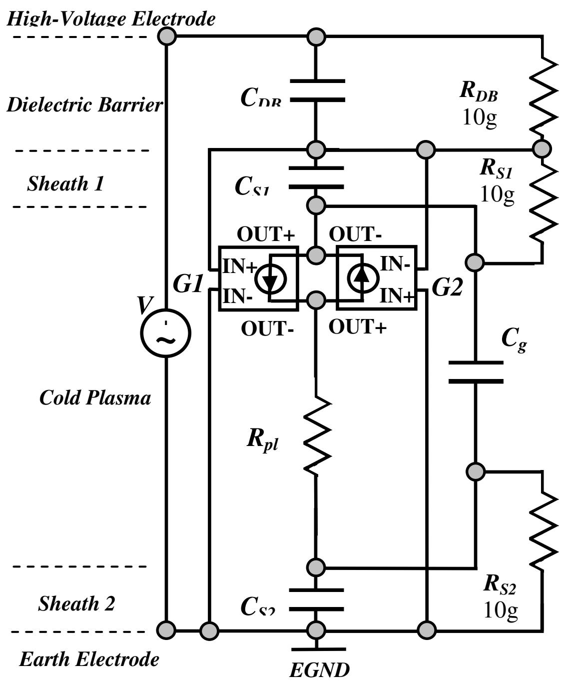


Fig. 2. General model for an *APAG* discharge plasma applicator system (with one dielectric barrier) for *PSpice* circuits simulation - *Roth-Chen*'s model

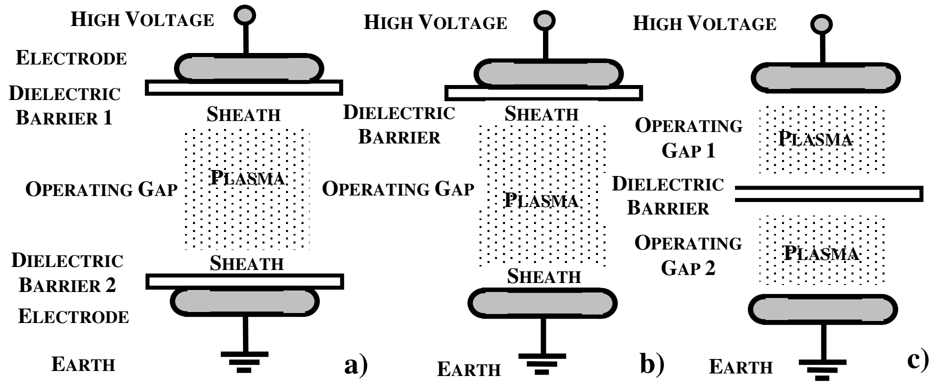


Fig. 3. Parallel-plate *APAG* discharges plasma applicators: **a** – two-dielectric-barriers discharge plasma applicator; **b** – one-dielectric-barrier discharge plasma applicator; **c** – two-operating-gap plasma applicator.

III. NEW PSPICE MODEL FOR *APAG* DISCHARGE PLASMA APPLICATOR SYSTEMS

The adopted simulation of an *APAG* discharge plasma applicator by using the *Roth-Chen's schematic* model, Fig. 2, has demonstrated some negative aspects: for instance, the presence of indeterminacy when the voltage applied across the operating gap is equal to the critical ignition voltage: $U_g < U_{pi}$; negative values of current I at $n = 2\,k - 1$ (odd number); unstable operation of the simulator due to indeterminacy (the upper and lower limits of the variation range are not fixed) of the difference $(U_g - U_{pi})^n$; the ignition current is never equal to zero – it has an exactly determined critical value $I_{cr} > 0$ [4].

All this required the creation of a new *PSpice* schematic model for an *APAG* discharge plasma applicator, which is governed by the following general relationship reflecting the stationary burning regime of an *APAG* discharge [3]:

$$\begin{cases} I_{av} - I_{cr} = 0 \;\; for\, U_{eff} < U_{cr} \\ I_{av} - I_{cr} \;\alpha \left(U_{eff} - U_{cr}\right)^n \;\; for\, U_{eff} > U_{cr} \end{cases}, \tag{2}$$

where $U_{cr} > U_{br}$ (U_{br} is the discharge burning voltage) and I_{cr} are the critical parameters of ignition of an *APAG* discharge.

The simulation model (schematics) is constructed on the basis of using a dependent current source GVALUE (from the library abm.slb - *Analog Behavioral Models*), for which the control may be preset as a mathematical expression connecting the output current and the input variable – an *e.m.f.* formed at the output of another dependent *e.m.f.* source EVALUE, for which the control is also preset as an analytical expression relating the input voltage U_g - ΔU_b, where ΔU_b is the voltage drop across the barrier, to the controlling GVALUE *e.m.f.,* Fig. 5.

Controlling the dependent current sources *G1* and *G2*, one for each half-wave of the harmonically varying voltage of the ideal voltage source *V*, is realized by the dependent *e.m.f.* sources *E1* and *E2* in accordance with Eqn. 2, introducing the value of the critical ignition voltage U_{cr} of the discharge.

The limiting element LIMIT with parameters 0 and 500 is placed at the output of dependent voltage sources. The upper limit is in conformity with the maximal real values that can be assumed by the difference $(U_g - U_{cr})^n$, (voltage is in kV).
The model behavior in time (*Transient Response*) as a reaction to the application of a voltage across the electrodes, which varies with different frequency (50 Hz, 10 kHz, 30 kHz) in time, is investigated.

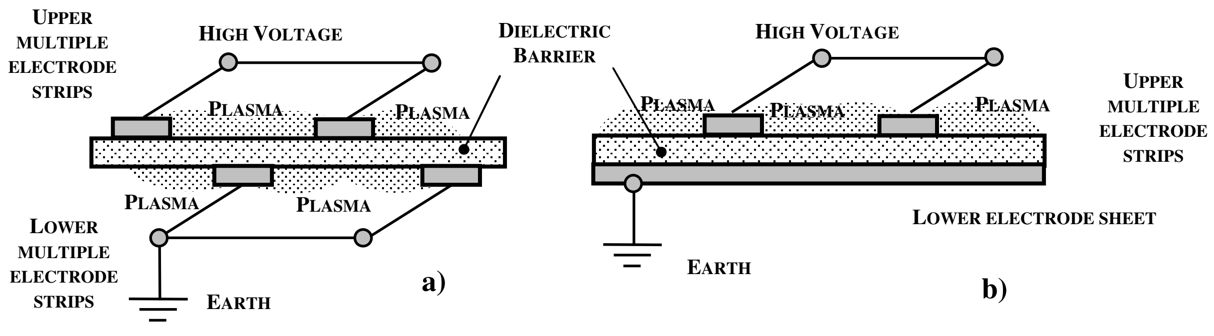


Fig. 4. Coplanar *APAG* discharge plasma applicators: **a** – two-multiple-planar-electrode-strips plasma applicator, or two-(upper and lower)-side plasma-energized applicator; **b** – one-multiple-planar-electrode-strips plasma applicator, or one-(upper)-side plasma-energized applicator.

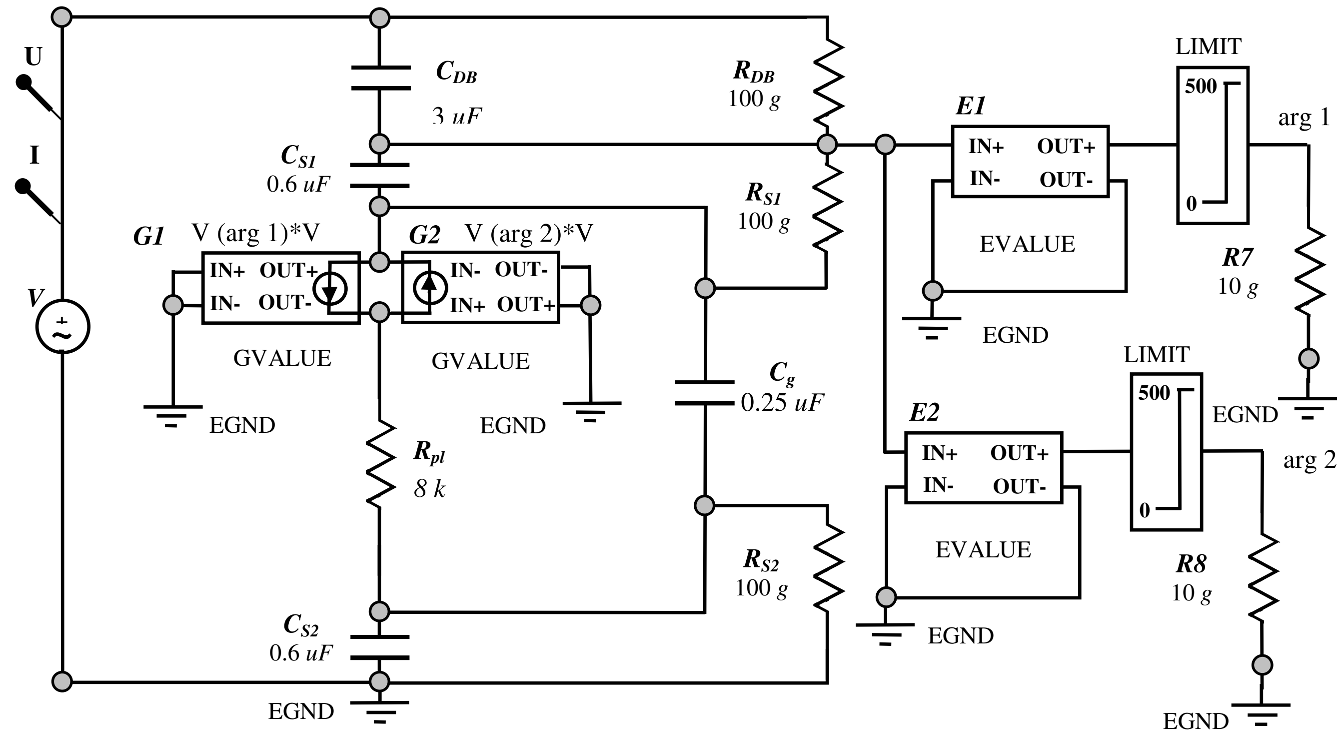


Fig. 5. New *PSpice* circuits simulation model for an *APAG* discharge plasma applicator system (with one dielectric barrier)

IV. Discussion and Conclusion

The simulation verifies that the amplitude of the simulated discharge current is determined by four independent variables - the values of the critical parameters - the voltage U_{cr} and current I_{cr} of ignition**;** the dielectric barrier capacitance, and the applied voltage across the electrodes. It has been found, too, that the power-law exponent n has a minor effect on the modeled discharge current waveform, Fig. 6.

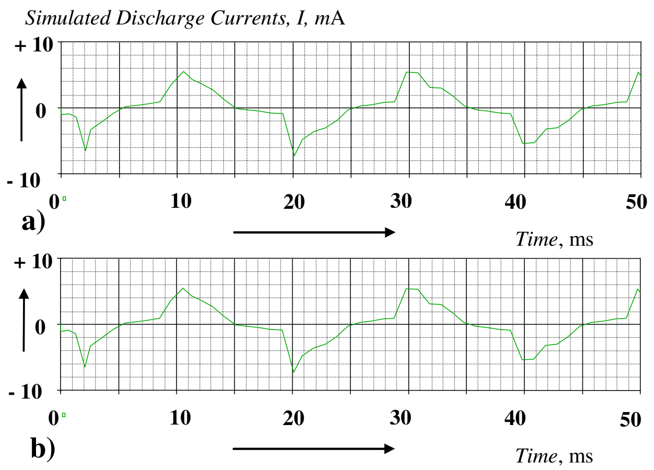


Fig. 6 – Comparison of the simulated discharge currents for an *APAG* discharge plasma applicator system (with one dielectric barrier) with different power-law exponent: **a** - n = 1.4; **b** - n = 12 (U_{eff} = 20 kV; δ_{bar} = 3 mm; δ_{gap} = 3mm).

We have developed a satisfactory PSpice circuit model for *APAG* discharge plasma applicator systems.

Acknowledgement

The National Science Fund, Ministry of Education and Science of Bulgaria, is gratefully acknowledged for the financial support of research project VUF 9/2005 Plasma Assisted Technologies and Devices for Fire Protection of Polymeric and Wood Materials.

References

[1] U. Kogelschatz. Dielectric-barrier Discharges: Their History, Discharge Physics, and Industrial Application. *Plasma Chemistry and Plasma Processing*, Vol. 23, No 1, March, 2003, pp. 1÷46.
[2] Z. Chen. PSpice Simulation of One Atmosphere Uniform Glow Discharge Plasma (OAUGDP) Reactor Systems. *IEEE Transactions on Plasma Science*, Vol. 31, No 4, August, 2003, pp. 1÷9.
[3] D. Gospodinova. Technological Processing and Apparatus for Plasma Surface Treatment of Materials at Atmospheric Pressure. - *Dissertation Thesis*. Sofia, Technical University, 2006.
[4] E. Gadjeva, T. Kouyomdjiev, and S. Farhi. *Computer Modelling and Simulation of Electronic and Electrical Circuits with OrCAD PSpice*. Sofia, Heron Press, 2001.

Stand for Investigation of Hydrogen-Storage Materials

Ruslan Ivanov[1], Viktor Anchev[1], Dimitar Trifonov[2], Seryozha Slavev[2]

Abstract – **A stand for determination of thermodynamic characteristics of hydrogen-storage materials is presented. Structure, action principle as well as a mathematical method for calculation of the quantity absorbed hydrogen, are described.**

Keywords – **hydrogen-storage materials, stand, investigation, thermodynamic characteristics, mathematical method.**

I. INTRODUCTION

During the last years at accelerated rates begins development of new branch in the industry – Hydrogen energetics and technologies [1] [2] [3] [4]. These rates are determined by the comprehensive application of the hydrogen not only like fuel but also like necessary raw material in many technological processes. The hydrogen is the perfect ecological type of fuel but because of the large interval of concentrations, which it forms with the air, arise important technical difficulties for its storage, transportation and distribution. Economical unprofitable and not safe, are the methods for storage and transportation of hydrogen in gas vessels under high pressure or in liquid state in thermal insulated vessels. Alternative way for solving these problems is the use of hydrogen-absorbing materials for storage of big quantities of hydrogen in comparatively small volumes and under low pressure, Fig.1 [5].

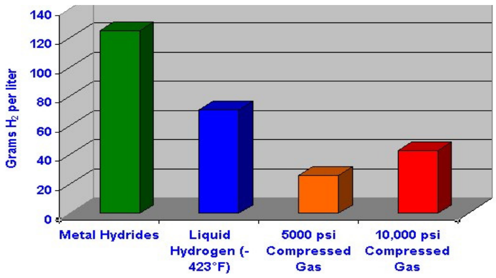


Fig.1 Storage capacity at the separate methods

The process of creating of new materials for storage of hydrogen requires development of methods for determination of their thermodynamic characteristics. In many cases, the used calorimetric and gravimetric testing methods of hydrogen-absorbing materials do not allow the determination of the real operating characteristics of the materials. Similar to the real conditions of exploitation of the reactive hydrogen tanks are the manometric testing methods of hydrogen-absorbing materials [6].

[1]Ruslan D. Ivanov is with the Faculty of Electrical Engineering, Technical University, Kliment Ohridski 8, 1000 Sofia, Bulgaria,
E-mail: rus_ivanov@tu-sofia.bg
[1]Viktor H. Anchev is with the Faculty of Machine Technology, Technical University, Kliment Ohridski 8, 1000 Sofia, Bulgaria,
E-mail: vanchev@tu-sofia.bg
[2]Dimitar I. Trifonov is with the Technovacsystem Ltd, Treti Mart 40, 7000 Rousse, Bulgaria, E-mail: office@technovacsystem.com
[2]Seryozha S. Slavev. is with the Technovacsystem Ltd, Treti Mart 40, 7000 Rousse, Bulgaria, E-mail: office@technovacsystem.com

This paper considers stand for determination of thermodynamic characteristics of materials with hydrogen – absorbing ability, created by teams of Technical University – Sofia and Technovacsystem Ltd by PHARE project [7]. The work of the stand is based on manometric method.

II. STAND FOR INVESTIGATION – CONSTRUCTION AND DESCRIPTION

The worked our stand for investigation of hydrogen-storage materials is shown on Fig. 2. The principle scheme of the stand is shown on Fig. 3. The basic elements of the stand are: reactor; high-vacuum pump aggregate; furnace; tanks for hydrogen with different volumes (R1-R3) and referent vessel for hydrogen (Ref); sensors for temperature (T1-T5) and pressure (P1-P9); electronic calculating machine and digital microprocessor devices for reading temperature and pressure; pipe connections and seals; electro-magnetic valves (V1-V8); reducing valves V2,V3,V4; manual valves (V9-V12). The use of electro-magnetic valves gives the opportunity for subsequent automation of the control process.

Fig. 2 Stand for testing of hydrogen-storage materials

The reactor is one of the critical components of the stand, because it has to give possibility for enough rapid heat transfer to and from the test, and in that way to assure one implementation of the test near to the isotherm, despite the permanent releasing, respectively increasing of the heat from the reactions which take place in the reactor during hydriding respectively dehydriding of the test body. In order to avoid losses of hydrogen and generating a mistake during the experiment, the material from which the reactor is made must not absorb hydrogen. It is made a reactor from chromium-nickel alloy with cylindrical form and volume 6,4 cm^3. The small working volume allows decreasing the time for evacuating of the system as well as the use of small quantity of test body. When the material is used in the form of powder is possible partial sucking of the powder from the vacuum pump during evacuating of the reactor. The detached powder from the test increases the mistakes at the measurement of the quantity of absorbed hydrogen and it causes damage in the vacuum system – vacuum pumps, sensors, valves and etc. With aim of preventing the above-described problems, it is foreseen the initial evacuating of the reactor to be done through needle valve. At the entry of the reactor is placed porous filter plate from sintered material which does not absorb/adsorb hydrogen. The plate prevents the sucking of the test from the vacuum pump, and simultaneously with that the filter assures enough good penetration and going out of hydrogen. For the most precise measurement of the temperature of the test body is implemented measurement of the temperature out of the reactor and the temperature inside of the reactor. For the calculations of the quantity of absorbed hydrogen is used the average value from the two measurements. The measurement of the temperature inside of the reactor is implemented with temperature sensor, placed in the middle of the test in chromium-nickel hood, so that the sensor to be entirely wrapped by the investigated material. The construction of the reactor allows its repeated use. The connecting of the reactor with the system is implemented through threaded connection with copper seal. The reactor is welded in its lower part.

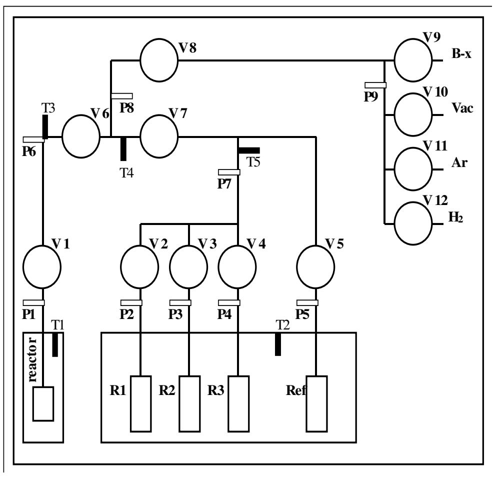


Fig. 3 Scheme of stand for determination of PCT diagrams of metal hydrides

The high-vacuum pump aggregate is consisted of three pumps: rotary pump, roots pump and diffusion pump, and it reaches limit vacuum in the range from 10^{-5} – 10^{-6} Pa. In the conditions of high vacuum and appropriate temperature depending on the material of the test, the oxides are decomposed and active metal surfaces are created, and the kinetics of the absorbing and desorbing process is improved [8].

The referent vessel (Ref) is made of chromium-nickel steel and it has volume 0,3 dm^3.

Depending on the type of the material which will be investigated and with aim of implementing economic profitable experiments, it is necessary in the system to be provided enough quantity of hydrogen. This is done through combination of the vessels for hydrogen (R1-R3) having different volumes 0,3, 0,5 and 1dm^3. The vessels are placed in water bath, in order to minimize the temperature gradient, and also the temporary temperature changes in them during the experiments. The temperatures of the hydrogen in Ref, R1, R2 and R3 are measured with heat-sensitive element placed inside of the relevant tank. During the experiment is obtained information about the temperature of the water bath. The average value of the two measurements is taken at the determination of the hydrogen-absorbing capacity of the test. Although the undertaken measures, at the time of the experiment are obtained minimum changes in the temperature and the pressure of the gas in vessels Ref, R1, R2 and R3 leading to mistakes at calculating of the final quantity of hydrogen, absorbed by the test body. This imposes the introducing of correction coefficient at calculating of the absorbed by the test body hydrogen.

The pipelines of the system are made of stainless steel. The connecting of the separate sections is done through VCR threaded connections. For sealing the connections pipeline – pipeline, pipeline – sensors, pipeline – valves, are used copper seals which prevent the flowing of gas from the system under high pressures - 160 bar.

Sensors Pt100, class B, working temperature to 100°C are used for reading of the temperature. The pressures in each part of the system are measured by sensors for absolutely pressure having the following characteristics – body from stainless steel, working pressure 0-160 Bar, output current 4 -20 mA, exactness 0,25%, working temperature to 85°C. The use of this type of sensors allows specifying the quantity of hydrogen for unit of volume with maximum mistake 0,02 %. The total maximum mistake which is obtained at neglecting the mutual compensating of the mistakes is 0,18 % hydrogen for the whole stand. The information from the sensors is visualized and processed by programmable microcontrollers type TC800. It is foreseen processing and graphic visualization of the results by personal computer, and deducing the results on a paper through printing device. For this aim is used interface type RS485 assuring the communication between the microcontrollers and the electronic calculating machine (ECM). The processing of the information received from the sensors is implemented through software POLIMONITOR. The use of ECM allows the storage of the results of the implemented experiment, which can be used during subsequent experiments. There is a possibility for forming of data base from the results and the methods of implementation of the experiment facilitating the operator's work at subsequent experiment, as well as for analysis of the results.

The used furnace is vertical type. It can maintain temperature up to 700°C and gives opportunity to be investigated high-temperature hydrogen-storage materials. The characterristic of the furnace allows uniformity heating of the reactor and eliminating mistakes as a result of temperature gradient in the test. Its temperature is controlled by microcomputer type RT290.

The system is divided into separate parts which volumes are preliminary determined by using the methods – on the base of isothermal extension of one, accepted for ideal, gas at room temperature and low pressures (< 0.3MPa) from one known referent volume in volume which will be determined. It is specified the volume of the not filled with test body reactor's part, as well as the volume of the valves, the vessels for hydrogen and the referent vessel. The determination of the volume of the referent vessel is done by weighing it in empty state and after filling it with water.

The action principle of the stand with the above-described construction is on the base of the law for storage of the mass in closed isochoric system. The work with the stand is done in the following sequence:

- The test is weighed and it is charged in the reactor;
- The volume of the reactor unoccupied by the test body is determined;
- The test is activated;
- Evacuating of the whole system is done, as the heating of the test begins before reaching vacuum 10^{-5} Pa, until the test is completely desorbed;
- The type of the experiment is determined and the defining of specific for experiment parameters, for example what gas tank to be used (R1-R3), the number of the steps during implementation of the experiment and etc.;
- The desired end pressure in the reactor is determined;
- The tanks are filled with hydrogen;
- The temperature in the furnace is increased to the desired one, depending on the material which is investigated after beginning of each experiment;
- A static absorption (desorption) is implemented - it consists of leading in (leading out) hydrogen by steps: mostly at the beginning of every step one constant volume is filled (emptied) to an exact pressure and depending on the test, so that to absorb (desorb) the relevant quantity of hydrogen. At every step it has to be awaited until reaching approximately thermodynamic equilibrium to which the system approaches asymmetrically. This operation is repeated till the completely hydriding (dehydriding) of the test. The values of the pressure and the temperature are recorded, as well as the quantity of absorbed hydrogen in the beginning and in the end of every step;
- The experiment is stopped after reaching the end assigned pressure and all valves are closed.

The quantity of hydrogen absorbed by the test is calculated on the base of the hydrogen balance of the accepted closed isochoric equipment: when neglecting the losses of hydrogen the value of the quantity of hydrogen in the equipment during each experiment remains constant. This means that the changes of the quantity of absorbed hydrogen by the test, cause relevant precise defined changes in the hydrogen pressure. With known working volumes of the equipment and through measuring of relevant temperatures and pressures, and also by using an appropriate equation for the state of the hydrogen gas, is possible to calculate the quantity of hydrogen absorbed by the test at any time. The quantity of hydrogen absorbed by the test is calculated by reckoning the hydrogen content in every separate section of the system before and after each step of the absorption or desorption. For this purpose is used the equation:

$$p \cdot V = n \cdot R \cdot T \qquad (1)$$

where:

p is the pressure of the gas in the respective module of the system, Pa;

V – volume of the respective module of the system, m^3;

n – quantity gas, mol;

R – gas constant

T – temperature, K

The total quantity of gas in the system in definite moment is:

$$n_\Sigma = \sum_{i=1}^{k} n_i = \sum_{i=1}^{k} \frac{p_i(t) \cdot V_i}{R \cdot T_i(t)} \qquad (2)$$

where k is the number of the working sections of the system.

The total quantity of hydrogen absorbed by the test for one step is:

$$n_x = (n_{\Sigma 1} - n_{\Sigma 2}) - n_{korr} \qquad (3)$$

where:

$n_{\Sigma 1}$ – the quantity of hydrogen in the working sections in the beginning of every step;

$n_{\Sigma 2}$ – the quantity of hydrogen in the working sections at the end of every step;

n_{korr} – coefficient of mistake;

n_x – is the difference between the quantity of gas before and after the absorption, i.e. in the beginning and at the end of every step. It determines the quantity of hydrogen which is absorbed by the test.

The total quantity of hydrogen which is absorbed by the alloy is an amount of the quantity of hydrogen absorbed during each step at absorption. In analogous way is calculated the quantity of deabsorbed hydrogen at static deabsorption.

The correction coefficient n_{korr} is introduced at calculating of n_x eliminating the mistakes as a result of reading the change of the pressures and the temperatures in the separate parts of the system during the experiment, and also in consequence of inaccurate measuring of the relevant quantities.

III. Conclusion

With the created stand can be investigated hydrogen-absorbing materials with enough for the research work exactness. The presented stand assures a possibility of refueling hydrogen reservoirs on the base of hydrogen-absorbing materials. Thus the presented installation gives an opportunity for static trial the tests in conditions of absorption and desorption but it is foreseen a possibility for additional building in of regulator for the flow – allowing dynamic investigation of the experimental model. The construction of the stand allows automation of the control.

References

[1] E. Tzimas, C. Filiou, S.D. Peteves, *Hydrogen Storage:state of art and future perspective*, Petten The Netherlands, 2003.

[2] HILTech, *Hydrogen and storing hydrogen*, 2005.

[3] M. Dorheim, T. Klassen, R. Bormann, *Hydrogen storage materials*, Hamburg, Technical University Hambur –Harburg, 2004.

[4] www.whitehouse.gov/infocus/energy/

[5] www.ovonic-hydrogen.com/solutions/technology.htm

[6] Jurgen U. Keller, Erich Robens, *Thermogravimetric and sorption measurement techniques/instruments.*

[7] Dimitar Trifonov, project PHARE, *Technologies for obtaining new materials with hydrogen-storage capacity*, BG 0102.02.02.043, 2003-2005.

[8] Gerardo Friedlmeier, *Charakterisierung von hochtemperatur-metall-hydriden auf magnesium-basis*, 1997.

Calculation of Electromechanical Characteristics on Overband Magnetic Separator with Finite Elements

Mirka Popnikolova Radevska[1], Violeta Manevska[2] and Kristijan Petkovski[3]

Abstract – **In this paper will be presented an approach to improved nonlinear magnetic field analyses of the Overband Magnetic Separator, on the basis of FEM as a represent of the numerical methods. By using iterative procedure Finite Element Method, it will be calculated the nonlinear distribution of magnetic field, under rated excitation of the Separator. The electromagnetic field on the basis of the fluxes and flux densities in the particular domain of the separator will be defined. The electromagnetic forces on the x and y directions will be calculated.**

Keywords – **Electromechanical characteristics on overband magnetic separator.**

I. Introduction

The development of different Overband Magnetic Separator, has caused an appearance and an expansion of special different types of Magnetic Separator with enormous possibilities for their application. One of these particular is the Overband Magnetic Separator OMS. The exact analytic methods are almost improper.

The rated data of the Overband Magnetic Separator OMB which is going to be analyzed in this paper are: 7 kW input power and 220 V operating voltage. Maximum clearance distance is x=420 mm.

The cross-section of the separator and its dimensions are presented in Fig. 1.

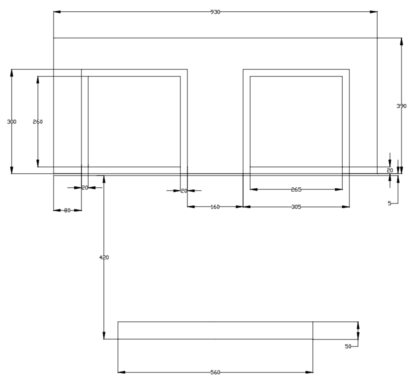


Fig. 1 Cross-section and dimensions on Overband Magnetic Separator

[1]Mirka Popnikolova Radevska is with the Faculty of Technical Engineering, “St. Climent Ohridski” University, Ivo Lola Ribar b.b., 7000 Bitola, Macedonia, E-mail: mirkara@mt.net.mk

[2]Violeta Manevska is with the Faculty of Technical Engineering, “St. Climent Ohridski” University, Ivo Lola Ribar b.b., 7000 Bitola, Macedonia, E-mail: violeta.manevska@uklo.edu.mk

[3]Kristijan Petkovski is with the Faculty of Technical Engineering, “St. Climent Ohridski” University, Ivo Lola Ribar b.b., 7000 Bitola, Macedonia, E-mail: kristijan.petkovski@uklo.edu.mk

In order to determine electromagnetic characteristics of the Overband Magnetic Separator as accurate as possible, the Finite Element Method (FEM) is used. The nonlinear interacttive procedure is applied. The calculation are carried out quasistatically, at given band position.

At the beginning, FEM is used at separately energized windings. The calculations continue when two couple winding are energized with rated current as well as with several other values.

Flux density in different Overband Magnetic Separator parts is calculated. Additionally electromagnetic forces and torque is calculated.

II. Modeling of Electrmagnetic Field

A. Preprocessing

FEM 4.0 has possibility to solve magnetic vector potential and consequently magnetic flux density by solving relevant set of Maxwell equations for magnetostatic case as well as for time harmonic case. In magnetostatic case field intensity $\mathbf{H}$ and flux density $\mathbf{B}$ must obey:

$$\nabla\, x\, \mathbf{H} = \mathbf{J} \qquad (1)$$

$$\nabla\, x\, \mathbf{B} = 0 \qquad (2)$$

subject to a constitute relation between B and H for each material:

$$\mathbf{B} = \mu\, \mathbf{H} \qquad (3)$$

and for nonlinear material (saturating iron) permeability μ is actually function of $\mathbf{B}$.

FEM goes about finding a field that satisfies Eqs. (1)-(3) via a magnetic vector potential. Flux density is written in terms of the vector potential $\mathbf{A}$, as:

$$\mathbf{B} = \nabla x \mathbf{A} \qquad (4)$$

This definition of $\mathbf{B}$ always satisfies Eq. (2). Then Eq. (1) can be rewritten as:

$$\nabla\, x \left(\frac{1}{\mu(\mathbf{B})} \nabla\, \mathbf{A} \right) = \mathbf{J} \qquad (5)$$

The advantage of using the vector potential formulation is that all the conditions to be satisfied have been combined into a single equitation. If **A** is found, **B** and **H** can be deduced by differentiating **A** .

In order to determine electromagnetic characteristics, in FEM pre-processing separator geometry should be first input. Afterwards all materials in all separator domains must be defined (Fig. 2).

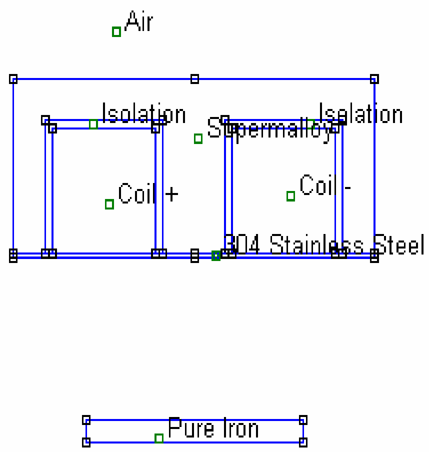


Fig. 2 Defining of material properties in different area of the OMS

This is enabled with program blocks which contain all necessary information regarding electrical and magnetic materials including magnetization curve (Fig. 3), as well as material conductivity.

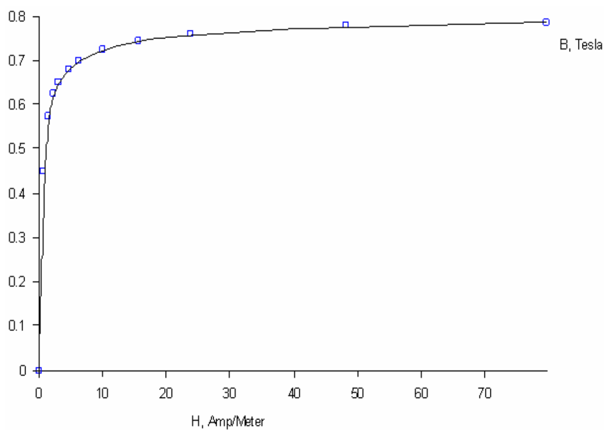


Fig. 3 Nonlinear magnetization characteristic

Afterwards a mesh of finite elements is generated. In this particular case mesh contains 31 418 nodes and 62480 finite elements (Fig.4).

In case when calculation of magnetic vector potential should be calculated more accurate mesh density should be increased especially on interface between two different materials.

In that case contour of integration passes at least two elements away from any interface or boundaries. Greater mesh density increases the computation time. So the good way to find mesh which is "dense enough" in order necessary accuracy to be achieved and still computation time to be reasonably small is comparation of results from different mesh densities. Than can be picked smallest mesh which gives convergence to the desired digit of accuracy.

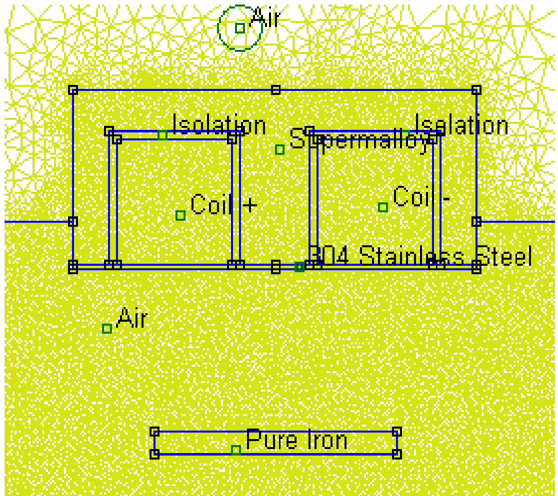


Fig. 4 Detailed cross section of the separator with increased mesh density

B. Post processing

Quasistatic analyses of the Separator at load condition is carried out in rated current.

Nonlinear iterative procedure enables "to enter inside the separator" and to carry out a deepened analyses of magnetic field distribution, taking into consideration the typical magnetic values, as magnetic flux and magnetic flux density.

From the flux distribution when the calculations continues when two couple winding are energized with rated current is presented in (Fig. 5).

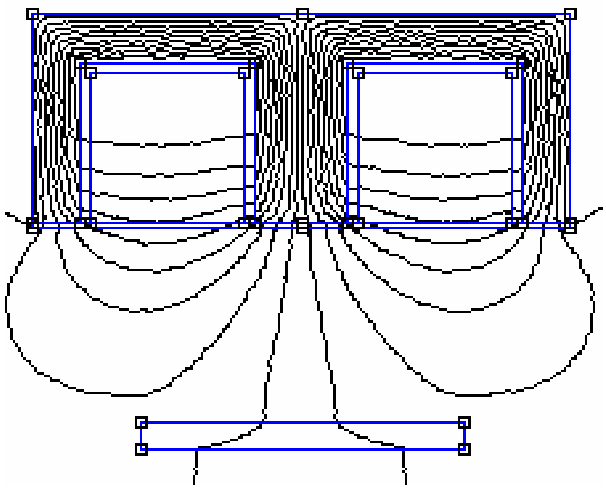


Fig. 5 Magnetic field distribution on Overband Magnetic Separator

On Fig. 6 are presented values of magnetic flux density when the coil is excited with rated current.

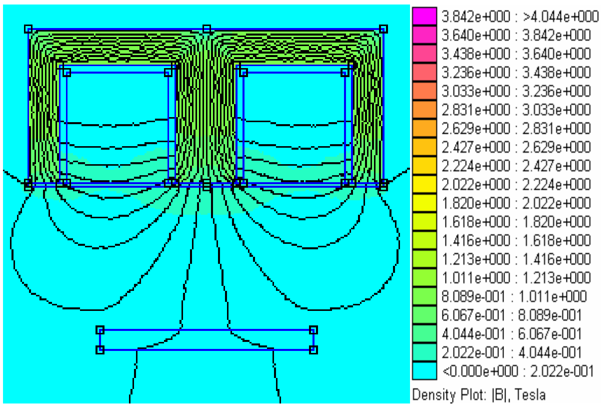


Fig. 6 Magnetic field density on Overband Magnetic Separator

The distribution of magnetic flux density when clearance distances are x=0 mm and x=420 mm is presented on Fig.7 and Fig. 8.

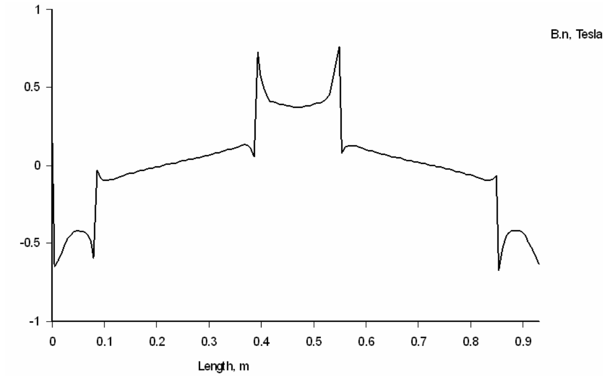


Fig.7 Distribution of flux density when clearance distance is x=0 mm.

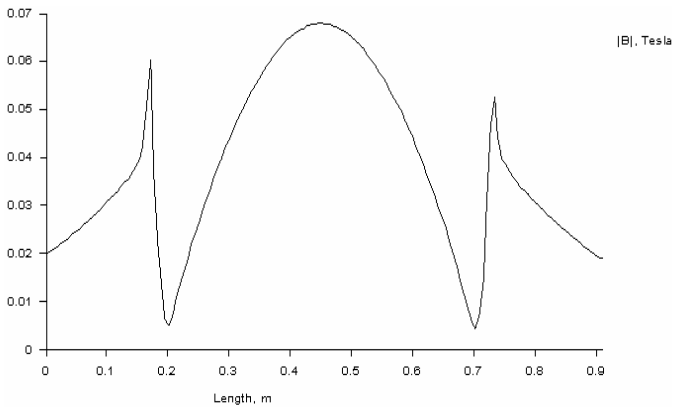


Fig. 8 Distribution of magnetic flux density when clearance distance is x=420 mm.

In this version of the FEM programme, this volume integral greatly simplifies the computation of forces and torques. Mearly select the blocks upon which force or torque are to be computed and evaluate the integral. No particular “art” is required in getting good force or torque results.

On limitation of the Weighted Stress Tensor integral is that the regions upon which the force is being computed must be entierly surrounded by air and/or abutting a boundary.

The differential force produced is:

$$dF = \frac{1}{2}\left(H(B \cdot n) + B(H \cdot n) - (H \cdot B)n\right) \tag{6}$$

where n denotes the direction normal to the surface at the point of interest. The net force on an object is obtained by creating a surface totally enclosing the object of interest and integrating the magnetic stress over that surface Fig. 9.

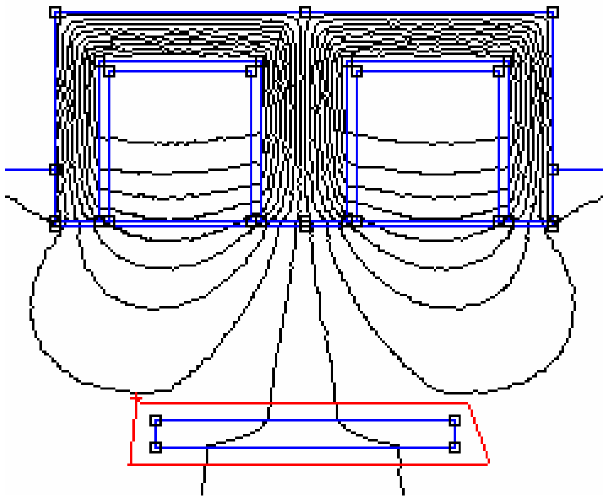


Fig. 9 The regions upon which the force are calculated

The electromagnetic forces in x-direction = 7.25658 N, and force in y- direction = 450.063N.

The upshot is that if stress tensor is evaluated on the interface between two different materials, the results be particularly erroneous. However, the stress tensor has the property that, for an exact solution, the same result is obtained regardless of the path of integration, as long as that path encircles the body of interest and passes only through air (or at least, every point in the contour is in a region with a constant permeability).

The second rule of getting good force result is used as fine a mesh as possible in problems where force results are desired.

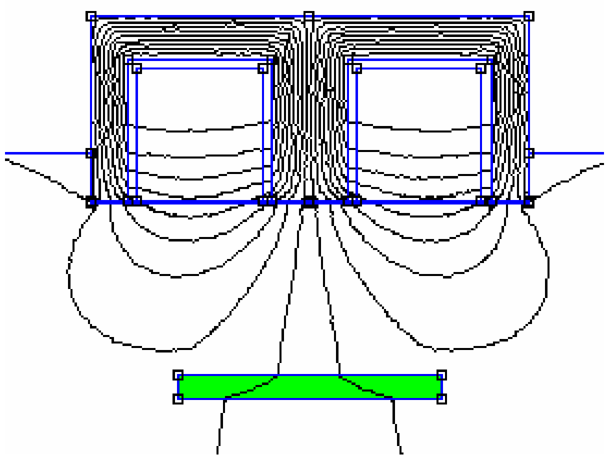


Fig. 10 Force and torque calculated via weighted stress tensor integral

The electromagnetic forces in x-direction =7.25658 N, and force in y- direction=450.063N.

The value of the electromagnetic torque about (0, 0) =3.12 Nm.

III. CONCLUSION

In this paper the non-linear magnetic field analyses and computation of electromagnetic, electromechanical characterristics are presented. For this purpose as the most suitable Finite Element Method is applied. This contemporary method enables exact magnetic quantities such as flux or flux density distribution to be evaluated in any part of the separator. Additionally electromagnetic force can be calculated for rated load current.

REFERENCES

[1]. M. Popnikolova-Radevska, M. Cundev, L. Petkovska, "Modeling of Three-Dimensional Magnetic Field in Solid Salient Poles Synchronous Motor", *ISTET International Symposium on Electromagnetic Fields in Electrical Engineering*, pp. 70-73, Thessaloniki, Greece, 1995.

[2]. M. Cundev, L. Petkovska, M. Popnikolova-Radevska, "Nonlinear Magnetic Field Analyses in Electrical Machines Synchronous Type", *ISEM'95*, paper A-29, Cardif, Great Bitain, 1995.

[3]. M. Popnikolova-Radevska, M. Cundev, L. Petkovska, "An Analyses of Electrical Machines Synchronous Type Based on Electromagnetic Field Analyses and Computation of Electromagnetic Characteristics of the Solid Salient Poles Synchronous Motor", *ICEM'98*, pp. 707-709, Turkey, 1998.

[4]. M. Popnikolova-Radevska, M. Cundev, L. Petkovska, "Modeling the Configuration of Solid Salient Poles Synchronous Motor for 3D-FEM", *Scientific Bulletin of Lodz Technical Univeristy*, *Nr. 788. Elektrika*, pp. 178-186, Lodz, Poland, 1998,

[5]. M. Popnikolova-Radevska, V.Sarac, M. Cundev, L. Petkovska, "Computation of Solid Salient Poles Synchronous Motor Parameters by 3D-Finite Element Method ", *EPNC Symposium on Electromagnetic Phenomena in Nonlinear Circuits*, p.p. 111-114, Leuven, Belgium, 2002.

[6]. D. Meeker, "Finite Element Method Magnetics Version 4.0 ", *Users Manuel*, 2004.

Bolted Busbar Connections with Particularly Slotted Bolt Holes

Raina T. Tzeneva[1] and Peter D. Dineff[2]

Abstract – **The paper discusses how introducing particularly slotted bolt holes in bolted busbar connections could increase significantly the true contact area and reduce the contact resistance. The profile presented has 4 slots with circular holes at the ends. The slots are 3mm long and lie on mutually perpendicular axes, rotated at an angle of 45 degrees in relation to the busbar axes. This hole shape is compared with the classical one of bolted busbar connections by the help of several computer models. It has been established that the new shape case leads to a considerable increase in the contact pressure and penetration in the contact area between the busbars.**

Keywords – **bolted busbar connections, new hole shape, contact resistance, contact pressure, contact penetration**

I. INTRODUCTION

The boost to transmission and distribution lines loading is due primarily to the amplified consumption of electrical energy worldwide. This explains the elevated reliability and service requirements of high power connections in the power-utility industry. The fundamentals to reliable bare overhead high-power line connections design are given in [1] and include: maximization of true area of electrical contact, optimization of frictional forces, minimization of creep and stress relaxation, minimization of fretting and galvanic corrosion, minimization of differential thermal expansion along and normal to interfaces. Summarizing the major criteria, it is worthwhile to note that all of them can be met simultaneously if an outline could achieve a sufficiently large contact load and metal-to-metal contact area (increased number and area of α-spots) as well as sufficient elastic energy storage in the connection so as to maintain acceptable load intensity throughout the connection service life.

! THE OBJECTIVE of this study is to analyze the influence that four bolt hole slots with circular ends exert on the busbars contact penetration and pressure, when the slots are arranged in pairs on mutually perpendicular axes, rotated at an angle of 45 degrees in relation to the busbar axes.

II. THEORETICAL BACKGROUND

All joint surfaces are rough and their surface topography shows summits and valleys. Thus under the joint force F two joint surfaces get into mechanical contact only at their surface summits. Electrical current lines are highly constricted at the contact spots when passing through, as presented schematically in Fig. 1a. This constriction amplifies the electric flow resistance and hence the power loss. Obviously, the more the contact spots, the smaller the power loss at the interface of the conductors. Power connections with superior performance are designed to maximize both the number and the life of the contact spots

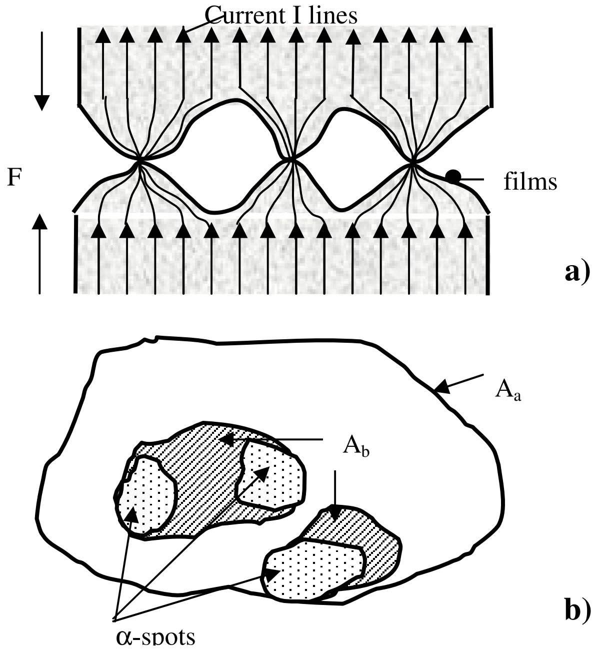


Fig. 1. a) Contact interface and the constricted current lines; b) Apparent contact area A_a, bearing area A_b with quasimetallic contact zones and α-spots

Hence the load bearing area A_b in an electric joint measures only a fraction of the overlapping, the so called apparent area A_a. Metal surfaces, e.g. those of copper conductors, are often covered with oxide or other insulating layers. As a result the load bearing area A_b may enclose regions that do not contribute to the current flow. So a fraction of A_b may have metallic or quasimetallic contact and the real electric contact area A_c, i.e. the conducting area, could be smaller than the load bearing area A_b (Fig. 1b) [2].

A conducting area is referred to as quasimetallic when it is covered with a thin (< 20 Å) film that can be tunneled through by electrons. This quasimetallic electric contact results in a relatively small film resistance R_F.

The summits of the two electric joint surfaces, being in metallic or quasi-metallic contact, form the so called α-spots where the current lines bundle together causing the constriction resistance R_c. The number n, the shape and the

[1]Raina T. Tzeneva is with the Department of Electric Apparatus, Faculty of Electric Engineering, Kliment Ohridski 8, 1000 Sofia, Bulgaria, E-mail: tzeneva@tu-sofia.bg

[2]Peter Dineff is with the Department of Electric Apparatus, Faculty of Electric Engineering, Technical University of Sofia, Kliment Ohridski 8, 1000 Sofia, Bulgaria, E-mail: dineff_pd@abv.bg.

area of the α-spots are generally stochastic and depend on the material parameters of the conductor, the topography of the joint surfaces and the joint force. For simplicity it is often assumed that the α-spots are circular. Looking at one such spot, its constriction resistance R_c depends on its radius *a* and the resistivity ρ of the conductor material. Under the assumption that the bulk material above and under the α-spot is infinite in volume, the value of the constriction resistance can be calculated by means of the Holm´s ellipsoid model

$$R_c = \frac{\rho}{2a} \tag{1}$$

If a single α-spot is completely covered with a thin film, of resistivity ρ_f and thickness *s,* its film resistance R_f is given by

$$R_f = \frac{\rho_f s}{\pi a^2} = \frac{\sigma_f}{\pi a^2} \tag{2}$$

where σ_f is the tunnel resistivity i.e. the resistance of the film across one cm².

The total resistance R_1 of an α-spot, referred to as contact resistance, results in the sum of the constriction resistance R_c and the film resistance R_f

$$R_1 = R_c + R_f \tag{3}$$

III. MODELING THE BOLTED BUSBAR CONNECTION

If the deeper contact penetration increases α-spots both in numbers and dimensions, which in turn expands the true contact area and decreases contact resistance, then a new hole-shape could be introduced for this connection. A typical bolted busbar connection is shown in Fig. 2.

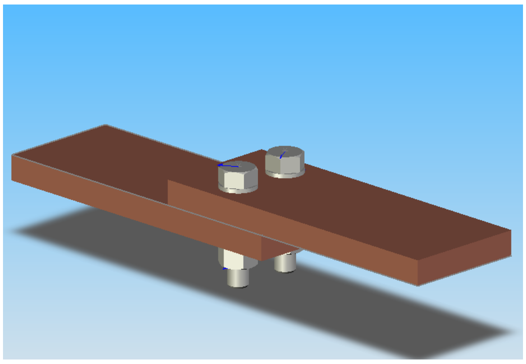


Fig. 2. A bolted busbar connection

The profile scheme of small circular holes at the slot ends arises from [3]. This new case is modeled and compared with the classical one in [4].

For that purpose there have been investigated 11 different models. The first one is a classical one of two copper busbars with 2 bolts (case 1). The next cases introduce the new concept where all bolt holes have four slots, of length 3mm arranged in such a way that the pairs of slots lie on mutually perpendicular axes that are rotated at an angle of 45 degrees in relation to the busbar axes. The slots end with small circular holes. This new shape form is presented in Fig. 3.

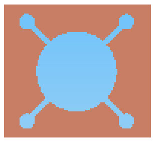


Fig. 3. New hole shape for bolts in bolted busbar connections

Table I describes the 11 investigated cases of different slot width and radius of the small circular holes.

TABLE I INVESTIGATED CASES

Case No	1	2	3	4	5	6	7	8	9	10	11
Slot width, mm	0	0.3	0.3	0.3	0.3	0.5	0.5	0.5	0.7	0.7	1.0
Radius of the small hole, mm	0	0.3	0.5	0.7	1.0	0..5	0.7	1.0	0.7	1.0	1.0

This is suggested to decrease radial loadings on bolts that emerge after the connection is assembled, increase the contact penetration in the busbars near the bolts area and maximize the true area of metal to metal contact at the electrical interface.

Copper busbars (Young's modulus E = $1.1.10^{11}$Pa, Poisson's ratio μ = 0.34, width 60mm, height 10mm, length 160mm, busbars' overlap 60mm) are investigated. The connection has 2 holes of ∅10.5mm.

Fasteners: bolts – Hex Bolt Grade B_ISO 4015 – M10 x 40 x 40 – N, steel E = 2.10^{11}Pa, μ = 0.3; nuts – Hex Nut Style1 Grade AB_ISO 4032 – M10 – W – N, steel E = 2.10^{11}Pa, μ = 0.3; washers – Plain Washer Small Grade A_ISO 7092 – 10, steel E = 2.10^{11}Pa, μ = 0.3. Tension in each bolt F = 15000N.

Models for the contact pressure and penetration in the electrical interface between busbars are obtained using the software package ANSYS Workbench 9.

The aspect of model meshing is distinguished as a key phase for proper analysis of the problem. This is because on the one hand it is an established certainty that the reason for the good quality of physical space triangulation is closely related to the consistent mapping between parametric and physical space. On the other hand a properly meshed model will present a fairly close-to-reality detailed picture of stress distributions which is a hard task for analytical solution and is usually an averaged value. It is evident from Fig.4 and Fig.5, for the uneven allocation of pressure and penetration, that the slotted cases bring even more complexities.

The meshed model incorporates the following elements: 10-Node Quadratic Tetrahedron, 20-Node Quadratic Hexahedron and 20-Node Quadratic Wedge. Contacts are meshed with Quadratic Quadrilateral (or Triangular) Contact and Target elements.

The 10-Node Quadratic Tetrahedron element is the basic mesh constituent. It is defined by 10 nodes with corresponding *x*, *y* and *z* translations and is recognized as well suited to mesh irregular geometries especially the ones produced by CAD systems like Solid Works, the CAD-system employed here.

The 20-Node Quadratic Hexahedron and Wedge type elements are used for meshing the washers. Like the previously described element, this one possesses the same features but offers 20 nodes that allow any spatial orientation. It is assumed that using it to mesh the washers guarantees full transfer of washer deformations and loading to the busbars, the primary researched target.

In order to represent contacting and sliding between the 3-D surfaces, the Quadratic Quadrilateral (or Triangular) contact element is applied. With an 8-node surface-to-surface contact capability and taking the same parameters as the solid element to which it is connected, this contact element is given the presumed behavior and friction of the contact region. Busbars are selected to be in a bonded type contact so as not to overload the hardware with calculations and at the same time ensure a theoretically full area contact with no initial penetrations. In such a way the study could search for the percent enlargement of the busbar contact area with higher penetration and compare this value for the new slotted cases with the one for the classical not slotted case.

The Quadratic Quadrilateral (or Triangular) target segment element associates with the corresponding contact element and is used due to the 3-D behavior it offers.

Fig. 4 shows contact pressure for case 2. It is obvious that the pressure in the area surrounding the slots is increased significantly.

Contact penetration for case 10 is presented in Fig.5. The 4 particularly slotted bolt holes provide the extended high penetration zone that covers the area around the hole between the slots.

The eleven cases have been evaluated by comparing the max values of pressure and penetration for each one of them as well as the percent participation of the 8 zones according to the legends. With that end in view, all zones are set to have equal upper and lower limits. The zones of the highest pressure or penetration are set to equal lower limits while the max values define their upper limits. This comparison procedure is performed by the help of the Adobe Photoshop software, where each colored zone is identified with a certain number of pixels. The results obtained are summarized in Table II and Table III.

TABLE II CONTACT PRESSURE RESULTS

Case No	p_{max}, MPa	3.02 - 8.14, %	8.14 - 13.26, %	13.26 - 18.38, %	18.38 -23.5, %	23.5 - 28.62, %	28.62 - 33.74, %	33.74 - 38.86, %	>38.86, %
1.	43.99	21.82	15.39	13.46	11.18	12.24	12.66	11.08	2.17
2.	55.01	21.89	15.59	13.39	10.02	8.96	8.62	7.98	13.55
3.	74.3	19.89	13.94	11.88	10.12	8.36	6.58	5.54	23.69
4.	74.06	20.92	14.52	11.37	8.9	8.24	6.28	6.03	23.74
5.	71.22	19.24	12.85	11.44	10.27	8.13	6.66	6.27	25.14
6.	85.71	21.56	14.83	12.98	9.17	7.98	7.94	6.5	19.04
7.	87.04	19.25	12.22	10.68	9.84	8.64	7.11	6.47	25.79
8.	74.38	18.321	12.81	10.1	9.64	8.6	7.06	6.6	26.87
9.	93.13	19.83	13.89	11.28	11.07	9.11	7.07	7.16	20.59
10.	91.02	18.19	12.52	11.43	10.22	8.31	6.33	5.89	27.11
11.	92.76	19.58	12.47	10.63	10.74	8.25	6.96	6.18	25.19

TABLE III CONTACT PENETRATION RESULTS

Case No	μ_{max}, µm	0.0084 - 0.0168, %	0.0168 - 0.0252, %	0.0252 - 0.0337, %	0.0337 - 0.0421, %	0.0421 - 0.0505, %	0.0505 - 0.0589, %	0.0589 - 0.0673, %	>0.0673, %
1.	0.076	20.21	15.23	13.38	11.79	12.85	13.27	11.18	2.09
2.	0.115	17.33	12.94	12.49	9.28	7.88	7.39	7.25	25.44
3.	0.278	12.13	7.6	7.46	5.68	4.39	4.93	4.85	52.96
4.	0.234	14.12	8.61	8.06	6.75	6.06	5.36	4.46	46.58
5.	0.201	13.58	8.88	7.83	6.51	6.15	6.63	5.28	45.14
6.	0.327	12.57	8.27	7.59	5.86	5.37	6.26	4.82	49.26
7.	0.283	12.42	8.34	7.29	5.37	4.96	5.55	4.48	51.59
8.	0.207	13.59	8.4	8.06	5.91	5.33	6.13	5.35	47.23
9.	0.296	22.56	7.2	6.99	5.8	4.99	5.31	4.63	42.52
10.	0.266	12.68	8.17	7.44	6	6.08	6.06	5.64	47.93
11.	0.296	13.11	8.01	7.27	5.65	5.15	5.74	4.88	50.19

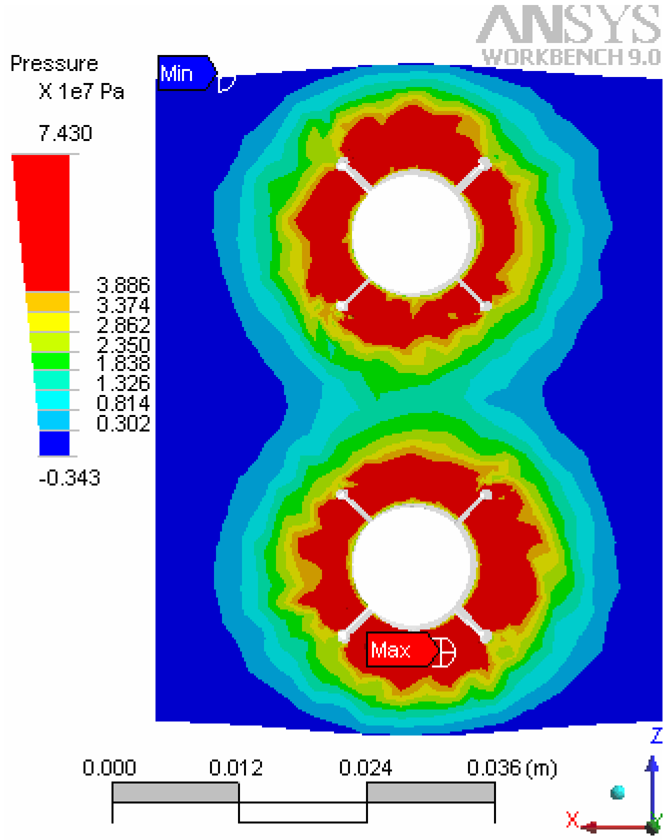


Fig. 4. Contact pressure for case 2 (new hole shape with 4 slots of length 3mm and width 0.3mm and small circular holes at the end of the slots of radius 0.5mm)

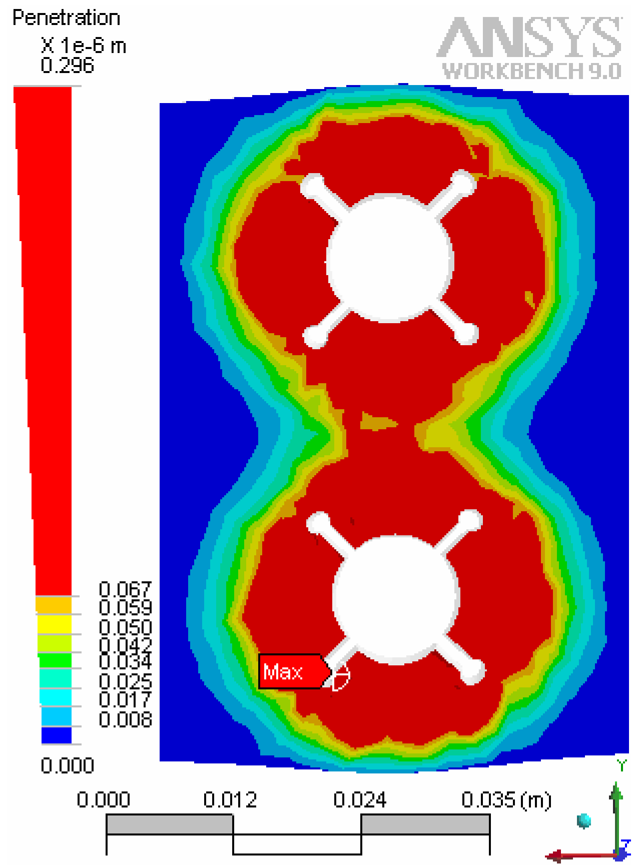


Fig. 5. Contact penetration for case 10 (new hole shape with 4 slots of length 3mm and width 1.0 mm and small circular holes at the end of the slots of radius 1.0 mm)

IV. Results and Discussion

When bolt holes of bolted busbar connections have slots ending with circular openings, then a zone of considerably high contact penetration and pressure concentration around the slots is observed. It is confirmed by the models, presented in Fig.4 and Fig.5.

Based on the contact pressure data for the eleven cases, summarized in Table II, it is obvious that, the introduced cases from 2 to 11, where four slots of 3mm length and width of 0.3; 0.5; 0.7 and 1mm, ending with small circular holes of radius 0.3, 0.5, 0.7 and 1.0 mm and lying in pairs on mutually perpendicular axes, are characterized with increased max contact pressure, that is between 12.05 and 111.7% higher than the max pressure for the classic case. The zone of contact pressure > 38.86MPa occupies between 13.55 and 27.11% of the entire contact surface, while for the classic case it occupies 2.17%.

Similar results, summarized in Table III, are obtained for the contact penetration.

The max contact penetrations, for cases 2 to 11 with the above described geometry, are between 0.115 and 0.327µm. It is between 51.3 and 330.3% higher than that for the classic case. The zone with contact penetration > 0.0673µm, occupies respectively between 25.44 and 51.59% of the entire contact surface, which is between 12 and 24.7 times higher than the relevant zone for the classic case.

V. Conclusion

When bolt holes of bolted busbar connections have 4 slots, lying in pairs on mutually perpendicular axes and ending with small circular holes, then the values of contact pressure and penetration are noticeably high in the area of busbars contact.

The deeper contact penetration increases the number of α-spots and their dimensions as well as the true contact area and this in turn reduces the contact resistance.

References

[1] R.S. Timsit, The Technology of High-Power Connections: A Review, *20th International Conference on Electrical Contacts*, Zurich, Switzerland, 2002, p. 526.

[2] R. Holm, *Electric Contacts. Theory and application*. Berlin, Germany, Springer-Verlag, 1976.

[3] A. Beredihin, M. Homiakov. *Electrical contact connections*. Energia, Moscow. 1980. (in Russian).

[4] R. Tzeneva, P. Dineff and Y. Slavtchev, Bolted Busbar Connections, *XIV-th International Symposium on Electrical Apparatus and Technologies SIELA 2005, 2 – 4 June 2005, Proceedings of Papers* Volume I, pp. 207 – 211, Plovdiv, Bulgaria.

Electric Field Distribution in Bolted Busbar Connections with Slotted Holes

Raina T. Tzeneva[1], Yanko T. Slavtchev[2] and Mikel Octavio[3]

Abstract – **The paper discusses how the bolt holes, of bolted busbar connections, influence the change of electric field distribution, when the holes are slotted in an effort to increase significantly the true contact area and reduce the contact resistance. The new hole profile, featuring 4 slots of length 3mm on mutually perpendicular axes, rotated at an angle of 45 degrees in relation to the busbars axes, is compared with the classical one of bolted busbar connections by the help of several computer models. It has been estimated that the new case changes the electric field distribution in the connection. Additionally, the Joule heat distribution in the buses and in the rest of the connection parts is investigated.**

Keywords – **bolted busbar connections, slotted hole shape, contact resistance, electric field distribution, Joule heat distribution**

I. INTRODUCTION

During the past years, reliability operation requirements of high power connections have grown significantly due to the rapidly increased consumption of electric energy worldwide. The design fundamentals for reliable high-power connections used in bare overhead lines are given in [1] and they include: maximization of true area of electrical contact, optimization of frictional forces, minimization of creep and stress relaxation, minimization of fretting and galvanic corrosion, minimization of differential thermal expansion along interfaces and normal to interfaces. A new bolt-hole shape for bolted busbar connections is proposed [3]. It features 4 slots on mutually perpendicular axes, rotated at an angle of 45 degrees in relation to the busbar axes. There were built 11 models to study the contact pressures and penetrations within the busbars contact zone. It is observed that there is a considerable raise in the values of the two studied parameters, in comparison to the classic bolted busbars case, which leads in turn to an enlarged true contact area and reduced contact resistance. In short, the reliability of the connection will be improved.

THE OBJECTIVE of the present work is to study the effect that the slotted bolt holes in bolted busbar connections exert on the electric field distribution and Joule heat distribution within these connections.

[1]Raina T. Tzeneva is with the Department of Electric Apparatus, Faculty of Electric Engineering, 8 Kliment Ohridski, 1000 Sofia, Bulgaria, E-mail: tzeneva@tu-sofia.bg

[2]Yanko T. Slavtchev is with the Department of Logistics and Materials Handling, Faculty of Mechanical Engineering, Technical University of Sofia, 8 Kliment Ohridski, 1000 Sofia, Bulgaria, E-mail: blamail@abv.bg

[3] Mikel Octavio is with the University of Navarra, 31080 Pamplona, Spain, E-mail: octavio@retena.com

II. THEORETICAL BACKGROUNDS

Current conduction analysis is used to analyze a variety of conductive systems. Generally, the quantities of interest in a current conduction analysis are voltages, current densities, electric power losses (Joule heat).

The problems of current distribution are described by the Poisson's equation for scalar electric potential U;

$$\frac{\partial}{\partial x}\left[\frac{1}{\rho_x}\frac{\partial U}{\partial x}\right]+\frac{\partial}{\partial y}\left[\frac{1}{\rho_y}\frac{\partial U}{\partial y}\right]+\frac{\partial}{\partial z}\left[\frac{1}{\rho_z}\frac{\partial U}{\partial z}\right]=0 \quad (1)$$

where the components of the electric resistivity tensor ρ_x, ρ_y and ρ_z are constant.

The electric current density j can be obtained from the equation

$$j=-\rho^{-1}E=\rho^{-1}gradU \quad (2)$$

where ρ^{-1} is the inverse tensor of the electric resistivity.

Power losses (Joule heat produced) in a volume are

$$Q=EjdV \quad (3)$$

III. MODELING THE BOLTED BUSBAR CONNECTION

If a deeper contact penetration increases α-spots both in numbers and dimensions, enlarges the true contact area and decreases the contact resistance, then a new hole-shape might be introduced. A typical bolted busbar connection is shown in Fig. 1.

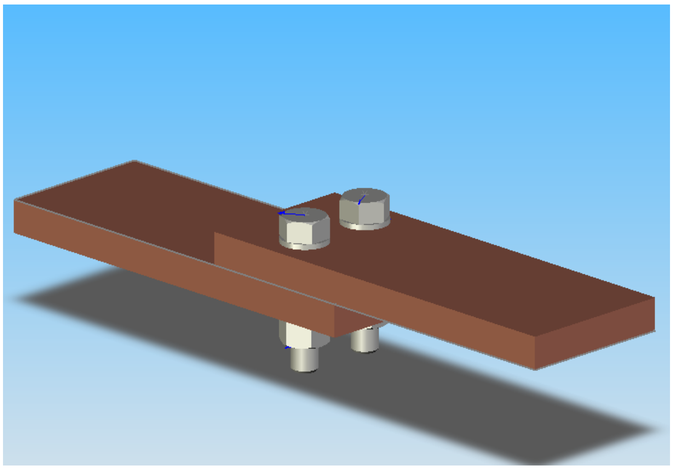


Fig. 1. A bolted busbar connection

The slotted bolt hole shape arises from [2] where longitudinal slot of width 3-4mm and length 50mm between the holes of the buses is proposed in order to increase the true connection area. This case is modeled and compared with the classic case in [3].

For that purpose there have been investigated 11 different models.

case 1 – the classical case – copper busbars with 2 bolt holes;

case 2 – the slots are parallel to the busbar axis;

case 3 – the slots are perpendicular to the busbar axis;

case 4 – mixed case – one of the busbars in the connection is of case 2 and the other one is of case 3;

For cases 2 to 4 all bolt holes have two slots of length 3mm and width 1mm.

For cases 5 to 8 the busbar holes have 4 slots of length 3 mm and varying width, arranged in such a way that the pairs of slots are on mutually perpendicular axes, rotated at an angle of 45 degrees in relation to the busbar axes.

case 5 – width 0.3mm;

case 6 – width 0.5mm;

case 7 – width 0.7mm;

case 8 – width 1mm;

case 9 – the 4 slots are not rotated;

case 10 – mixed case – the first busbar corresponds to case 8 and the second to case 9;

case 11 – a busbar hole with 8 slots of length 3mm and width 1mm;

Fig. 2 shows the hole-shape of the two-slot cases.

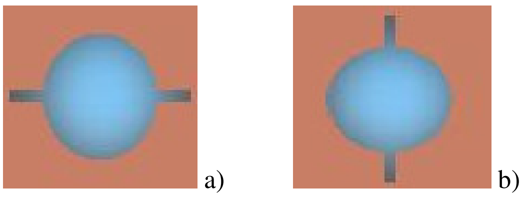


Fig. 2. Hole-shape with 2 slots

Fig. 3 presents the new hole-shape with 4 and 8 slots.

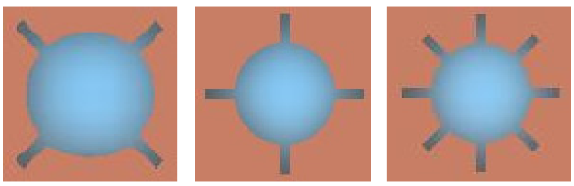


Fig. 3. Hole-shape with 4 and 8 slots

The investigated assembly consists of:

- Copper busbars (Young's modulus E = $1.1.10^{11}$Pa, Poisson's ratio μ = 0.34, width 60mm, height 10mm, length 160mm, busbars' overlap 60mm) with 2 holes of ∅10.5mm;
- Fasteners: bolts – Hex Bolt GradeB_ISO 4015 – M10 x 40 x 40 – N, steel E = 2.10^{11}Pa, μ = 0.3; nuts – Hex Nut Style1 GradeAB_ISO 4032 – M10 – W – N, steel E = 2.10^{11}Pa, μ = 0.3; washers – Plain Washer Small Grade A_ISO 7092 – 10, steel E = 2.10^{11}Pa, μ = 0.3. Tension in each bolt F = 15000N.

Several computer models smooth the research progress of the current density and Joule heat distribution changes that take place within the components of the bolted busbar connection, due to the introduced slotted bolt holes. The FEA package ANSYS 10 is employed in the analysis of the electric field and the Joule heat distributions. The model is meshed with the SOLID 98 element – Tetrahedral Coupled-Field Solid. It is defined by ten nodes with up to six degrees of freedom at each node.

The current density distribution in the bolted busbar assembly with 4 slots of length 3 mm and width 0.5 mm (case 3) is shown in Fig. 4.

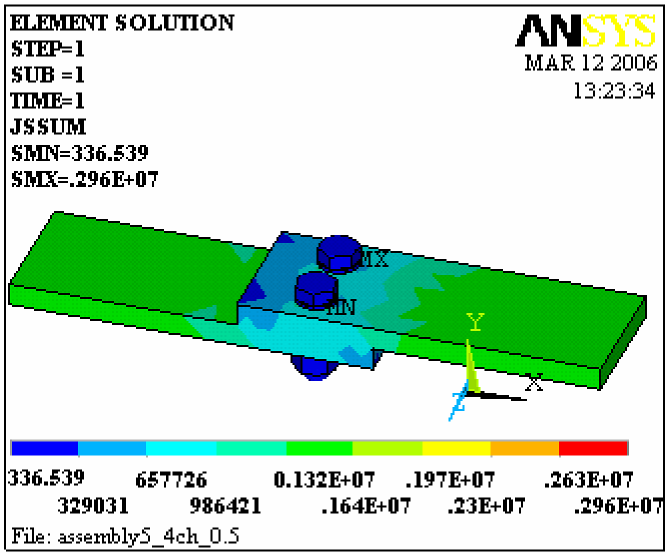


Fig. 4. Current density distribution in the bolted assembly for case 3

Fig. 5 represents the Joule heat distribution in the assembly for the same case.

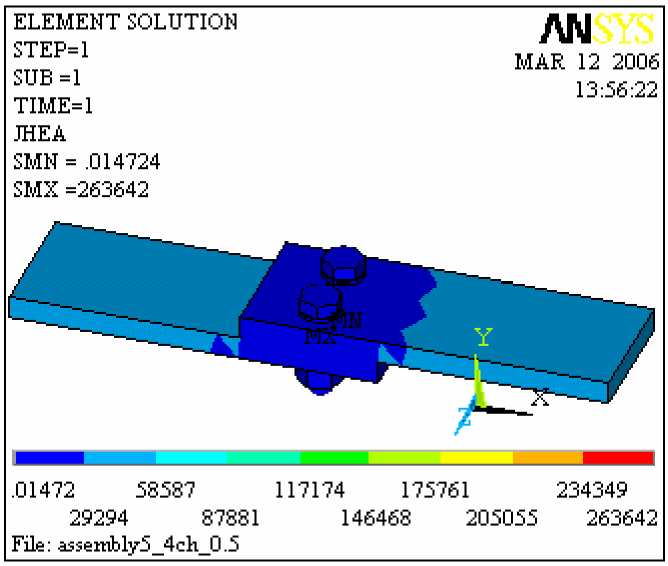


Fig. 5 Joule heat distribution in the bolted busbar assembly for case 3

The maximum current density value for each of the investtigated cases is shown in Fig. 6.

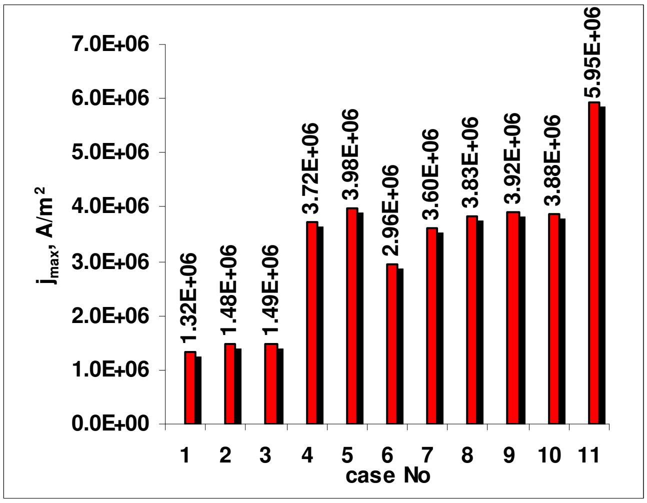


Fig. 6. Maximum value of the bolted busbar assembly current density

Additionally, the impact of the hole slots on the current density within the fasteners (bolts, nuts and washers) and the busbars is considered. Fig. 7 reviews the model current density distribution in the bolt, nut and washer for case 5 while Fig. 8, Fig. 9 and Fig. 10 summarize graphically the corresponding fastener results for each of the eleven studied cases.

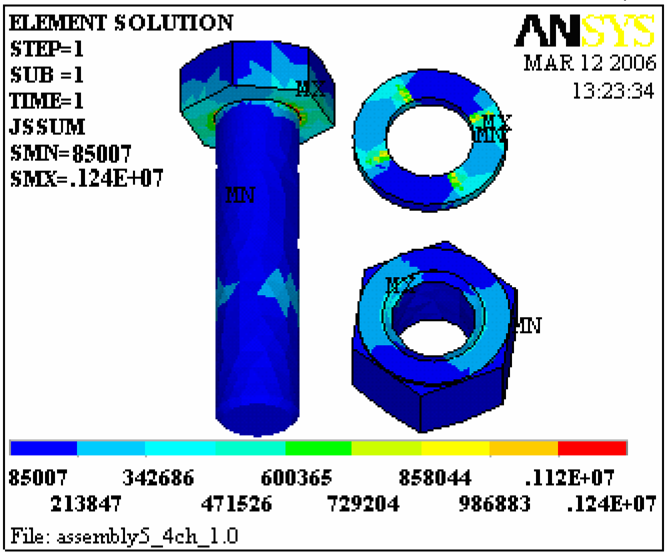


Fig. 7. Bolt, nut and washer current density distribution for case 5

IV. Discussion and Conclusions

It is observed that the introduction of slotted bolt holes raises significantly the contact pressure and penetration within the busbars contact area [3]. This will help to decrease the contact resistance and will provide an opportunity for more reliable assembly operation.

The analyses of the fasteners current density distribution for the entire range of the studied cases confirm that:

- the max assembly current density remains unaffected regardless of the two introduced slots orientation – either parallel or perpendicular to the busbars axes (Fig. 6); the same holds for the bolts (Fig. 7), nuts (Fig. 8) and washers (Fig. 9).

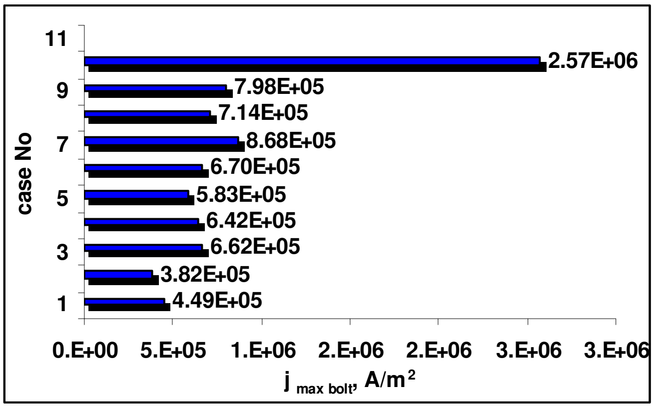


Fig. 8. Maximum value of the bolt current density

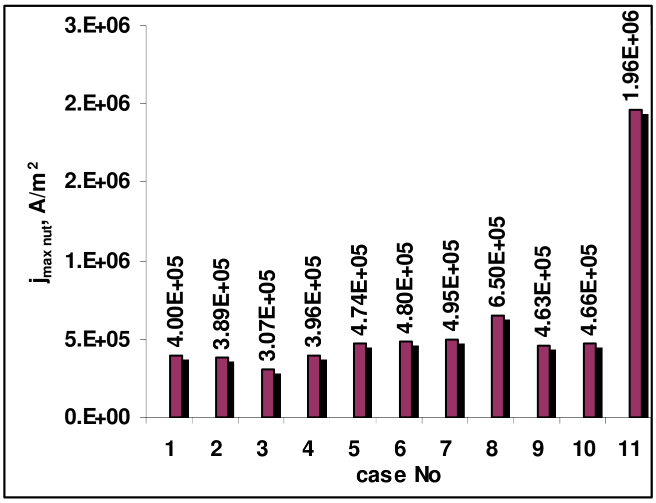


Fig. 9. Maximum value of the nut current density

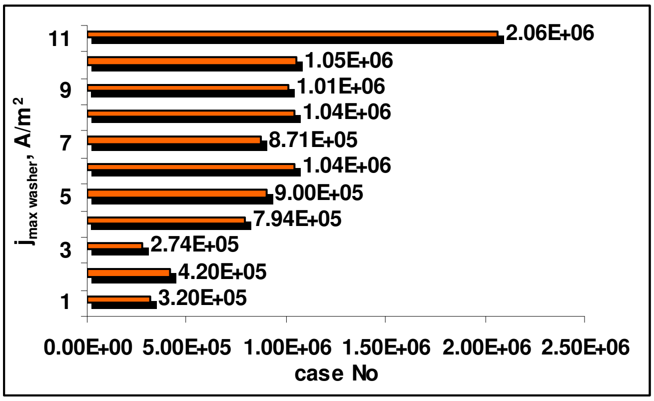


Fig. 10. Maximum value of the washer current density

- the max assembly current density value is between 2.5 and 3 times higher for the 4-slot cases than for the classic case, regardless of the slots width and position; in most cases the max value is concentrated within very small zones at the slot ends and edges; the max value within the bolts is situated near the edge of the bolts head-to-neck transition and is about 1,5 to 1,8 times higher compared to the classical case; the nuts max current density value in most cases is approximately the same or a little (1,6 times) higher compared to the classical case; the washers max current density value is 2,5 to 3 times higher than for the classical case and is located in the zones contacting the slots near the bolt holes;

- the max current density values increase significantly within all of the assembly parts when introducing 8 slots of width 1mm and length of 3mm; it is 4.5, 5.7, 4.9 and 6.4 times higher respectively for the assembly, the bolts, the nuts and the washers.

When the bolt-holes of bolted busbar connections are slotted, the max current density value in the assembly and the fasteners increases significantly. Considering the busbars, it is concentrated in very small zones, around the slot ends and edges. As for the bolts, it is around the edge under the head of the bolts while for the washers – it is in the slots-washer contact zone.

The 8-slot case adds rapidly to the max current density value in both the assembly and the fasteners. That is why this case is not recommended for use despite the fine contact pressure and penetration values.

ACKNOWLEDGEMENT

This research was encouraged and financially supported by “Quality 21-st Century” foundation.

REFERENCES

[1] R.S. Timsit, The Technology of High-Power Connections: A Review, *20th International Conference on Electrical Contacts*, Zurich, Switzerland, 2002, p. 526.

[2] A. Beredihin, M. Homiakov. *Electrical contact connections*. Energia, Moscow. 1980. (in Russian).

[3] R. Tzeneva, P. Dineff and Y. Slavtchev, Bolted Busbar Connections, XIV-th International Symposium on Electrical Apparatus and Technologies SIELA 2005, 2 – 4 June 2005, Proceedings of Papers Volume I, pp.207 – 211, Plovdiv, Bulgaria

.

Broken Bar Detection In Induction Machines Using ADC 42

Krstevski Toni

Abstract – **This study describes broken bar detection in induction motors before actual breakdown occurs in purpose maintenance to be predictable. It is based on sampling stator current on motor with high resolution A/D converter (ADC 42), and three phase current waveforms are analyzed using Fast Fourier Transformers (FFT).**

Key words – **ADC 42, FFT.**

I. Introduction

Squirrel-cage induction motor is low priced, robust and rugged, simple and easy to maintain, it has become the most commonly used electrical rotating machine in industry. So, the need for detection of rotor faults as in rotor bar or in end-ring become increasingly important. Defect in rotor will result in a high resistance which will overheat that area, so cracking or small hale may than occur in rotor bar [4]. It is known that rotor bar and end-ring faults yield asymmetrical operation of induction machines causing unbalanced current, torque pulsation, increased losses and decreased average torque. Vibration analyses, thermal analyses and current spectrum analyses have been applied to monitor rotor bar faults, focusing on current spectrum analyses as the current signal is easily accessible for induction motor. In industry however, low cost and high sensitivity of motor diagnostic system is highly demanded.

II. Model of Asymmetric Rotor of Induction Cage Motor

For the purpose of analyses, each rotor bar and segment of the end ring is replaced by an R-L series equivalent circuit representing the resistive and inductive nature of cage. Such an equivalent circuit is shown at Fig. 1. Assuming the rotor of squirrel- cage induction machine to be symmetric, an equivalent model of a wound - rotor machine may be obtained in synchronously rotating dq reference frame [2] and [3]:

Toni Krstevski is student at SS. Cyril and Methodius University, Faculty of Electrical Engineering, Karpos II bb, P.O.B. 574 1000 - Skopje, Macedonia, E-mail: toni77kr@yahoo.com

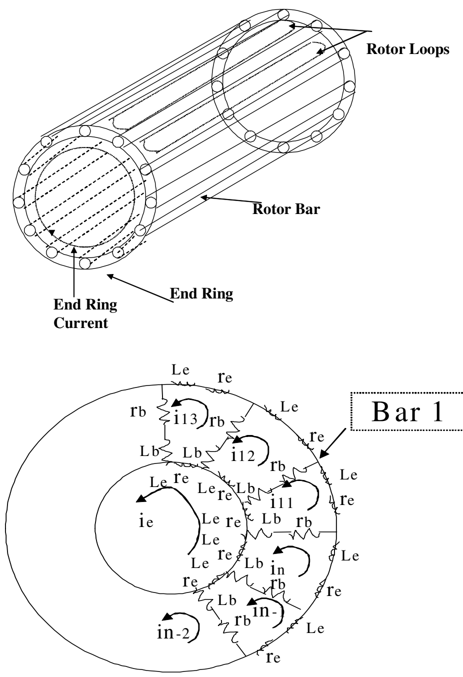


Fig. 1. Rotor cage equivalent circuit

$$v_{ds} = r_s i_{ds}^e + p\lambda_{ds}^e - \omega_e \lambda_{qs}^e \tag{1}$$

$$v_{qs} = r_s i_{qs}^e + p\lambda_{ds}^e + \omega_e \lambda_{ds}^e \tag{2}$$

$$v_{dr} = 0 = r_s i_{dr}^e + p\lambda_{dr}^e - \left(\omega_e - \omega_r\right)\lambda_{qr}^e \tag{3}$$

$$v_{qr} = 0 = r_s i_{qr}^e + p\lambda_{qr}^e - \left(\omega_e - \omega_r\right)\lambda_{dr}^e \tag{4}$$

$$\lambda_{ds}^e = L_{ls} i_{ds}^e + L_m\left(i_{ds}^e + i_{dr}^e\right) \tag{5}$$

$$\lambda_{qs}^e = L_{ls} i_{qs}^e + L_m\left(i_{qs}^e + i_{qr}^e\right) \tag{6}$$

$$\lambda_{dr}^e = L_{lr} i_{dr}^e + L_m\left(i_{ds}^e + i_{dr}^e\right) \tag{7}$$

$$\lambda_{qr}^e = L_{lr} i_{qr}^e + L_m\left(i_{qs}^e + i_{qr}^e\right) \tag{8}$$

Where:
P- Number of poles;
w_r -rotor speed;
T_e -torque;
r_s - equivalent stator resistance;
r_r - equivalent rotor resistance;
L_{ls} - equivalent stator leakage inductance;
L_{lr} - equivalent rotor leakage inductance;
L_m - equivalent magnetizing inductance;

In many studies such as the work of Krause [5], is convenient to express the voltage and flux linkage equations from Eq. (1) to Eq. (8) in terms of reactance. From this studies determination of torque, in synchronously rotating reference frame is made with neglect the time rate of change of all flux linkages and then employ the relationships:

$$\sqrt{2}\,\tilde{F}_{as} = F_{qs}^{e} - jF_{ds}^{e} \tag{9}$$

$$\sqrt{2}\,\tilde{F}_{ar} = F_{qr}^{'e} - jF_{dr}^{e} \tag{10}$$

Torque can be expressed in terms of currents in phasor form by first writing the torque in terms of currents in the synchronously rotating reference frame and utilizing Eq. (9) and (10) to relate synchronously rotating reference frame and phasor quantities. The torque expression becomes:

$$T_e = 3\left(\frac{P}{2}\right)\left(\frac{X_m}{\omega_b}\right)R_e\left[j\tilde{I}_{as}^{*}\tilde{I}_{ar}^{'}\right] \tag{11}$$

where:

ω_b is base electrical angular velocity used to calculate the inductive reactances;

$\tilde{I}_{as}^{*}$ is the conjugate of $\tilde{I}_{as}$;

The vast majority of singly excited machines are the squirrel cage rotor type, where :

$$\tilde{I}_{ar}^{'} = -\frac{j(\omega_e/\omega_b)X_m}{r_r^{'}/s + j(\omega_e/\omega_b)X_{rr}^{'}}I_{as}^{'} \tag{12}$$

where:

$$X_{rr}^{'} = X_{lr}^{'} + X_m \tag{13}$$

Substituting Eq. (12) into Eq. (11), yield:

$$T_e = \frac{3\left(\frac{P}{2}\right)\left(\frac{\omega_e}{\omega_b}\right)\left(\frac{X_m}{\omega_b}\right)\left(\frac{r_r^{'}}{s}\right)\left|\tilde{I}_{as}\right|^2}{\left(\frac{r_r^{'}}{s}\right)^2 + \left(\frac{\omega_e}{\omega_b}\right)^2 X_{rr}^{'2}} \tag{14}$$

where the slip s is defined:

$$s = \frac{\omega_e - \omega_r}{\omega_e} \tag{15}$$

In general all n-loop rotor current ($i_{l1}, i_{l2},..., i_{ln}$) are mapped into a n-dimensional vector space. This new space vector is defined by the transformation matrix T such that:

$$\begin{bmatrix} i_{r1}^{'} \\ i_{r2}^{'} \\ i_{rn}^{'} \end{bmatrix} = T\begin{bmatrix} i_{l1}^{'} \\ i_{l2}^{'} \\ i_{ln}^{'} \end{bmatrix} \tag{16}$$

where:

$i_{r1}^{'}$ - real part of the rotor current space vector;

$i_{r2}^{'}$ - imaginary part of the rotor current space vector;

$i_{r3...rn}^{'}$ - zero sequence components of the rotor current space vector;

$$\begin{bmatrix} f_d \\ f_q \\ 0 \\ . \\ . \\ . \\ 0 \end{bmatrix} = \frac{n-1}{n}\begin{bmatrix} \cos(\theta) & . & . & . & \cos\left(\theta + \left\{\frac{n-1}{n}\right\}2\pi\right) \\ \sin(\theta) & . & . & . & \sin\left(\theta + \left\{\frac{n-1}{n}\right\}2\pi\right) \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ Z_{n1} & . & . & . & Z_{nn} \end{bmatrix}\begin{bmatrix} f_1 \\ f_2 \\ . \\ . \\ . \\ f_n \end{bmatrix} \tag{17}$$

The T transformation is generated from a very simple algorithm. The first two lines from Eq. (17) correspond to n-phase dq transformation, plus a constant. Since the vectors formed by the two lines are linearly independent, the null space of this 2xn sub-matrix has a dimension (n-2). A base for the null space is defined by (n-2) linearly independent vectors $\underline{Z}$ such that $T_{dq}\cdot\underline{Z}=0$. Taking θ equal to zero, for simplicity:

$$\begin{bmatrix} 1 & \cos\left(\frac{2\pi}{n}\right) & . & . & . & \cos\left(\frac{n-1}{n}2\pi\right) \\ 0 & \sin\left(\frac{2\pi}{n}\right) & . & . & . & \sin\left(\frac{n-1}{n}2\pi\right) \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \\ . \\ . \\ . \\ z_n \end{bmatrix} = 0 \tag{18}$$

From [1], [2] and [3]:

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 1 & \cos\left(\frac{2\pi}{n}\right) \\ 0 & \sin\left(\frac{2\pi}{n}\right) \end{bmatrix}^{-1} \times \begin{bmatrix} \cos\left(\frac{2}{n}\cdot 2n\right) & . & . & . & \cos\left(\frac{n-1}{n}\cdot 2\pi\right) \\ \sin\left(\frac{2}{n}\cdot 2\pi\right) & . & . & . & \sin\left(\frac{n-1}{n}\cdot 2\pi\right) \end{bmatrix}\cdot\begin{bmatrix} z_3 \\ . \\ . \\ z_n \end{bmatrix} \tag{19}$$

$z_3, z_4,..., z_n$, can be arbitrary chosen. Making $z_3=1$ and $z_4= z_5=... z_n=0$ z_{31} and z_{32} are determined, resulting in third vector of the null space:

$$\begin{bmatrix} z_{31} & z_{32} & 1 & 0 & . & . & . & 0 \end{bmatrix} \tag{20}$$

Taking $z_4=1$ and $z_3= z_5=... z_n=0$ z_{41} and z_{42} are determined, and the fourth vector of hull space is calculated as:

$$\begin{bmatrix} z_{41} & z_{42} & 0 & 1 & . & . & . & 0 \end{bmatrix} \tag{21}$$

The rotor current complex vector ($\underline{i}_r$) is computed from the symmetric model. With i_r referred to a rotor fixed reference frame, the n- loop currents (i_{li}) are then computed as:

$$\begin{bmatrix} i_{l1} \\ i_{l2} \\ i_{l3} \\ i_{ln} \end{bmatrix} = T^{-1}\begin{bmatrix} i_{dr} \\ i_{qr} \\ 0 \\ 0 \end{bmatrix} \tag{22}$$

Let assume that kith bar is broken. The null current through this bar is obtained by modifying the n-loop rotor current vector according to:

$$\begin{bmatrix} i_{l1}^{new} \\ i_{l2}^{new} \\ \cdot \\ \cdot \\ \cdot \\ i_{lk}^{new} \\ i_{l(k+1)}^{new} \\ \cdot \\ \cdot \\ \cdot \\ i_{ln}^{new} \end{bmatrix} = T \begin{bmatrix} i_{l1} \\ i_{l2} \\ \cdot \\ \cdot \\ \cdot \\ \dfrac{i_{lk} + i_{l(k+1)}}{2} \\ \dfrac{i_{lk} + i_{l(k+1)}}{2} \\ \cdot \\ \cdot \\ \cdot \\ i_{ln} \end{bmatrix} \tag{23}$$

III. Instrument Description

According on mathematical model that is represented above for the expectation is to have increasing of the amplitude of sideband components, while other harmonics maintained their amplitude when some bar of the squirrel - cage induction motor is broken. For the purpose to show that, stator current is measured on the terminals of the motor with broken bar and FFT analyses are made. Analog to digital converter is used (ADC 42) that is shown on Fig. 2, with characteristic which are given in Appendix I.

Fig. 2. Analog to digital converter (ADC 42)

For measuring stator current of the motor is used digital clump meter, shown on Fig. 3 which pick up the current flowing through the conductor. Technical specifications of clump meter are given in Appendix II.

Fig. 3. Digital clump meter

Because measured current on transformer jaws is with small velocity it was necessary to amplify the signal so the ADC 42 could measured. An electronic device as amplifier was made and its electronic circuit is given at Appendix III.

IV. Experimental Results

It was used 3 phase induction motor from SIEMENS type 1LA 3106-4AA21-Z with characteristics that are given in Appendix IV. Also, Pico Scope Software application [6] is used. Fig. 4. shows the current frequency spectrum with the motor in healthy conditions. Some harmonics are visible in the measured current spectrum, with emphasis on the third, fifth and seventh harmonics, caused by inherent asymmetries of three phase windings. Besides that, the sideband component at frequency (1-2s)f_e can also be seen, which is caused by inherent asymmetries of the cage. Fig. 5. shows the harmonic current spectrum for broken bar. It is notorious the increase in the amplitude of the sideband components, while other harmonics components maintained their amplitudes. Fig. 6. shows the broken bar of the squirrel - cage of induction motor.

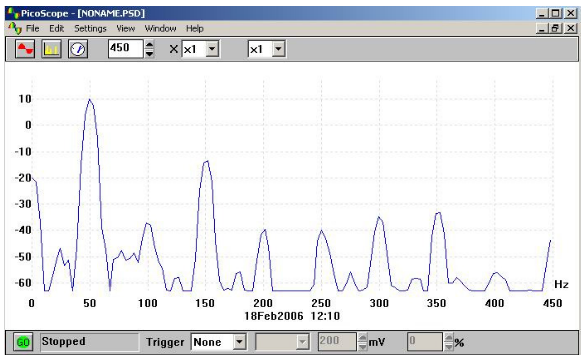


Fig. 4. Current frequency spectrum with motor in healthy conditions

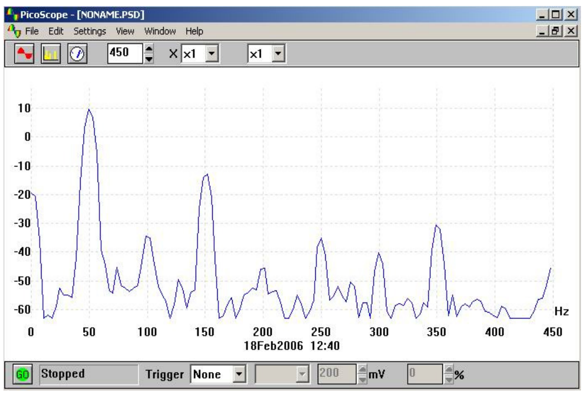


Fig. 5. Current frequency spectrum with broken bar

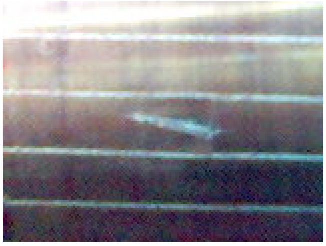


Fig. 6. Broken bar of squirrel cage of induction motor

V. Conclusion

In this paper, is given mathematical model of classical fourth-order transient model of symmetrical induction machines, with additional computation limited to the transformation of the rotor current vector to a rotor fixed reference frame. Also are made measurement on stator current on motor with healthy conditions and with a broken bar. Experimental results show increasing of amplitude of sidebands components and theoretical mathematical model confirm that.

Appendix I

ADC 42 Characteristic

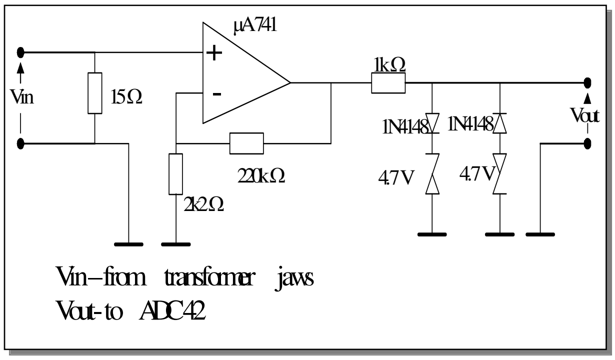
Resolution 12 bits; Channels 1 x BNC; Voltage ranges ±5 V; Overload protection ±30 V; Input impedance 1 MΩ; Sampling rate 15 kS/s; Accuracy ±1%; Scope time bases 500 µs/div to 50 s/div; Spectrum ranges 0 to 7kHz, 65 dB dynamic range; Analog bandwidth 7.5 kHz; Buffer Size- No buffer**;** Power supply-Not required; Dimensions 55x55x15mm (2.17x2.17x0.6 in); Output connector D25 to PC parallel port; Supplied software Pico Scope; DOS and Windows (3.x, 95/98/XP, NT/2000);

Appendix II

Technical specification of Clump meter

- Max. Voltage between terminals and earth ground: 600 Vrms,
- Operating principle: dual slop integration,
- Sample rate: 2 times/ sec for digital data;

Appendix III



Appendix IV

3 phase squirrel-cage induction motor from Siemens, type: 1LA 3106-4AA21-Z; IEC 100L;

Nr. E6482 1169 04 012; B5 IP 54; Rot KL16; 220/380 V; Δ/Y; 9/5.2A; 2.2 kW; ICL B cosφ 0.82 ; 1415/min; 50Hz; VDE 0530

References

[1] Carla C. Martines Cunha, Renato O. C. Lyra, Braz J. Cardaso Filho, "Simulation and Analysis of Induction Machine with rotor Asymmetries" *IEEE Transaction on industry applications,* Vol. 41, pp 18-24, No. 1, January/ February 2005

[2] Alfredo R. Muňoz, Thomas A. Lipo, "Complex Vector Model of the Squirrel-Cage Induction machine Including Instantaneous Rotor Bar Current" *IEEE Transaction on industry applications,* Vol. 35, pp 1332-1339, No. 6, November/ December (1999)

[3] Elkasagby, N. M., Eastham, A. R. and Dawson, G. E. "Detection of broken bar in the Cage rotor on an Induction Machine" *IEEE Transaction on Industry application,* Vol. 28, No. 1, pp 165-171 (1992)

[4] Hakan Çaliş, Abdülkadir Çakir, Müslüm Arkan, "Sensorless broken bar detection in induction machines" eleco.emo.org.tr/eleco2003/ELECO2003/asession/A2-05.pdf

[5] P.C. Krause, O. Wasynezuk, S.D. Sudhoff, "Analysis of the electrical machinery" *The institute of electrical and electronics engineers,* pp New York, 1995

[6] http://www. pico tehnology.com

Parallel operation of transformers in practice

Natasa D. Mojsoska[1]

Abstract— **In this paper is analyzed the parallel operation of two or more transformers when there is not fulfilled two conditions $u_{kI} \neq u_{kII}$ and $k_{12I} \neq k_{12II}$. The phase diagram is given. It is explained the differences between theory and practice. Also are referred the other conditions that influence on decision the type of transformer which should work on parallel operation. Conclusion is given for what are the most important values for analyses of the parallel operation. I must mention that this work is only small part of a whole project for researching the parallel operation of two transformers in practice, the conditions, distribution of rated powers, max allowed power, equalizing currents est.**

Keywords—**Transformer, Parallel operation, Impedance voltage, Transformation ratio.**

I. Introduction

When technical and economical justifications exist, total power which had to be transferred to the consumers of EES is split to more transformers units. In practice at substations in Macedonia there are two and more transformers in parallel operation. Causes for this situation could be following:

- increasing of the loading of substation caused by turning on new consumers or resizing the capacity of the old one
- construction of the substation was in more phases in a long period
- variable loading depends on certain timeline period (peak loading, overloading, winter loading etc)
- one transformer is back up in case of breakdown or blow of the first transformer

All this provides continuality, reliability and security in supplying with electrical energy to the consumers and also maintenance costs are lower.

II. Conditions for parallel operation of transformers

In theory is provided that optimal parallel operation of transformers can be done if are fulfilled following conditions:

1. rated voltages of primary windings of all trans-formers in parallel operation to be equal to network voltage Eq.(1)

$$U_{1I} = U_{1II} = U_{1III} = \text{..........} = U_{1n} = U_m \tag{1}$$

2. rated voltages of primary windings of all trans-formers in parallel operation to be equal to network voltage Eq.(2)

$$U_{1I} = U_{1II} = U_{1III} = \text{..........} = U_{1n} = U_m \tag{2}$$

3. rated voltages of secondary windings of all transfor-mers to be equal Eq.(3)

$$U_{2I} = U_{2II} = U_{2III} = \text{..........} = U_{2n} \tag{3}$$

This condition is come down to equality of trans-formation ratios of transformers Eq. (4)

$$k_{12I} = k_{12II} = k_{12III} = \text{..........} = k_{12n} \tag{4}$$

where:

$$k_{12I} = \frac{U_{1I}}{U_{2I}}, \quad k_{12II} = \frac{U_{1II}}{U_{2II}}, \ ...., \quad k_{12n} = \frac{U_{1n}}{U_{2n}}$$

4. the phase displacement between the primary and secondary winding should be the same i.e. the transformer connections should have the same clock hour figure
5. equality of impedance voltages Eq.(5)

$$u_{kI} = u_{kII} = u_{kIII} = \text{..........} = u_k \tag{5}$$

or the components of active and reactive impedance voltage to be equal Eqs. (6) and (7)

$$u_{kaI} = u_{kaII} = u_{kaIII} = \text{..........} = u_{kan} \tag{6}$$

$$u_{k\sigma I} = u_{k\sigma II} = u_{k\sigma III} = \text{..........} = u_{k\sigma n} \tag{7}$$

From all recent analyses, both theoretical and practical, the conclusion is that parallel operation of transformers is possib-le if completely is fulfilled the condition that transformers should have the same or appropriate clock hour figure, while with the other conditions (equality of secondary voltages, impedance voltages and transformation ratios) is allowed certain tolerance which is limited by defined frames.

III. Allowed frames of tolerance

A. Transformation ratio

If the condition for equality of transformation ratios is not fulfilled allowed frames are:

- for transformers with transformation ratios $\mathbf{k_{12} \le 3}$ and for transformers for self necessities is allowed $\mathbf{\Delta k \le 1\%}$, Eq. (8):

$$\Delta k = \frac{k_{12II} - k_{12I}}{k} .100 \tag{8}$$

where:

$$k = \sqrt{k_{12II} . k_{12I}} \tag{9}$$

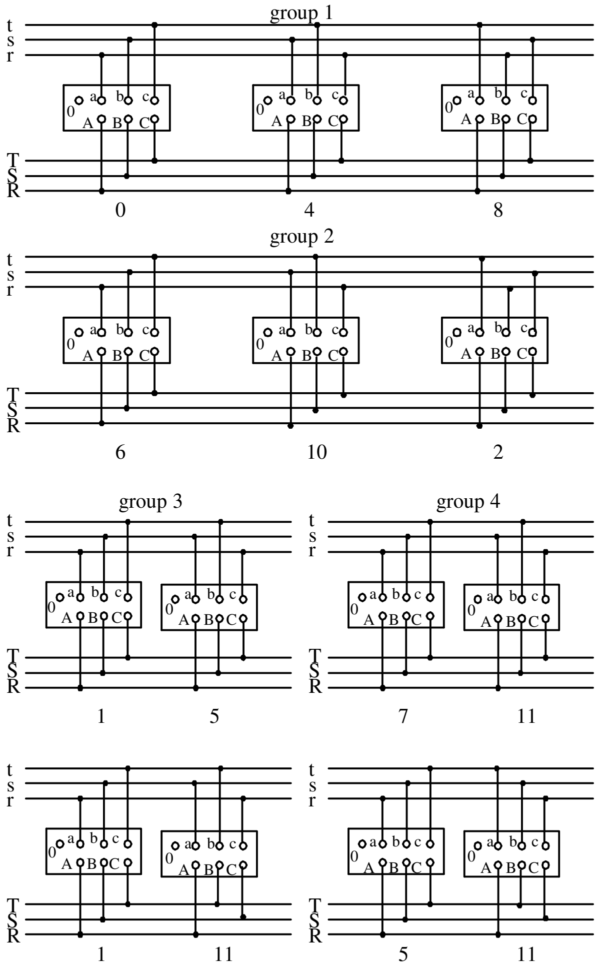
- for others transformers $\mathbf{\Delta k \le 0.5\%}$.

In cases when transformers which should work in parallel operation are from even or odd class of connection group, but do not belong to the same group, parallel operation is possible to be done. Even groups of connection are 0 and 6 and odd are 5 and 11. Even figures of combination are Yy, Dd and Dz and odd are Yd, Dy and Yz.

Connections of transformers are split at 4 groups:

[1] Natasa D. Mojsoska is with the Faculty of Technical Sciences, I.L.Ribar bb. 7000 Bitola, Macedonia E-mail: natmojso@freemail.com.mk

group 1	clock hour	0	4	8
group 2	clock hour	2	6	10
group 3	clock hour	1	5	
group 4	clock hour	7	11	



All mention connections are given on the following Fig. 1.

Fig. 1. Survey of connection groups of transformers in parallel operation

B. Impedance voltage

Tolerance of impedance voltage of each transformer in parallel operation should be within the limits of **±10%** from middle arithmetical value of others transformers Eqs. (10) and (11).

$$u_{ksr} = \frac{u_{kI} + u_{kII}}{2} \tag{10}$$

$$\Delta u_{kI} = \frac{u_{kI} - u_{ksr}}{u_{ksr}} \cdot 100 \le (\pm 10)\% \tag{11}$$

$$\Delta u_{kII} = \frac{u_{kII} - u_{ksr}}{u_{ksr}} \cdot 100 \le (\pm 10)\% \tag{12}$$

It is recommended transformer with lower power to have bigger impedance voltage and conversely. Also ratio of rated powers of the biggest and the smallest transformer should not exceed **3:1**

IV. Parallel Operation of Transformers with Different Transformation Ratios and Impedance Voltages

Some authors have made analyses of parallel operation of transformers with presumption that only one of conditions for optimal work is not fulfilled, while the others are. In practice the most common case is when two or all transformers working in parallel operation do not satisfy two conditions at the same time ($k_{12I} \neq k_{12II}$ and $u_{kI} \neq u_{kII}$).

In this paper work the subject of analyses is parallel work of two transformers with different transformation ratios and impedance voltages. Phase diagram of these transformers is given on Fig. 2.

For this concrete case presumptions are:

$$U_{1I} = U_{1II} = U_m$$

$$k_{12I} < k_{12II} \quad \Rightarrow \quad U_{2nI} > U_{2nII}$$

$$u_{kI} < u_{kII}$$

$$S_{nI} > S_{nII} \quad \Rightarrow \quad I_{opI} > I_{opII}$$

First transformer is mark with I and second with II.

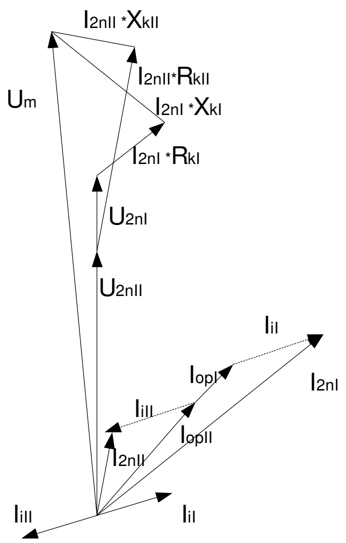


Fig. 2. Phase diagram of two transformers in parallel operation with different k_{12} and u_k

U_{2nI} is secondary voltage of the first transformer working in parallel operation

U_{2nII} is secondary voltage of the second transformer

I_{iI} is equalizing current of the first transformer

I_{iII} is equalizing current of the second transformer

I_{2nI} is rated current of the first transformer

I_{2nII} is rated current of the second transformer

I_{opI} is current of loading of the first transformer

I_{opII} is current of loading of the second transformer

U_m is voltage of network.

Because transformation ratios of transformers are not equal consequently difference in the secondary voltages Δu (%) exists. That difference will cause equalizing current to flow between the transformers. Phase diagram show that equalizing current increase current of loading of the first transformer I_{opI} and decrease current of loading of the second transformer I_{opII}. This caused difference at values of the rated currents I_{2nI} and $I_{2nII.}$ Voltage of network U_m is primary voltage for both transformers. But presumption that $k_{12I} < k_{12II}$ means that there is difference in secondary voltages and different voltage drop at firs and second transformer. They are $I_{2nI}{\cdot}R_{kI}$ ohmic voltage drop, $I_{2nI}{\cdot}X_{kI}$ reactive voltage drop for first transformer and $I_{2nII}{\cdot}R_{kII}$ ohmic voltage drop, $I_{2nII}{\cdot}X_{kII}$ reactive voltage drop for second transformer.

The values of the current are given in the following Eqs. (13) and (14) in % from the rated currents of individual transformer.

$$i_{iI}(\%) = \frac{I_{iI}}{I_{2nI}} \cdot 100 = \frac{\Delta k}{u_{kI} + u_{kII} \cdot \dfrac{P_{nI}}{P_{nII}}} \cdot 100 \qquad (13)$$

$$i_{iII}(\%) = \frac{I_{iII}}{I_{2nII}} \cdot 100 = \frac{\Delta k}{u_{kII} + u_{kI} \cdot \dfrac{P_{nII}}{P_{nI}}} \cdot 100 \qquad (14)$$

In worst case allowed values are:

$$P_{nI} : P_{nII} = 3 : 1$$

$$\Delta k = 1\% \qquad \Rightarrow \qquad u_{kI} = 0.9 \cdot u_{kII}$$

With substitution of this values in Eqs. (13) and (14) the results are given in Eqs. (15) and (16):

$$i_{iI} = \frac{0.26}{u_{kII}}(\%) = \frac{0.23}{u_{kI}}(\%) \qquad (15)$$

$$i_{iII} = \frac{0.77}{u_{kII}}(\%) = \frac{0.75}{u_{kI}}(\%) \qquad (16)$$

Because impedance voltages are moving in frames from 3 to 10 equalizing currents will be about from 0.23% to 2.6% from rated currents.

V. Parallel operation of transformers in practice

Mainly in practice we have slightly different picture about parallel operation of transformers.

With turning on of one transformer within parallel operation with already existing transformer necessary power of the group has played important role. For ex. if we have transformer with power of 100kVA and necessary power which should provide the group is 128kVA in that case we can not installed transformer with that power of 28kVA because that special design will be too expensive and economically unprofitable. So it should be installed transformer with standard power or in this case the smallest is with 50kVa. Totally installed power will be 150kVA. Because of unfulfilled conditions for parallel work with calculation can be find allowed max power those transformers which are connected can transmit in the aggregate with none of them being overloaded.

So as a first condition for parallel operation of these transformers is following:

If by limiting caused by any of basic conditions would be limited max power of the group under 100kVA in that case there is not necessity of installation the second transformer because that power we already have it.

In practice transformer with bigger rated power has bigger impedance voltage. It can be get with measuring of voltage when no-load and flown have rated current between the transformer windings.

From other side different manufacturers produced transformers with same rated power but different impedance voltages [4] from 3.7% to 5% which is a big difference.

It is also important how will be divided the power among individual transformers Eq. (17).

$$S_I = S_{nI} \cdot \frac{u_k}{u_{kI}} \qquad (17)$$

where

$$u_k = \frac{S}{\dfrac{S_{nI}}{u_{kI}} + \dfrac{S_{nII}}{u_{kII}} + \ldots} \qquad (18)$$

The main issue is max power Eq. (18) that transformers can transmit in the aggregate with none of them being overloaded:

$$S_{max} = S_{nI} \cdot \frac{u_{kn}}{u_{kI}} + S_{nII} \cdot \frac{u_{kn}}{u_{kII}} + \ldots \qquad (19)$$

where:

u_{kn} is the smallest of all impedance voltages of the individual transformers

S_{max} is max allowed power that transformers can transmit in the aggregate with none of them being overloaded

Calculation is made from S_{max} how divided this power to each of transformers and value of efficiency Eq. (20) of the whole group is found:

$$\eta = \frac{S_{max}}{S} \qquad (20)$$

η is coefficient of efficiency of the whole group.

A. Mathematical analysis

One of many calculations that are made for this research is given in following example with the final results. Two transformers with following data are taken:

SnI=100kVA	SnI=250kVA
$U_m = U_{1I} = U_{1II} = 10$kV	
$U_{2I} = 395$V	$U_{2II} = 400$V
$k_{12I} = 25.3$	$k_{12I} = 25$
$u_{kI} = 3.7\%$	$u_{kI} = 4.1\%$

According to above mentioned equations the results are given on Eq.(21) and on Table I:

$$S_{max}=325.61\text{kVA} \qquad \Delta k=1.193\% \qquad \eta=0.93 \quad (21)$$

TABLE I
RESULTS OF MATHEMATICAL CALCULATION
FOR PARALLEL WORK OF TWO TRANSFORMERS

	T I	T II
I_n [A]	146.068	360.84
I_{op} [A]	146.068	325.64
I_i [A]	32.5885	32.5885
i_i [%]	22.31% I_{opI}	10% I_{opII}
I [A]	I_I=113.448	I_{II}=358.22

The scheme of connection of these two transformers with all data is given on Fig. 3.

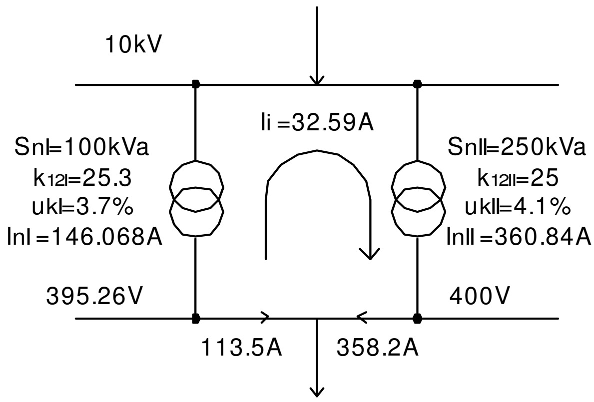


Fig. 3. Scheme of connection of two transformers in parallel operation

I must mention that current limiting for second transformer is necessary. Equalizing current increase the value of current and make it bigger than the current of loading. It can cause smaller max. allowed power S_{max}.

B. Experimental verification

Experimental verification is made on smaller transformers in laboratory with rated power to 3000VA. The results match the calculations.

Purpose of this research are bigger energetic transformers. We demanded a permission from Electro-Bitola to allow measurement at two substations where parallel operation exists. We got information that it will be possible in near future. This measurement will verify our mathematical model for calculating parallel operation.

It is planed the measurement to be performed with measure information system that can follow all relevant data at one substation. All results are on computer or we can get it on flash memory. Then the comparison will be done.

V. CONCLUSION

When we have parallel operation of transformers in substation where already exists two or more transformers analysis can be made in direction of determination of equalizing currents.

In case when in some substation there is necessity of new transformer this calculations are very useful and important. Calculation of max allowed power and divided power of each transformer will give us information about value of rated power that should be selected.

Main goal is proper selection of transformer which will be used more, have bigger coefficient of efficiency of the whole group and have the smallest loses. All this, if is fulfilled it will justify transformer installation. In this paper work subject of wide research is parallel operation of transformers in practice, allowed frames of tolerance of recommended values important for calculation.

As a conclusion from this part the following can be established:

- all limitations caused by allowed deviation does not limit the power under value of rated power of any transformer because there is no need of one of transformers.
- values of equalizing currents to be as lower as possible, so that the losses will be smaller
- the transformers should be chosen from same manufacturers or if it is not possible, the difference between impedance voltages to be smaller
- loses will be smaller if efficiency of each transformer and whole group is bigger.

VI. ACKNOWLEDGEMENT

The measurements on smaller transformers were performed at the Technical Faculty of Bitola, for which the author would like to thank this institution.

Also we acknowledge to “Electro Bitola” – Bitola for their interest about these measurements and their readiness for help in every way.

REFERENCES

[1] M. P. Kostenko, L. M. Piotrovski "Эlektriчeskie машiны" - Leningrad, 1973
[2] Dragutin Kaiser "Elektrotehnički Priručnik" Zagreb, 1964
[3] Prof. d-r Lidija Petkovska "Electrical machines and trans– formers" - lectures, Skopje, 2000.
[4] "Technical handbook" - Rade Koncar, Zagreb, 1984.

Dimming of High Pressure Discharge Lamps

A. Pachamanov[1], G. Todorov[2], N. Ratz[3]

Abstract - **The paper discusses a low cost approach for dimming of high-pressure discharge lamps (HPDL) used for street and tunnel lighting. Possibilities of utilizing of variable inductance reactor for reducing the discharge lamps flux are analyzed. The results are compared with three other methods used for group dimming.**

Keywords - **high-pressure discharge lamps, dimming, reactor**

I. Introduction

It is a useful practice to reduce the luminous flux of the street lighting late at night. The traffic during these late hours is low and luminance approximately 30-50 % of the rated is necessary to provide for safety traffic. Saving of 30-50 % of electric energy is realized in consequence of dimming. There is another profit in dimming – the exploitation of lighting devices with reduced supply voltage prolongs the life of the HPDL and its starting devices.

Different methods for dimming ware applied – control device with variable autotransformers, device with thyristor switch, transistor inverters with PWM. It was analyzed [1,2] that utilization of inverters with PWM gives best performance characteristics and stability of the high-pressure discharge lamps. This method may be chosen as a best decision for high-power lighting systems, but the relatively high price of the inverter makes it too expensive for application with relatively low power. Approximately the same characteristics are obtained at dimming devices with variable autotransformers but they are with very big weight and dimensions. In the present paper a method for dimming is proposed with utilization of variable inductance reactor and the performance characteristics of HPDL are simulated and compared to the other methods.

II. Dimming Methods

Dimming devices utilizing booster transformer and variable autotransformer to control the voltage supply of the street lighting are known. The devices are produced to control the luminous flux for single-phase and three-phase power supply in wide range [4]. With the aid of the autotransformer the device can increase or reduce the voltage supply.

The three-phase version includes three booster transformers in series with the load and three single-phase variable autotransformers, connected across the corresponding line wire and neutral – Fig.1.

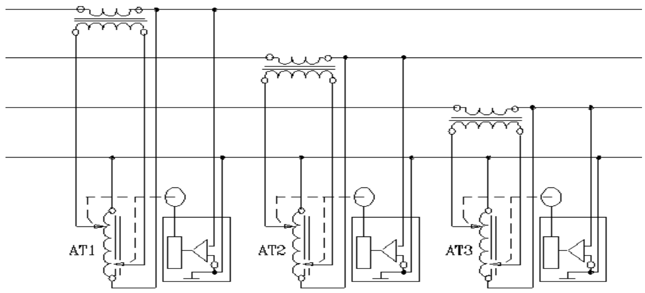


Fig.1. Dimming device for variation of voltage supply with booster transformers and variable autotransformers

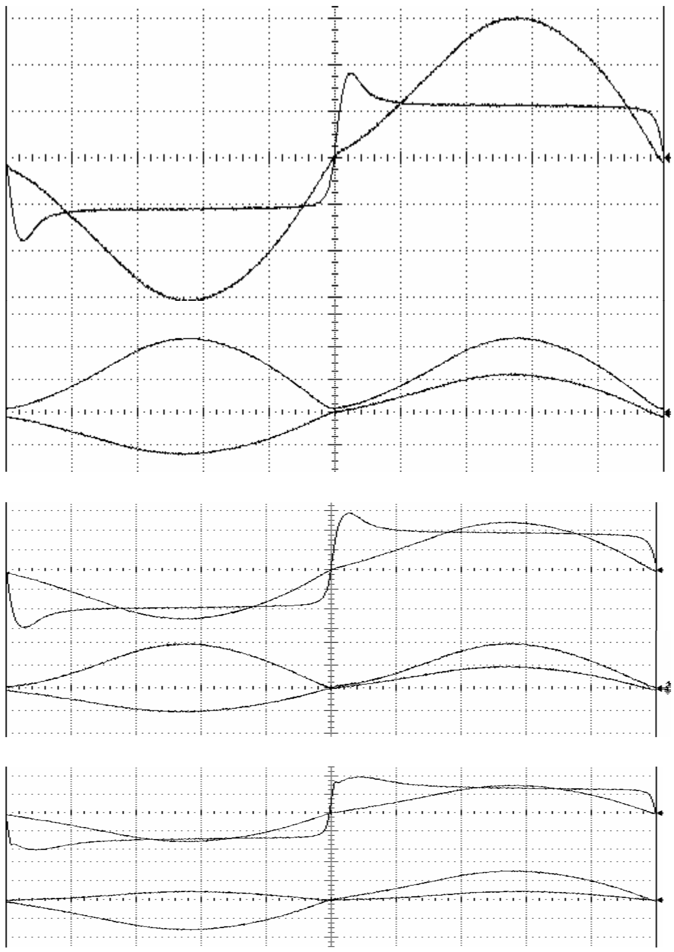


Fig. 2. U_{supply}- I_{lamp} , U_{lamp} -I_{lamp} and Φ- I_{lamp} oscillographies for 400W HPDL at dimming with boost and autotransformer at 230, 180 and 120V~

[1] **Angel Pachamanov** is with The Faculty of Electrical Engineering, Technical University, Kliment Ohridski 8, 1000 Sofia, Bulgaria, E-mail: pach@tu-sofia.bg

[2] **George Todorov** is with The Faculty of Electrical Engineering, Technical University, Kliment Ohridski 8, 1000 Sofia, Bulgaria, E-mail: gtto@tu-sofia.bg

[3] **Neli Ratz** is with The Technical Colleage at the Technical University, Kliment Ohridski 8, 1000 Sofia, Bulgaria

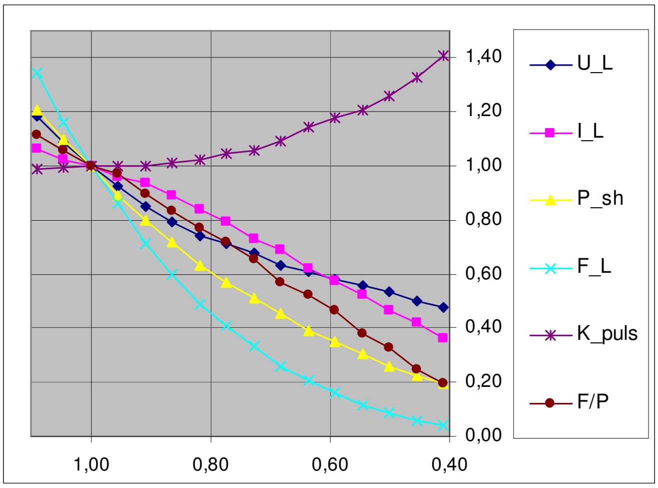
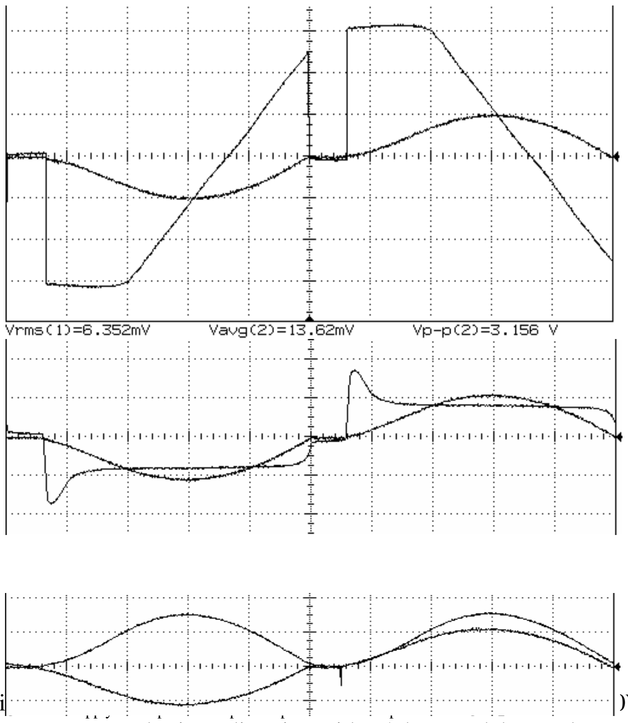


Fig. 3. Performance characteristics of 400W HPDL with inductive ballast at dimming with variation of the supply voltage amplitude via boost transformer and autotransformer

In the figure:
U_L is the voltage across the lamp, I_L is the lamp current, P_sh is the load power, F_L is the luminous flux, K_puls is the flux pulsation coefficient , F/P is the luminous flux yield.



Fi)W
HPDL at thyristor dimming with 75 degrees delay angle

Dimming devices with thyristor switch has been widely used due to its simplicity and low price. Different level of dimming is achieved controlling the delay of firing angle. An experimental example for application of this method is shown in Fig.4. The delay of firing angle is 75 degrees. It is seen that the current flowing through the lamp is equal to zero for a certain period of time, which causes corresponding pulsation in the luminous flux.

The bigger is the delay angle the bigger is luminous flux pulsation and this is main disadvantage of the method (Fig. 5). Recently some frequency converters with PWM (Fig. 6.) were developed and applied for dimming tunnel lighting [3].

Simulations and experimental investigations for 400 W HPDL with inductive ballast were made. The frequency converter supplies the lamp with a variable frequency AC voltage. Controlling the frequency between 50 Hz and 140 Hz the luminous flux were reduced from the rated value to approximately 10-20 % of the rated. The results of simulation and experimental investigations prove very smooth and lossless dimming process.

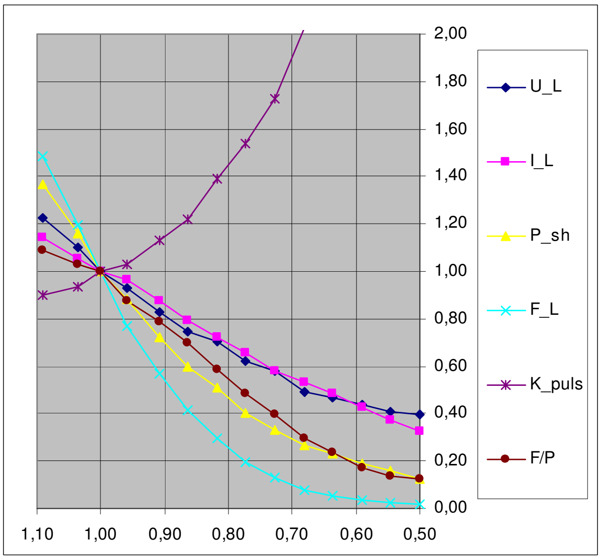


Fig. 5. Characteristics of HVDL 400W at thyristor dimming

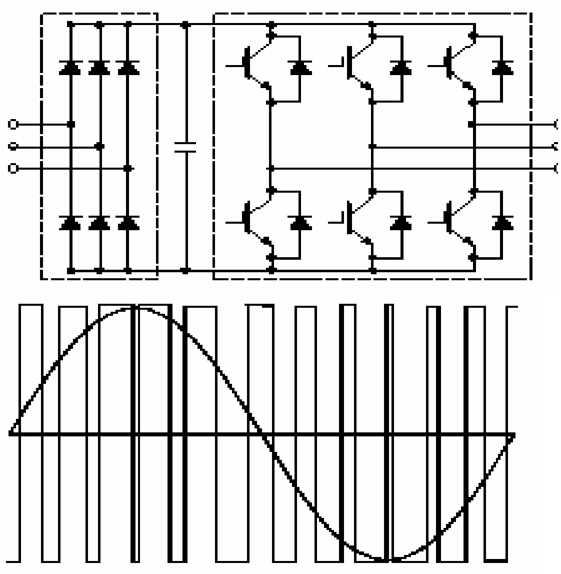


Fig.6. Sinusoidal inverter output voltage via PWM

The disadvantage of this tape of dimming device is its relatively high price, compared to the other components of the system. For this reason it can be used for applications in long tunnels or streets with high-power lighting systems. It is possible to use conventional frequency converters (manufactured for AC motor drives) for applications in low power lighting systems to avoid the development expensive application oriented inverter. Unfortunately the conventional three-phase inverters have been developed for symmetrical machines and do not have neutral wire. In case of non-symmetrical load the neutral will "float" and it is pretty possible some lamps to operate at supply voltage bigger than the rated. This will result in shorten lamps life and will increase the exploitation costs of the lighting.

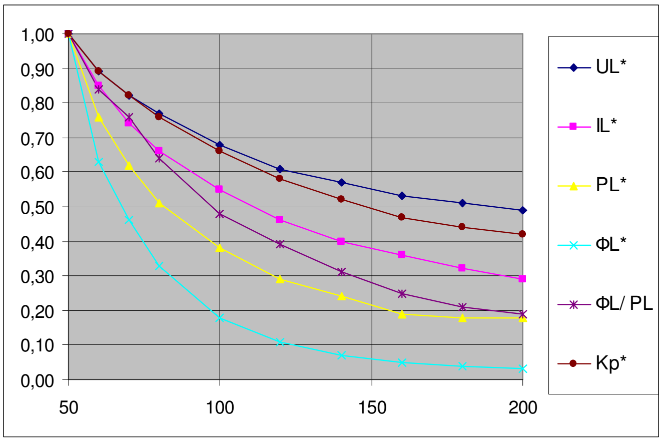


Fig. 7. Characteristics of 400 W HVDL at frequency dimming

The performance characteristics of the HPDL under operation with autotransformer dimming device are like characteristics with frequency converter as it is shown in Fig.3 and Fig.7. A comparison of pulsation for the three dimming methods is given in Fig. 8. It is clearly seen that the frequency inverter dimming makes less pulsation – the increase in frequency of the supply voltage leads to pulsation coefficient reduction while the thyristor switch makes pulsation to increase rapidly.

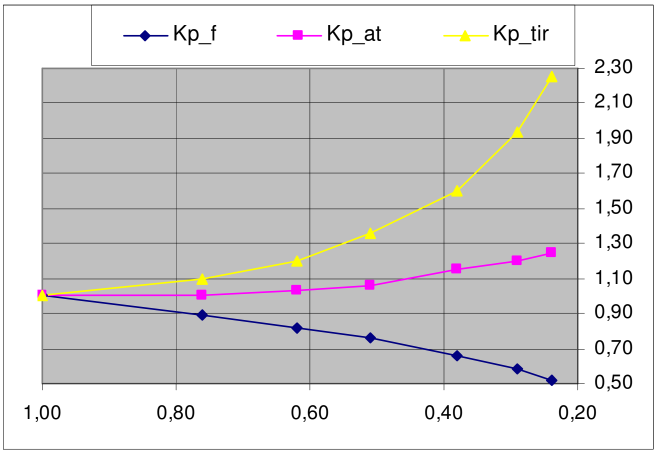


Fig. 8. Pulsation coefficient at dimming with autotransformer (at), thyristor switch (tir) and frequency inverter (f)

The voltage across the lamp for the three dimming methods is given in Fig. 9. It is seen that under thyristor switch dimming the reduction of the luminous flux increases the voltage across the lamp (arc discharge voltage). This causes intensive disintegration the lamp electrodes and shortens the lamp life. According this feature dimming with reduction the voltage supply with booster transformer and autotransformer is superior to frequency inverter dimming. The luminous flux yield for both methods is equal (Fig. 10).

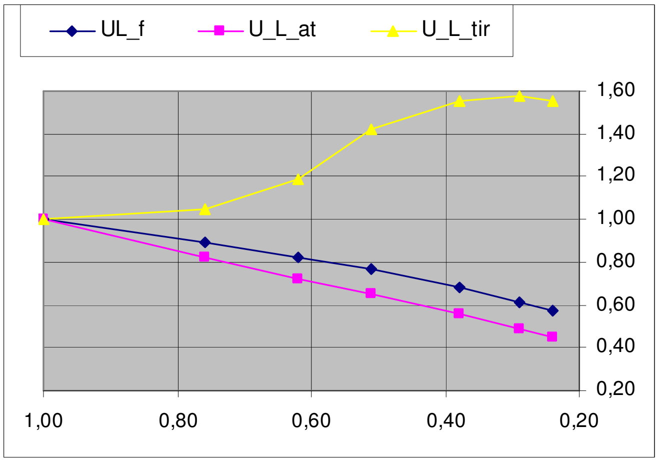


Fig. 9. Voltage of the discharge lamp at dimming

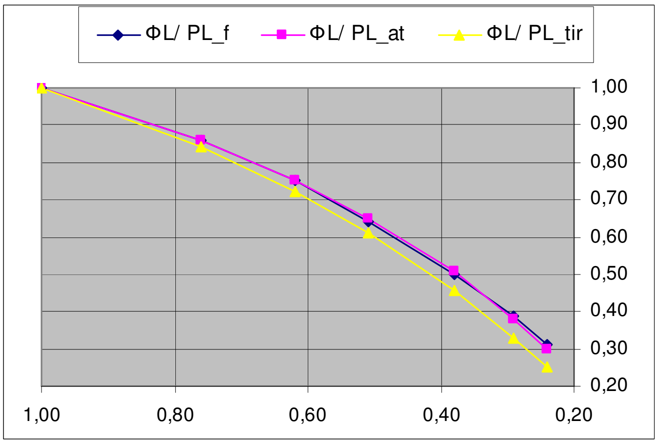


Fig.10. Luminous flux yield at dimming with autotransformer (at), thyristor switch (tir) and frequency inverter (f)

According to the analysis and comparison given above it is reasonable to conclude that the method with reduction of the supply voltage with booster transformer and variable autotransformer can be chosen for dimming the street and tunnel lighting. An approach to reduction the price of the dimming device is proposed in the present paper – to replace the booster transformer and variable autotransformer with a variable reactance reactor for group dimming of a set of lamps.

III. Reactor Dimming

Normally the dimming is used to make the illumination operate with voltage equal to the rated and lower than the rated. There is no need to increase the voltage supply. This makes possible to simplify the dimming device and use only a variable reactor connected in series with the lamps. With a control of the reactor's reactance the current flow through the lamps is controlled similar to the autotransformer's device. Simulations for a set of six HPDL were made. Each lamp is with power of 400 W and is equipped with 146 mH inductive ballast and 45 uF capacitive compensation. The results are shown in Table 1 and Fig. 11. It was found that a single reactor with 70 mH reactance can provide dimming of the luminous flux from rated value to 20 % of the rated for the whole set of lamps.

TABLE 1
DIMMING WITH A VARIABLE REACTOR

L_R	U_{supply}	I_{lamp}	U_{lamp}	I^*_{lamp}	U^*_{lamp}	Φ^*
mH	V	A	V	-	-	-
0	230	4,15	110	1	1	1
10	220	3,97	105	0,956	0,955	0,87
17	210	3,79	100,5	0,914	0.914	0,76
23	200	3,62	96	0,872	0,872	0,66
29,5	190	3,44	91	0,828	0,827	0,57
36	180	3,25	86	0,784	0,784	0,48
43	170	3,06	81	0,737	0,737	0,40
49	160	2,89	77	0,697	0,697	0,34
57	150	2,71	72	0,652	0,652	0,28
64,5	140	2,53	67	0,610	0,610	0,23
70	134	2,42	64	0,582	0,582	0,20

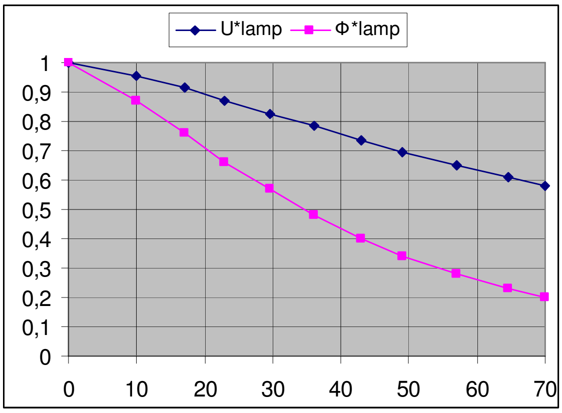


Fig.11. U_lamp and F_lamps at dimming with reactor 0-70 mH

These results were compared to the experimental investigations made with autotransformer's dimming device. The per-unit luminous flux versus per-unit lamp voltage is shown in Fig.12 for both methods. The relative difference is given as "eps" in the same figure.

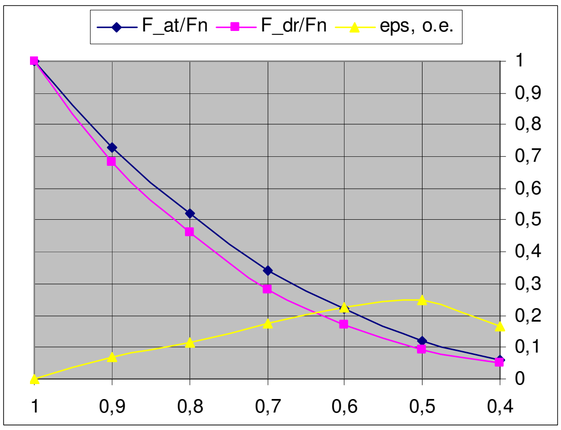


Fig.12. F_lamps/F_nom_lamps as function of supply voltage at dimming with an autotransformer (at) and with a variable reactor (dr)

The advantage of reactor dimming is the possibility for lamps to operate with individual capacitive compensation, which makes the line current approximately 50 % of the current through the lamp and allows reducing the supply wire's cross section. For comparison it should be mentioned that at dimming with frequency converter, group power factor capacitive compensation for the whole set of lamps is used as the capacitors are included in the inverter circuit.

Some problems have to be solved for practical application of the proposed method:

- to chose the type of reactor – multi-tap reactor or direct-current controllable reactor;
- to implement adequate feedback for keeping the constant voltage across the set of lamps in case of accidental change load, e.g. disconnecting one of the lamps.

Utilization of multi-tap reactor is cheaper decision but the tap-changing should provide continuous current flow through the lamps. Application of direct-current controlling reactor will make the device to operate smooth and safe but this reactor is more expensive. Additional economic calculations and experiments should be made to make the final decision.

IV. CONCLUSION

Methods for dimming the street and tunnel lighting are analyzed and compared in the paper. It is obtained that dimming with reduction of the supply voltage amplitude is expedient method for application at relatively low power lighting systems. A method for group dimming with reduction of the supply voltage amplitude is proposed. Variable reactance reactor is used to control the level of dimming for a set of HPDL with individual inductive ballast and capacitive compensation. Simulation of the lamp's performance characteristics is made and compared to the experimental results for dimming device with booster transformer and variable autotransformer. Good agreement in the results is obtained. Both methods provide smooth and reliable operation of the lighting but the proposed dimming device with variable reactor is with reduced dimensions, weight and price.

REFERENCES

[1] Pachamanov A., B.Pregyov, D.Pachamanova, N.Ratz, Dimming of Artificial Lighting in Threshold and Transition Zones of Road Tunnels. Balkan Light'02" "Energy Saving and New Trends in Lighting",3-4 October 2002, Istanbul.

[2] N.Ratz, A.Pachamanov. Research of possibilities for dimming of high pressure discharge lamps. XII National Conference „Light'2004", 15-17 June 2004, Varna-Bulgaria

[3] www.tunnellighting.com

[4] ILUEST – Stabilizer-Reducer for Discharge Lamps Lighting Systems - Catalogue, Barcelona, Spain.

New Method for Allocation of Losses in Distribution Systems with Dispersed Generation

Metodija B. Atanasovski[1] and Vesna P. Borozan[2]

Abstract – **This paper presents a new method for allocation of losses in distribution networks with dispersed generation (DG). The method establishes direct relationship between losses in each branch of the network and injected active and reactive power in the nodes on which path to the power supply point (PSP) the branch is placed. It is assumed that the power flows in each branch of the distribution network are constant and equal to the sum of injected active and reactive power in each node on which path to the PSP, the branch is placed. DG in this method is treated as negative power injection. Application of the proposed method on real distribution network will be presented.**

Keywords – **Allocation of losses, dispersed generation, distribution systems.**

I. INTRODUCTION

Electricity supply industries worldwide are undergoing major structural changes with the basic objective of introducing competition and choice in electricity supply. An important feature of the market-based structure of electricity industry is the separation of the generation, transmission, distribution and supply segments of the power systems into autonomous business units. Prices in the generation and supply segments are determined through market mechanisms, whereas those in the natural monopoly segments of transmission and distribution are regulated. An essential condition for developing of competition is open access on a nondiscriminatory basis, to transmission and distribution networks. The central issue in open access concept is setting an adequate price for transmission and distribution services and clear identification and allocation of costs for these services to their users, avoiding or minimizing any possibilities for temporal or spatial cross-subsidies. Main operative cost of any transmission or distribution network are electric energy losses and the issue of their appropriate allocation to network users is of great importance for networks efficiency in market conditions.

In parallel with the structural changes described, another not less important development is the growing penetration of DG into power distribution systems. The Working Group 37-23 of CIGRE has summarized in literature [1] some of the reasons for an increasing share of DG in different countries. The presence of DG in distribution systems alters radically the way these networks should be considered from both technical and commercial aspects. This is because DG changes distribution networks from passive networks, with unidirectional power flows from higher to lower voltage levels, into active ones with multidirectional power flows.

This paper is primarily concerned with the allocation of variable network losses in distribution network with DG. According to structural changes mentioned above and introducing DG into distribution networks, the problem of allocation of losses become very important problem. In literature [2], requirements for ideal loss allocation method are summarized as follows: 1) *Economic efficiency:* Losses must be allocated in a way to reflect the true cost that each user imposes on the network; 2) *Accuracy, consistency and equity:* Loss allocation method must be accurate and equitable i.e. must avoid or minimize cross subsidies between users and between different time of use; 3) *Utilization of metered data:* From a practical standpoint it is desirable to base allocation of losses on actual metered data; 4) *Simplicity of implementtation:* For any proposed method to find favor it is important that the method is easy to understand and implement.

A large number of different methods for loss allocation have been published. Most of them are dedicated to transmission networks. In literature [3] and [4] authors introduce a basic assumption of proportionality that they use to determine the proportion of active power flow in a transmission line contributed by each generator. This limits the applicability of these methods to distribution networks, because when reactive power is not considered in the allocation process, cross subsidies will emerge. Also this methods neglect the crossed terms, which due to the fact that the power losses of specific branch are quadratic function, must be carefully allocated and can not be neglected. Another approach for allocating losses is the substitution method, which has been used in England and Wales. There are number of problems associated with this method, which are summarized in literature [2]. In order to overcome the problems with the substitution method, authors in [2] present marginal loss coefficients (MLC) method. By definition MLC measure the change in total active power losses due to a marginal change in consumption/generation of active and reactive power at each node in the network. In [5] authors present a method that applies the same concept as in [2] but from a different point of view. The proposed method determines the prices at different nodes in the distribution networks using nodal factors. In addition the proposed method is compared with classical distribution network pricing scheme which is used in most regulatory concepts in the world.

The proposed method in this paper is branch-oriented as [3] and [4] and tries to resolve the referred difficulties mentioned above. Method establishes direct relationship between losses in each branch of the network and injected

[1]Metodija B. Atanasovski is with the Faculty of Technical Sciences, I.L.Ribar bb, 7000 Bitola, Macedonia, E-mail: metodija.atanasovski@uklo.edu.mk

[2]Vesna P. Borozan is with the Faculty of Electrical Engineering, Karpos II bb, 1000 Skopje, Macedonia, E-mail: vesnab@etf.ukim.edu.mk

active and reactive power in the nodes on which path to the power supply point (PSP) the branch is placed. PSP is the reference bus where the transmission network connects to the distribution network. DG in this method is treated as negative power injection. It is assumed that the power flows in each branch of the distribution network are constant and equal to the sum of injected active and reactive power in each node on which path to the PSP, the branch is placed. The network is considered lossless. Crossed terms are carefully allocated using the quadratic method for allocation of crossed terms proposed in literature [6]. Application of the proposed method on real distribution network is presented. Results are discussed according to the requirements for ideal loss allocation method.

II. PROPOSED METHOD

Let us consider distribution system with DG of Fig. 1. In order to simplify the problem, we first make the following assumptions:

- An AC power flow solution is available from on-line state estimation or off-line system analysis. The network is in steady state operation. There is only one path between each node of the distribution network and PSP, consisted of known set of branches.
- Appropriate indexing of network nodes and branches is done. Every branch of the network has its beginning and ending node. Beginning node is the node which contains less branches on his path to the PSP. The index of beginning node of one branch is smaller than the index of its ending node. Index of the branch is equal with the index of its ending node.
- Node voltage and injected active and reactive power in the network are known after power flow solution. A DG has the priority to provide power to the load on the same node. DG is treated as negative power injection. For the purposes of the method the following definition of injected active and reactive power in each node i , has been applied:

$$\begin{aligned} P_i &= P_{i(consumers)} - P_{i(DG)} \\ Q_i &= Q_{i(consumers)} - Q_{i(DG)} \end{aligned} \tag{1}$$

- Active and reactive power flow in each branch of the network is constant and is equal to the sum of injected active and reactive powers to each node of the network on which path to the PSP the branch is placed. The network is considered “lossless”.

According to these assumptions, active power losses on each branch of the network with index k, are calculated with following formula:

$$\Delta P_k = R_k \cdot \frac{\left[\left(\sum_{i=k}^{Z} P_i\right)^2 + \left(\sum_{i=k}^{Z} Q_i\right)^2\right]}{U_k^2} \tag{2}$$

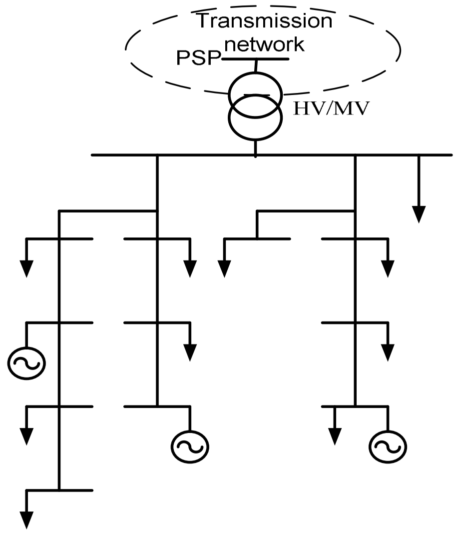


Fig. 1. Radial distribution network with DG

where:

R_k -is the resistance of the branch with index k;

U_k -is voltage magnitude of ending node of the branch with index k;

$\sum_{i=k}^{Z} P_i$ -is a sum of injected active power in the nodes on which path to PSP the branch k is placed;

$\sum_{i=k}^{Z} Q_i$ -is a sum of injected reactive power in the nodes on which path to PSP the branch k is placed;

Z –is a number of nodes on which path to PSP the branch k is placed.

The equation (2) can be developed in a form which shows that injected power in node i have impact in the following terms:

$$\Delta P_k^i = R_k \cdot \frac{\left[P_i^2 + 2 \cdot P_i \cdot \left(\sum_{\substack{j=k \\ j\neq i}}^{Z} P_j\right) + Q_i^2 + 2 \cdot Q_i \cdot \left(\sum_{\substack{j=k \\ j\neq i}}^{Z} Q_j\right)\right]}{U_k^2} \tag{3}$$

These terms have two different origins. Some are due to injected active and reactive power in node i, P_i^2 and Q_i^2. Other two terms are result of simultaneous influence of injected active and reactive power in node i and injected active and reactive power in other nodes on which path to PSP branch k is placed. These terms are called crossed terms and must be carefully allocated [6]. As referred in [6], these crossed terms can be allocated in a proportional, quadratic and geometric way. According to a quadratic relationship between losses and power, quadratic way of allocation of crossed terms has been adopted here. For better understanding of allocation of crossed terms the simplest case of single line will be considered, in which two transactions P_i and P_j create power flow respectively [6]:

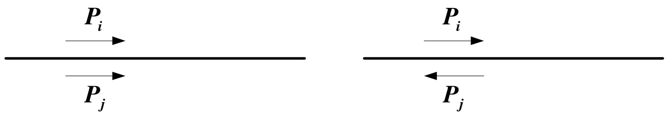


Fig. 2. Allocation of crossed terms

Two cases are considered in Fig. 2, when transactions have same and opposite directions. For both cases, the following approximate expression can be obtained:

$$\Delta P \cong \left(P_i \pm P_j\right)R = \left(P_i^2 + P_j^2 \pm 2P_iP_j\right)R \tag{4}$$

Allocation of losses between transactions is:

$$\Delta P^i = \left(P_i^2 \pm \beta_i P_i P_j\right)R \text{ ;} \qquad \Delta P^j = \left(P_j^2 \pm \beta_j P_i P_j\right)R \tag{5}$$

Mandatory balance of power can be simply written as:

$$\beta_i + \beta_j = 2 \tag{6}$$

Since power losses grow quadratic with power flows, it is also reasonable to impose the following constraint:

$$\frac{\beta_i}{P_i^2} = \frac{\beta_j}{P_j^2} \tag{7}$$

which, combined with (6), yields:

$$\beta_i = 2\frac{P_i^2}{P_i^2 + P_j^2}; \qquad \beta_j = 2\frac{P_j^2}{P_i^2 + P_j^2} \tag{8}$$

The same considerations can be applied for reactive power flows, which are not applied in [7], because dc power flow technique is used. Using the quadratic allocation of crossed terms (8), we can obtain losses of the branch *k*, allocated to node *i*:

$$\Delta P_k^i = R_k \cdot \frac{\left[P_i^2 + 2\cdot P_i \cdot \left(\sum_{\substack{j=k \\ j\neq i}}^{Z} P_j \frac{P_i^2}{P_i^2 + P_j^2}\right) + Q_i^2 + 2\cdot Q_i \cdot \left(\sum_{\substack{j=k \\ j\neq i}}^{Z} Q_j \frac{Q_i^2}{Q_i^2 + Q_j^2}\right)\right]}{U_k^2} \tag{9}$$

Total value of losses allocated to node with index *i*, will be a sum of losses allocated to it in each branch of the network placed on its path to the PSP, obtained with equation (9), as follows:

$$\Delta P^i = \sum_{k\in\alpha} \Delta P_k^i \tag{10}$$

where α is set of branches placed on the path of node *i* to PSP.

The assumptions and approximations made in the computation of proposed method give rise to very small differences between the losses calculated from application of the method and those calculated with power flow:

$$\Delta P \cong \sum_{i=1}^{n} \Delta P^i \tag{11}$$

where *n* is the total number of nodes of the network. These small differences are result of the neglecting generated reactive power from distribution lines because of the existence of their shunt capacitances b, with the proposed method. Distribution lines compensate part of the reactive power that flows through them and because of these, the sum of allocated losses with the proposed method is greater than the losses calculated with power flow. With this distribution network directly subsidize consumers, compensating part of the reactive power demanded by them. There is no need for fundamental reconciliation, but these small differences can be easily resolved by introducing constant multiplier reconciliation factor k_0. This factor is calculated as follows:

$$k_0 = \frac{\Delta P}{\sum_{i=1}^{n} \Delta P^i} \tag{12}$$

Finally total system active power losses allocated to nodes is:

$$\Delta P = \sum_{i=1}^{n} k_0 \cdot \Delta P^i \tag{13}$$

III. Application of the Proposed Method

Let us consider real radial 10 kV distribution network of Fig. 3, which is a part of the distribution network of Distribution Company-Bitola. There are different types of loads in the nodes of this network and to provide realistic load variations, several customer types are included in the analysis, using typical daily load profiles. Application of the proposed method is performed on hourly basis for two extreme cases: typical winter working day and summer Sunday. As input data for allocation of losses with the proposed method, power flow calculations are used. DG is working with constant output and constant power factor during the considered typical days.

Results for allocation of losses on hourly basis with the proposed method are given in Fig. 4 and Fig. 5. Fig. 4 shows the allocation of losses profiles in kW for network nodes, calculated with the proposed method for typical winter working day. Consumers located at nodes with positive allocation of losses will be charged for losses. Conversely, those consumers at nodes with negative allocation of losses will be rewarded. It is clear that between the hours 7:00 and 24:00 and between the hours 0:00 and 1:00 (19 hours in total), the DG at nodes HEC Filternica and TS Pumpi Vodovod are rewarded for there contribution to system loss reduction (signified by negative allocation of losses for this node during this period). However during early hours of the day between the hours 2:00 ans 6:00, when load is low, DG HEC Filternica contributes to increasing losses and therefore it is duly charged (signified by positive allocation of losses for this node during this period). During the same period DG in node TS Pumpi Vodovod is rewarded for its contribution to system loss reduction, because it directly supplies the load in the same node.

To illustrate spatial distribution of loss allocation with the proposed method, two time periods (operating points) were chosen, 2^{nd} and 12^{th} hour in a winter working day. Each node has different loss allocation, which can be positive or negative. At 2:00 the load is relatively low and therefore the DG at node HEC Filternica for example, increases total system losses and the allocation of losses is positive. At 12:00 hour, an opposite situation is presented. The load is relatively high and the DG at same node contributes to reducing the total system losses.

Fig. 5 shows the allocation of losses profiles in kW for network nodes, calculated with the proposed method for typical summer Sunday. It is clear that between the hours 0:00 and 24:00 (24 hours in total), the DG at node HEC Filternica is rewarded for its contribution to system loss reduction.

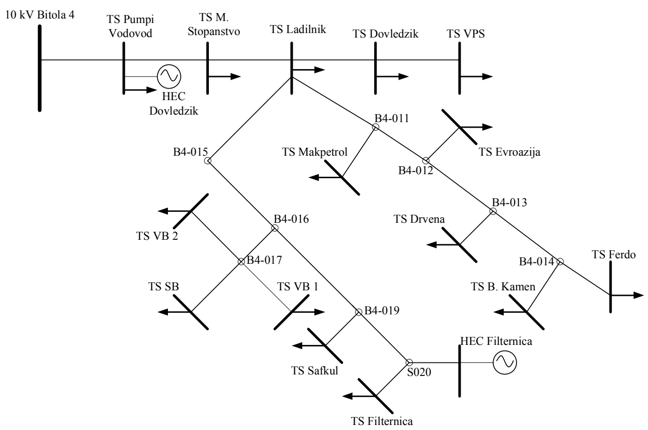


Fig. 3. Case studied 10 kV distribution network

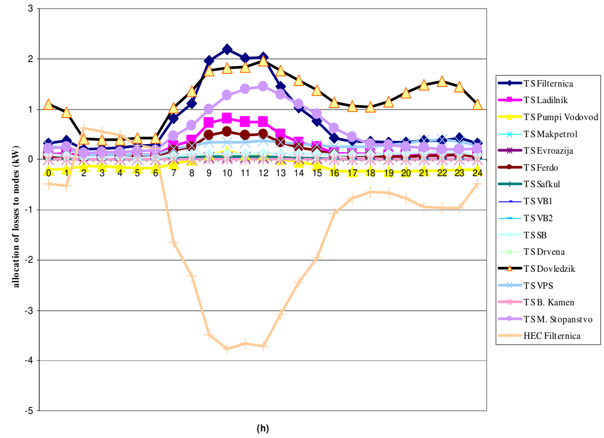


Fig. 4. Allocation of losses profiles for nodes for typical winter working day

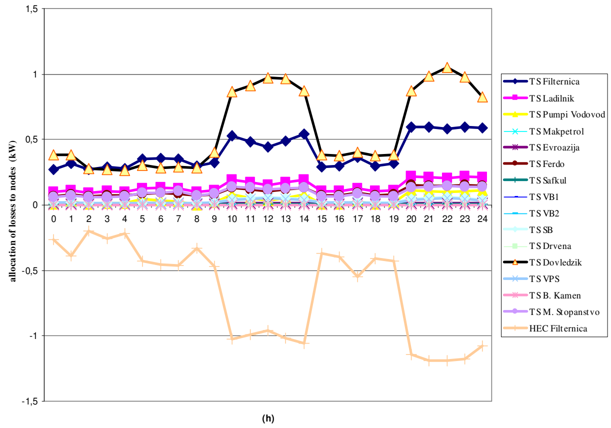


Fig. 5. Allocation of losses profiles for nodes for typical summer Sunday

IV. Conclusion

In this paper new method for allocation of losses in distribution systems with dispersed generation has been proposed. It has been demonstrated that the allocation of losses with this method is location specific and vary in time. Furthermore, allocation of losses with the proposed method can be positive or negative depending on the user particular impact on losses at any given time. The presented method is simple and easy for implementation and understanding. It can use all metered data in the network.

Application of proposed method has been illustrated on real distribution network. The results obtained clearly demonstrate that allocation of losses with this method change with time of day and from one location to another. The most appropriate approach for implementation of method is on hourly bases.

References

[1] Final Report WG 37-23, “Impact of Increasing Contribution of Dispersed Generation on the Power System”, *CIGRE Study Committee no 37*, September 1998.

[2] J. Mutale, G. Strbac, S. Curcic, and N. Jenkins, “Allocation of losses in distribution system with embedded generation", IEEE Proc.-Gener. Trans. Distrib., vol.147, no.1, pp. 7-14, Jan. 2000.

[3] J. Bialek, “Tracing the Flow of Electricity”, IEE Proc.-Gener. Transm. Distr., Vol. 143, no. 4, 1996.

[4] D. Kirschen, R. Allan and G. Strbac, “Contribution of Indiviadual Generators to Loads and Flows", IEEE Transactions on Power Systems, Vol. 12, no. 1, 1997.

[5] J. Mario Vignolo and Paul M. Sotkiewicz, “Allocation of Loss Costs in Distribution Networks: The Nodal Factor Pricing Method”, International Journal of Power Systems, Vol. 10, No. 10, 2004.

[6] A. G. Exposito, J. M. R. Santos, T. G. Garcia, E. R. Velasco, “Fair Allocation of Transmission Power Losses”, IEEE Transactions on Power Systems, Vol. 15, No. 1, February 2000.

Virtual Instrumentation Applied to Electrical Power Quality Measurement

Božidar B. Dimitrijević[1] and Milan M. Simić[2]

Abstract – **This paper describes advances in electrical power quality monitoring equipment and virtual instrumentation software tools for metrological support and analyzing power quality measurement results. Different types of power quality variations are described and the methods of characterizing each type are presented. Instruments for electrical power quality measurements are based on complex digital processing of input signals, whose waveforms are highly variable. The flicker and harmonics generated by network disturbances present a problem that is well known to both the manufacturing industry and electrical utilities. Virtual instrumentation tools can be used to record the voltage quality at various points in the system and save the data to a central database. By correlating the relevant variables and analyzing the time of events, the cause for a disturbance can be localized and eventually eliminated. Virtual instrument described in this paper, provides parameters of voltages, currents and phases, defined by electrical power quality protocols, by means of 8 channels DAQ card PCI NI 6713, comparing generated and measured values of voltage and current parameters, and automatically correcting generated values until the desired accuracy class is achieved.**

Keywords - **virtual instrumentation, electrical power quality measurement, LabVIEW**

I. INTRODUCTION

Electrical power quality, in recent years, has become an important issue and is receiving increasing attention by utility, facility, and consulting engineers. Present equipment setups and devices used in commercial and industrial facilities, such as digital computers, power electronic devices, and automated equipment, are sensitive to many types of power disturbances. The threatened limitations of conventional electrical power sources have focused a great deal of attention on power, its application, monitoring and correction [1].

Over the last twenty years, power quality has become an increasingly important issue. This is due to a number of reasons: the increased use of non-linear and pulsed loads, can leads to mains interference feedback; a growing number of electronically controlled devices, which require a supply voltage of a certain quality, respond very sensitively to mains pollution; with energy supply companies operating in a highly competitive market, the relationship between supplier and customer has become more important than ever [2]. Power economics now play a critical role in industry as never before. As the use of power electronics and other devices that put a strain on the network is on the increase, electricity suppliers need detailed data on the quality of their network.

[1]Božidar B. Dimitrijević is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia and Montenegro, E-mail: dimitrijevic@elfak.ni.ac.yu

[2]Milan M. Simić is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia and Montenegro, E-mail: milansm@elfak.ni.ac.yu

With the high cost of power generation, transmission, and distribution, it is of paramount concern to effectively monitor and control the use of energy. The increased concern for power quality has resulted in significant advances in monitoring equipment that can be used to characterize disturbances and power quality variations [3].

In the past, voltage quality measurements were only carried out when required, for example after a complaint by a customer who had experienced problems. In such cases, the quality was measured and the results were analyzed. If no special events such as a power failure occurred during the time of measurement, it was basically impossible to establish what had caused the problem in the first place. Today, the market and general good practice of continued quality control require electricity suppliers to permanently monitor the voltage quality in their networks [2]. Computer substations, and manufacturing or process plants now have the tools to monitor power quality and consumption to aid in achieving reliable and cost effective operations. Results from power quality monitoring allow for electrical power supplier selection and electronic load balancing within a plant.

II. ELECTRICAL POWER QUALITY VARIATIONS

It is important first to understand the kinds of power quality variations that can cause problems with sensitive loads. There is a standard: BS EN 50160:2000 Voltage characteristics of electricity supplied by public distribution systems [4], that provides the limits and tolerances of various phenomena that can occur on the mains. A summary of the criteria for the low voltage side of the supply network is given in Table 1.

TABLE I
A SUMMARY OF THE ELECTRICAL POWER QUALITY CRITERIA

Supply voltage phenomenon	Acceptable limits
Grid frequency	47Hz to 52Hz
Slow voltage changes	230V ± 10%
Voltage Sags or Dips (≤1min)	10 to 1000 times per year (under 85% of nominal)
Short Interruptions (≤3min)	10 to 100 times per year (under 1% of nominal)
Accidental, long interruptions (>3min)	10 to 50 times per year (under 1% of nominal)
Temporary over-voltages (line-to-ground)	Mostly < 1.5 kV
Transient over-voltages (line-to-ground)	Mostly < 6kV
Voltage unbalance	Mostly 2% but occasionally 3%
Harmonic Voltages	8% Total Harmonic Distortion (THD)

Some practical conclusions can be made from this data: the limits are wide, perhaps more than one would expect; it important to check that safety interlocks and relays can reset after a sag or dip, because such incidents can occur quite frequently; control and process equipment that is voltage sensitive should be applied with caution because the allowable voltage tolerance is very wide and can drift outside ±10% for 5% of the time; the transient tolerances are high and so the use of surge-protection devices should be carefully considered, especially where manufacturing of high-cost components or processes involving lengthy and expensive restart times are concerned.

Categories of these variations must be developed with a consistent set of definitions so that measurement equipment can be designed in a consistent manner and so that information can be shared between different groups performing measurements and evaluations. A voltage dip is defined as a sudden drop of the voltage to a level below 90% of the nominal voltage, followed by an increase to a level above 90% within a period of 10ms to 60s. Most dips in medium or high voltage networks are over within less than 0.2s. Short dips can lead to serious problems such as failure of machine control systems (loss of production) or damage to motors caused by the sudden torque change. Most dips are caused by faults in the insulation in medium and high voltage networks [2]. Other relatively frequent causes are snow, thunderstorms, frost or branches of trees touching lines. If such an event occurs, all customers in the network are equally affected. Many dips are however caused by the consumers themselves. Voltage dips can for example occur when a great load is switched on where there is insufficient short circuit capacity at the point of common coupling. While short circuits are a common cause, voltage dips can also be caused by the start-up of large machines or capacitor banks.

Transient voltages may be detected when the peak magnitude exceeds a specified threshold. RMS voltage variations may be detected when the RMS variation exceeds a specified level. Steady state variations include normal RMS voltage variations and harmonic distortion. These variations must be measured by sampling the voltage and/or current over time. Harmonic distortion of the voltage and current results from the operation of nonlinear loads and devices on the power system. The nonlinear loads that cause harmonics can often be represented as current sources of harmonics. The system voltage appears stiff to individual loads and the loads draw distorted current waveforms. Harmonic voltage distortion results from the interaction of these harmonic currents with the system Harmonic distortion levels can be characterized by the complete harmonic spectrum with magnitudes and phase angles of each individual harmonic component. It is also common to use a single quantity, the Total Harmonic Distortion - THD, as a measure of the magnitude of harmonic distortion [1,2].

III. ELECTRICAL POWER QUALITY MONITORING SYSTEM WITH VIRTUAL INSTRUMENTATION

To understand the power quality problem better requires a comprehensive monitoring and data capturing system that is used to characterize disturbances and power quality variations. Power quality monitoring and power metering will allow plants to perform predictive maintenance, energy management, cost management, and quality control. We need a configurable and programmable high-performance PC with high-speed data acquisition units and software to perform the required task [3]. The power line monitoring system based on high-performance software and hardware, acquires and displays the voltage and current waveforms. It monitors waveforms for distortion, glitches, and frequency deviations. Such system provides a range of valuable data [2]:

- Overview of the voltage quality, showing types and levels of disturbances experienced in the network.
- Quality reports.
- Detailed, time-related recording of disturbances. Based on the data recorded, the cause of a disturbance can be identified, events can be related to problems or known events and exceptional situations can be distinguished from general seasonal trends.
- The data provides the basis of predictive maintenance. Based on indicators, some disturbances can be predicted at an early stage.

Figure 1 illustrates a typical configuration of an electrical power quality monitoring system based on virtual instrumentation software tools [3].

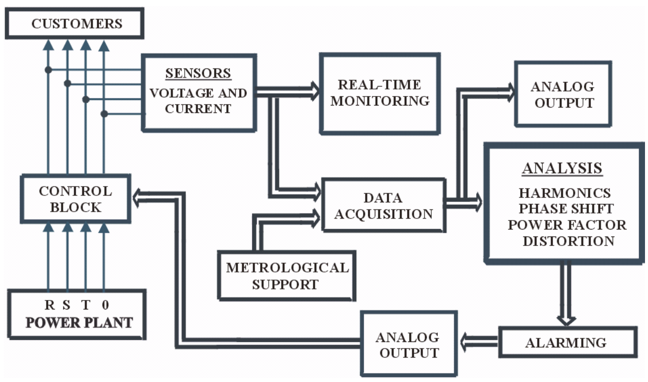


Fig. 1. Configuration of an electrical power quality monitoring system based on virtual instrumentation.

One major development resulting from the widespread adoption of the PC over the past twenty years is the concept of virtual instrumentation. A virtual instrument consists of an industry-standard computer or workstation equipped with application software, cost-effective hardware, such as plug-in boards, and driver software, which together perform the functions of traditional instruments. The development and use of programmable measurement systems have been widely explored. The possibility of modifying the measurement procedure simply by changing the algorithm executed by the computer-based architecture without replacing the hardware components makes the experimental activity easier. Virtual measurement systems have been introduced to simplify the design and implementation of programmable measurement systems by adopting a visual interface [5]. Networking has also been introduced successfully in measurement to interconnect different instruments and data processing sites into a distributed measurement system – DMS.

Currently the most popular way of programming is based on the high-level tool software. With easy to use integrated development tools, design engineers can quickly create, configure and display measurements in a user-friendly form, during product design and verification. Virtual instrumentation software tools are easily applied to electrical power monitoring applications. Control panel of virtual instrument developed in a graphical programming language LabVIEW [6], which provides parameters of voltages, currents and phases, defined by electrical power quality protocols, is shown on figure 2. A virtual instrumentation software tool provides a possibility for changing amplitude and phase of all voltage and current channels.

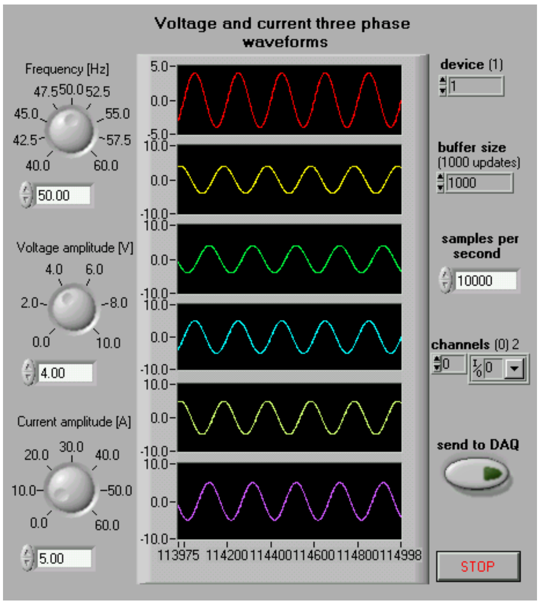


Fig. 2. Virtual instrument for generating voltage and current three phase waveforms.

Using a multi-channel microcomputer card DAQ PCI NI 6713 [7], for referent voltage and current generating, eight digital/analogue output channels are provided.. A program sequence developed using a graphical programming language LabVIEW, is shown on figure 3.

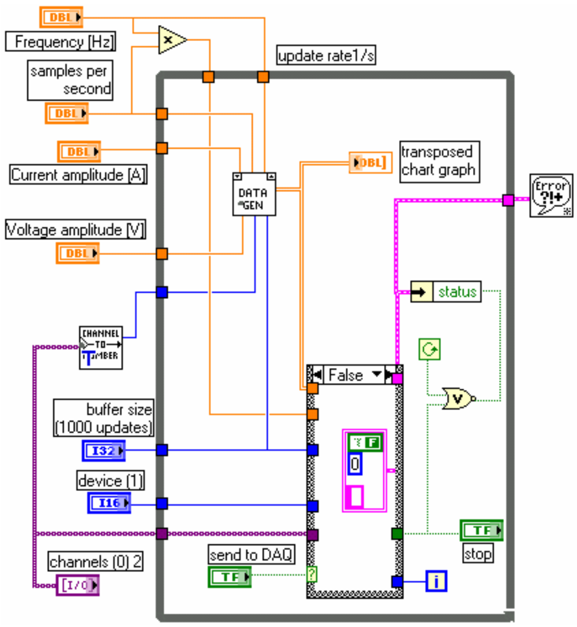


Fig. 3. A program sequence developed using a graphical programming language LabVIEW.

In most practical cases the network voltage variations influence measured quantities and that way increase measuring error. Precise setting of each current or voltage value, as well as the phase shift, requires a substantial time. In addition, most of quantities influence each other, and therefore there is need for multiple corrections. Described virtual instrument can be used as a part of the feedback loop, which in real time compares generated and measured values of voltage and current parameters, in order to correct generated values until the desired accuracy class is achieved.

Multi-channel acquisition card NI PCI 6713, generates voltage and current three phase waveforms. Measuring transducers provide information about momentary values of voltage and current, which are fit to input of acquisition card ADQ ED428. Converted values measured by acquisition card in the feedback loop are corrected for the values of obtained errors based on preset values of measured transducer coefficients [8]. Deviation of effective voltage values in time, that shows reducing of obtained error by feedback loop correction, is shown on figure 4.

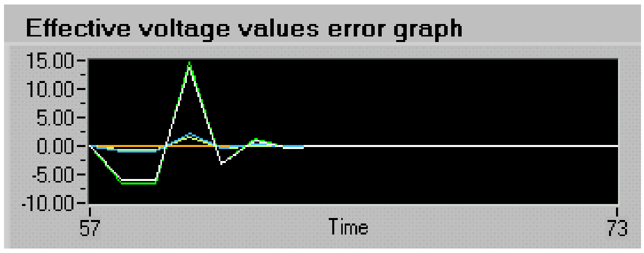


Fig. 4. Time deviation of effective voltage values.

The measurement uncertainty of voltage and current parameters settings depends both on measurement transducers coefficients stability (current shunts and voltage dividers), and the reference voltages stability of the converters in acquisition cards. Experimental measurement results of eighteen successive corrected effective voltage values, together with corresponding reduced error values, are given in table 2.

TABLE II
EXPERIMENTAL MEASUREMENT RESULTS OF CORRECTED EFFECTIVE VOLTAGE VALUES

Ordinal nomber of measurem.	Nominal effect. voltage values (V)	Measured effect. voltage values (V)	Measurement error (V)
1.	220,00	226.62	-6.62
2.	220,00	227.17	-7.17
3.	220,00	227.22	-7.22
4.	220,00	223.92	-3.92
5.	220,00	223.97	-3.97
6.	220,00	223.97	-3.97
7.	220,00	221.63	-1.63
8.	220,00	221.67	-1.67
9	220,00	221.69	-1.69
10.	220,00	219.08	0.92
11.	220,00	219.04	0.96
12.	220,00	219.00	1.00
13.	220,00	219.99	0.01
14.	220,00	220.00	0.00
15.	220,00	219.99	0.01
16.	220,00	220.00	0.00
17.	220,00	220.00	0.00
18.	220,00	220.01	-0.01

IV. CONCLUSION

Electrical power quality has become an important issue and is receiving increasing attention by utility, facility, and consulting engineers in recent years. Power quality problems show up as impacts within the end user facility but may involve interaction between all levels of the system. To understand the power quality problem better requires a comprehensive monitoring and data capturing system that is used to characterize disturbances and power quality variations. Electrical power quality monitoring allows plants to perform predictive maintenance, energy management, cost management, and quality control. The power line monitoring system based on high-performance virtual instrumentation software and additional hardware acquires and displays the voltage and current waveforms. With continuous power quality monitoring, it is very important to be able to summarize variations with time trends and statistics, in addition to characterizing individual events. Virtual instrument described in this paper is developed in a graphical programming language LabVIEW. It provides parameters of voltages, currents and phases, defined by electrical power quality protocols, by means of 8 channels DAQ card PCI NI 6713, comparing generated and measured values of voltage and current parameters, and automatically correcting generated values until the desired accuracy class is achieved.

REFERENCES

[1] Christopher J. Melhorn, Mark F. McGranaghan, ”Interpretation and Analysis of Power Quality Measurements”, Electrotek Concepts, Inc., Tennessee.

[2] Power quality measurement - more important than ever, LEM NORMA GmbH, Austria (www.lem.com).

[3] Building an Electrical Power Quality, Consumption, and Distribution Monitoring System with Virtual Instrumentation, National Corp, USA (www.ni.com).

[4] Standard BS EN 50160:2000 Voltage characteristics of electricity supplied by public distribution systems, Reo, UK (www.reo.co.uk).

[5] Viktor Smiesko and Karol Kovac, ”Virtual Instrumentation and Distributed Measurement Systems”, Journal of Electrical Engineering, vol. 55, no. 1-2, 2004, 50-56.

[6] LabVIEW User Manual, National Corporation, USA, 2000 (www.ni.com).

[7] DAQ 6711/6713/6715 User Manual, Analog Voltage Output Device for PCI/PXI/CompactPCI/PCMIA/1394 Bus Computers, National Corporation, USA, 2003.

[8] Dimitrijević B., Živanović D., Ranđelović I., ”Virtual Instrument Applied to Energy Meters Gauging”, ICEST 2005, Conference proceedings, pp. 390-393, Nis, June 29 - July 1, 2005.

Analysis of the Safety Conditions from High Touch and Step Voltages in the Grounding Systems

Nikolce Acevski[1], Risto Ackovski[2] and Mile Spirovski[3]

Abstract - **In this paper, beside the results of calculation, are described and elaborated mathematic models which are applied for solution of problem with transfer of potential. Appropriate algorithm is made as programmed package which is used for analysis or better says, for anticipation of the conduct of al grounding system (GS) in the area of the mine when faults to ground appeared in the system 110 kV for the present situation.**

Keywords – **groundings, transferred potentials, safety conditions, touch and step voltages.**

I. Introduction

Because of the nearness of the sources of REK Bitola as well as because of the meaning of the mine "Suvodol" problem of transferred potentials is especially expressed. Cables in the network are with rubber isolation that contains three-phase conductors as well as additional signal conductors made by cooper with relative big section. So, all objects in part of the mine (operative stations, diggers, transport tapes, engines and the other consumers on medium (MV) and low voltage (LV)) are mutual galvanic harnessed and together with their groundings, they formed GS of the mine. At appearance of fault to ground at station 110/6 kV/kV of the mine or the network 110 kV in nearness of the mine, that power is distributed at all GS (tower groundings on 110 kV transmission lines (TL), grounding on substation 110/6 kV/kV and groundings by individual MV consumers). For mutual galvanic connection of groundings from supplied objects at 6 kV networks, at appearance of fault to ground in 110 kV network, by metal pieces of equipment may occur potentials appreciably higher of potential by ambient soil. The problem with transfer of potentials also is expressed with the systems for transport of soil and chats, as and substation TS MV/LV by 6 kV network in mine, where groundings of that TS are sources of current field and dangerous voltage of touch and step.

II. Grounding System Model

Netlike Grounding - Model

Small part of the current (just 9,4 %) that is injected in GS of the mine will go to the earth over netlike grounding of TS 110/6kV/kV "Suvodol".

[1] Nikolce Acevski is with the Faculty of Technical Sciences, Bitola, Macedonia, member IEEE
E_mail: nikola.acevski@mofk.gov.mk
[2] Risto Ackovski is with the Faculty of Electrotechnical Sciences, Skopje, Macedonia, member IEEE, E-mail: acko@ieee.org
[3] Mile Spirovski is with the Faculty of Technical Sciences, Bitola, Macedonia, E_mail: mile.spirovski@yahoo.com

Accuracy of the final results of the analysis for the potentials of GS of the mine is little connected with the accuracy of the calculated R_{ts} of the netlike grounding.

For this calculation we will use simplified way the entirely netlike grounding will be modeled with one horizontal panel which equivalent diameter has the same parameter A as the netlike grounding itself.

$$D_{ek} = \sqrt{4 \cdot A / \pi} = 1{,}128 \cdot \sqrt{A} \qquad (1)$$

For calculation of the resistance we will use the following expression [8], where is $\rho = 100\Omega m$:

$$R_{TS} = \frac{\rho}{2 \cdot D_{ek}} = 0{,}725\Omega \qquad (2)$$

Modeling of Transmission Lines

When TL is performance with protective rope, it participates of the transfer of the currents and potentials when fault to ground appeared in the system 110kV. Therefore, in replacement scheme in the network of GS every TL ought to be presented on the same way, with so-called "π-scheme" [8]. In power network of the mine "Suvodol" exist only one TL provided with protective rope. That is TL 110kV TS 110/6 "Suvodol"-TS 400/110 Bitola2", who is consecution like two systematic (with two independent threat on phase conductors with length l=2,7 km, 2x3xA1/C240/40mm²), type of protective rope Fe III 50mm². According to [8], for impedance per kilometer length of protective rope we can get:

$$\underline{z} = \left(0{,}05 + \frac{1000\rho_{Fe}}{S_{Fe}}\right) + j\left(\log\frac{2D_e}{d} + 0{,}0157\mu_r\right) \qquad (3)$$
$$\underline{z} = (4{,}24 + j1{,}24)\Omega / km$$

With D_e is indicated equivalent depth of the return path of current in the earth, which is according to Carson model depend of ρ and frequency f:

$$D_e = 658\sqrt{\rho / f} = 930{,}6m \qquad (4)$$

As we know number of aperture 12, for average a=225 m than for impedance of average aperture we have

$$\underline{Z}_r = a \cdot \underline{z} = (0{,}954 + j0{,}279)\Omega / aperture \qquad (5)$$

TL with more than 10 apertures is treated like infinite TL, without making any important mistake in modeling. TL may be equivalent with it entering impedance:

$$\underline{Z}_{VL} = \sqrt{R_{ST} \cdot \underline{Z}_r} - 0{,}5 \cdot \underline{Z}_r = (2{,}04 + j0{,}22) \qquad (6)$$

Where $R_{st} = 6{,}5\Omega$ is average value of grounding resistance on the separates towers of TL.

Modeling of Cables 6 kV and Characteristics of their Models

Every cable looked together with returned path through earth, can be presented with I-removal scheme, i.e. with one ordinal impedance $\underline{Z}=\underline{z}\cdot l$. Longitudinal impedance that is impedance of unit length $\underline{z}$ will be [8]:

$$\underline{z} = r + jx = \left(\frac{1000}{\kappa \cdot S} + 0{,}05\right) + j\log\frac{D_e}{D_s} \qquad (7)$$

Modeling of Groundings on TS 6/X kV/kV

Each TS 6/x kV/kV which has own grounding with known grounding resistance *R*, in removal scheme of GS will be node, so-called "grounding place", and the grounding itself in removal scheme will be modeled with cross located active resistance *R*.

Accessory Groundings – Model

Accessory groundings of different types of mine objects and machines are modeling of identical way like groundings of substation TS 6/x kV/kV. We can say that it happens that more grounding are galvanic connected in one grounding place. In that case, in removal scheme of GS will appear, as parallel connected active resistance as there are different galvanic connected groundings.

Surface Groundings –Model

According to [9] grounding resistance of transporters/tracks with length *l* and equivalent diameter *d* on the surface of earth at average ρ is:

$$R = \frac{\rho}{\pi \cdot l} \cdot \ln\frac{2l}{d} \qquad (8)$$

According, to this, if the track is located on the area of earth and if it is linked on one end with GS and it is free on the other sight, than it can be treated like elementary grounding, that in equivalent scheme GS on individual grounding place will entered active resistance *R*. As both ends of track are galvanic connected for different grounding places, then in removal scheme of GS the track will have to introduce with π-scheme.

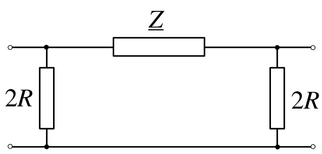


Fig. 1. π–removal scheme for transporters/track

Parameter $\underline{Z}$ of ordinal branch in π– scheme is:

$$\underline{Z}=[\rho_{Fe}\ l/S + 0{,}05] + j\cdot[0{,}1445\cdot\log(2D_e/d) + 0{,}0157\cdot\mu_r] \qquad (9)$$

Analyses shows that step upon areas of diggers, transporters and other equipment (which with their caterpillars realize good power contact with earth), satisfactory can be modeled with individual netlike GS, horizontally placed in earth on certain small depth h. Thereat, netlike GS will need to have the same geometry like geometry of step upon area of digger/machine, and own modeling itself can be successfully done if step upon area is changed with heavy network on horizontal tracks, placed on small distance from one another (e.g. *D*=50 sm.), buried in small depth (e.g. *h*=5 sm).

III. Estimation of Currents and Potentials in Grounding System

When one-phase fault to ground appear in 110 kV patch board of mine "Suvodol" that is of arbitrarily place of TL 110 kV TS 110/6 "Suvodol"-TS 400/110 "Bitola 2", comes to flux of currents per grounding of station, and the potential of grounding, V_z. The analyses show that most unfavorable case, from aspect of the quantity of potential of grounding, is for one-phase fault to ground produced in the station. The whole current of fault doesn't go into the earth, but just one part of it, which will be indicated with I_z. This current in different ways, flows through earth to source of current, i.e. to station 400/110 "Bitola 2". If with $\underline{Z}_z$ we indicate complex "enter impedance of the whole GS at station 110/6 than:

$$V_z = Z_z \cdot I_z \qquad (10)$$

The value of the equivalent impedance Z_z of whole GS needed for estimation of potential V_z may be received in two ways: with direct measuring or with estimation. In this paper it is calculated with computer programmed based on the method of independent voltages.

In the paper it is supposed that current on fault to ground is known, produced in TS 110/6 "Suvodol". Data for value of the current of one-phase fault to ground in power system of R.M. are received from competent offices. The current I_z=17196 A that is injected in GS in mine is calculated like in [6]. After according current I_z is estimated by known (calculate or measured) value of entered impedance $Z_z = Z_{ek}$, calculate the potential V_z=1215 V in netlike grounding, node SO.

After determination of potential V_z on the place of fault, we can determine potentials and for other groundings in the region of the mine. It's very useful, particularly when the calculations are made by computer, if there are used matrix methods. In the algorithm according to which is produced the programmed package Suvodol, is used matrix approach for solution of the problem of distributing of currents and transfer of potentials in GS in the mine. Solution of mentioned problem is done with help of matrix of impedances, [*Z*], of GS [11]. For this purpose, primary, according to known plot of feeder 6kV and their known parameters, is generated so-called matrix of admittances [*Y*] of GS, [11]. With its inversion [*Z*]:

$$[Z] = [Y]^{-1} \qquad (11)$$

Potentials $\underline{V}(i=1,N)$, of the separate groundings are:

$$V_i = Z_{is}\cdot I_Z\ ;\ i = 1,2, ..,N \qquad (12)$$

With $Z_{is}\,(i=1,N)$ are elements in the column "*s*" of matrix [*Z*], hereby, with "*s*" is indicated the index of node that is referred to grounding.

IV. RESULTS OF THE ESTIMATION AND MEASURES FOR ELIMINATING DANGERS

Whole network 6 kV is divided, on:

1. Coal system (SJ);
2. First system (S1);
3. Second system (S2) and
4. Zero system (S0).

Table 1 Distribution of injected currents and equivalent impedances per separated systems of the mine for total injected current I_z =1000A

	S i s t e m	I_r	I_x	I	φ		Z
/	/	(A)	(A)	(A)	(o)	%	Ω
1	Coal (SJ)	277.62	-19.71	278.32	-4.06	26.7	0.25376
2	First (S1)	217.72	-50.86	223.58	-13.15	21.5	0.31589
3	Second (S2)	151.09	-16.38	151.98	-6.19	14.6	0.46472
4	Zero (S0)	255.55	-1.06	255.55	-0.24	24.5	0.27637
5	Netlike grounding	70.64	67.11	97.44	43.53	9.4	0.72481
6	110 kV TL	27.37	20.89	34.43	37.35	3.3	2.05105
	Total	1000.00	0.00	1000.00	0.00	100.0	0.07063

Table 2. Nodes in network 6 kV in which $\Delta E_d > 200$ V

	System	node	Grounding type	ρ (Ωm)	U (V)	ΔE_c (V)	ΔE_d (V)	ΔU_d (V)
1	(SJ)	RP6	RP16	40	770	139	201	189
2	(SJ)	RP78	RP78	40	908	127	296	279
3	(SJ)	DOM.TRAFO	RP16	100	1168	210	305	264
4	(S2)	TS4	RP16	100	1184	213	309	267
5	(S2)	RP5	RP16	50	778	140	203	188
6	(S2)	ES10	BG_JL3	40	558	275	296	279
7	(S0)	RP910	RP78	40	903	126	294	277
8	(S0)	PRP	PRP	100	814	197	385	333

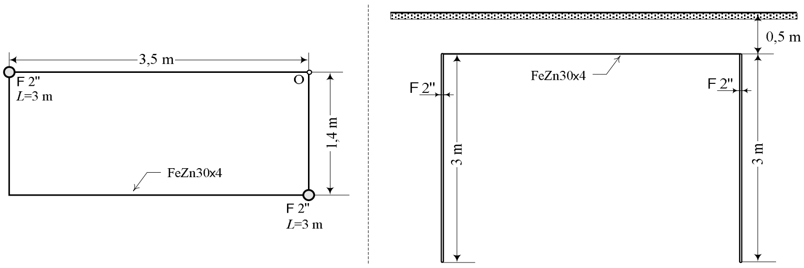


Fig. 2. Grounding type PRP

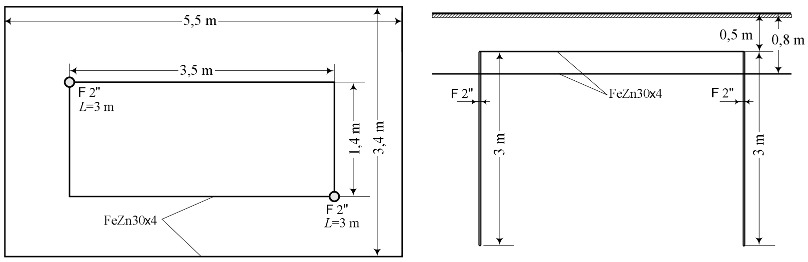


Fig. 3. New grounding type PRP 2

That is made, so we can see what is the stake of each of the systems of the mine in drop off of total current I_z although it's partial GS are galvanic connected in the central separation (node "SO") through it's grounding type "CENTRAL", and they are treated like that in the estimations as well. The results are shown in table 1. When we evaluate the attained results, if estimated voltages of touch and step are in the permitted borders, we need to respect time of duration of one-phase fault to ground and allowed voltages of touch and step.

According to the results from the studies [1], [2], time of disconnection of fault to ground made in the station 110kV is t=0,1 s. In the European and American standard (ANSI/IEEE Std80-1986) the lowest values are when person is in contact with metal areas, and voltage of touch for man with 70 kg is:

$$\Delta U_{ddoz} = \Delta U_{cdoz} = 157 / \sqrt{t} \qquad (13)$$

According to [5] if the time of disconnection is t=0,1s, allowed voltage is 300 V. From this point we can make the conclusion that if time of disconnection of fault to ground produced in 110 kV network is 0,1 s, allowed voltage of touch/step, is 300 V. From the received results, we can see that maximum voltages of steps are regular for 40-60 % smaller than maximum voltages of touch, so we can conclude that dangers in 6 kV network in mine when fault to ground appear in 110 kV system are voltages of touch. In table 2 are present nodes of network that have maximum potential difference of touch that exceeds the value of 200V. In it there are shown voltages of nodes U(V), ρ (Ωm) as well as the voltages of touch ΔU_d and potential difference of touch ΔE_d. But if the measure for danger is ($\Delta U_{d.doz} = \Delta U_{\sim.doz} = 300$ V), then in 6 kV network in the mine danger of more higher voltage of touch will appear one in the knot PRP in zero system ($\Delta U_d = 333$ V). Solution of this problem can be done in more than one way but the simplest are the following:

1. Installation of additional ring around the object, 1m of its margins and at depth of 0,8 m.
2. Asphalting 1 m path around the object with thickness 5sm.

Asphalting will occur if there isn't objective limitation for installation of the additional ring around the object itself. For that reason permanent grounding type PRP (figure 2) that is consisted of just one rectangular ring with sizes axb=3,5x1,4m buried in depth $h_1 = 0,5$ m and two vertical plug F2" that are 3 m long, we will add another rectangular ring with dimensions a2xb2=5,5x3,4m, buried in depth $h_2 = 0,8$ m and this way we will have new grounding that will be called type PRP2 (fig. 3). Characteristics of this new type of grounding are explored in ZAZEM and it is concluded that the biggest potential differences of touch and step are by the length, direction of 0-A.

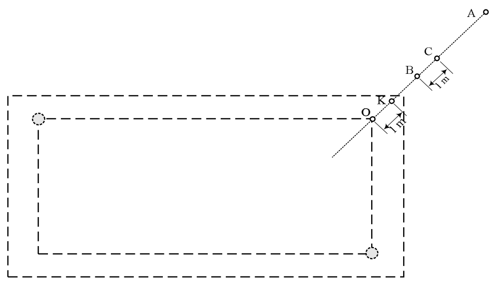


Fig. 4. Estimation of potentials in the surface of earth

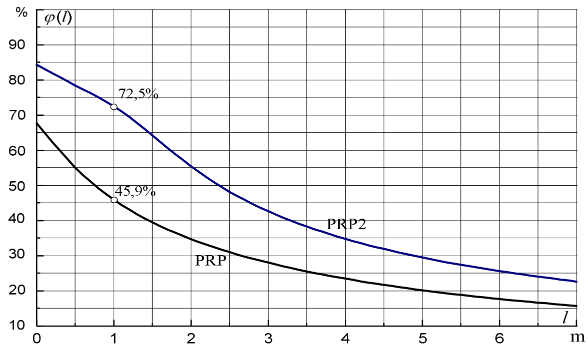


Fig. 5. Distribution of potentials per length of critical direction

On the fig. 5. φ_K ($\varphi_K = \varphi_{min}$) in the critical item k, that is located 1m from the edge of object, and the lowest when the potential difference of touch ΔE_d is difference between potential φ_0 of the object and φ_K, in item k:

$$\Delta E_d = \varphi_0 - \varphi_K \qquad (14)$$

So, after putting second, additional, contour, we obtain new, grounding type PRP2, at which maximum voltage of touch will be reduced of the value $\Delta U_d = 169$ V. Besides these analyses, there are calculated potentials differences of step ΔEc. Calculations show that the biggest potential difference of step is again in direction 0-A, between the points B and C which are 1m from one to another. Essential characteristics of groundings type PRP and PRP2 are shown in table 3.

Table 3. Characteristics of groundings type PRP and PRP2

Grounding type	R_Z, (Ω)	ΔE_c (%)	ΔE_d (%)
PRP	11,35	21,7	54,1
PRP2	8,05	17,2	27,5

V. CONCLUSION

1. The biggest danger of transferred potentials in 110kV network in region of mine "Suvodol" at REK –Bitola is when fault to ground appear in TS 110/6kV/kV.
2. Time of disconnection for the current that is produced is estimated $t = 0,1$s. For this time is suggested that as a criteria of danger should be accepted $\Delta U_{d.doz} = \Delta U_{c.doz} = 300$ V.
3. For this adopted criteria for safety is obtained that real danger of higher voltage of touch is noticed only with one grounding, in the node PRP ($\Delta U_d = 333$ V).
4. It is suggested the grounding in this TS 6/0,4 kV to be improved with adding one rectangular contour. Alternative, whole zone around the object should be asphalted with width of 1m and thickness 5 sm.
5. Currents of fault to ground after building of 400kV relation Bitola2-Amindeo, will be bigger at least for 40-60%. That will provoke enlargement of dangers from transferred potentials in MV network of the mine for approximately the same % and appearance of new dangerous places in MV network, table 2.

REFERENCES

[1] R. A~kovski, D. Raji~i}, R. Taleski i A. Dimitrovski, "Modelirawe na kusite vrski, zazemjuva~kiot sistem i presmetka na potencijalite na zazemjuva~ite vo reonot na rudnikot 'SUVODOL' pri pojava na kusa vrska vo sistemot 110 kV", ETF - Skopje, mart 1991.

[2] R. A~kovski, D. Raji~i}, A. Dimitrovski i R. Taleski, "Analiza na karakteristikite na tipskite zazemjuva~i vo reonot na 'Suvodol'", Skopje, juni 1991.

[3] D.Vidanovski, M. Stojanovski B. Dungovski, "Analiza za spre~uvawe na izvoz na opasni potencijali po transporterite vo Rudnik Suvodol". 1993 g.

[4] R. A~kovski, R. Taleski, A. Dimitrovski: Stru~na recenzija na elaboratot "Analiza za spre~uvawe na izvoz na opasni potencijali po transporterite vo Rudnik Suvodol", ETF-Skopje, 1994 g.

[5] Pravilnik za tehni~kite normativi na elektri~ni postrojki i uredi vo rudnicite so povr{inska eksploatacija na mineralni surovnini. "Sl. list na SFRJ br. 66/87".

[6] N. Acevski, "Prilog kon modelite za re{avawe i analiza na zazemjuva~i i zazemjuva~ki sistemi", doktorska disertacija ETF-Skopje, 2003.

[7] "M. Pijade"-Svetozarevo. Katalog za SVERPEN kabli 6 kV.

[8] J. M. Nahman: "Uzemljenje neutralne ta~ke distributivnih mre`a" Nau~na knjiga Beograd, 1980.

[9] A. Huki}: "Proraun iznoenja potencijala iz TS 110/35/6 HAK putem metalnog pla{ta 6 kV kablova. ^asopis "Elektrotehnika" ELTHB2 29, br. 1-2/1986.

[10] D. Salamon: "Analiza tipi~nih uzemljiva~a na povr{inskom kopu rudnika uglja REIK "Kolubara". "Elektrotehnika" ELTHB2 27, br. 1-2/1984.

[11] H. E. Brown: Solution of Large Networks by Matrix Methods, John Wiley&Sons, Inc. N.Y. 1975.

Power Consumption Control System in Copper Tubes Factory - Majdanpek

Vladimir Despotović[1] ,Viša Tasić[2], Dragan R.Milivojević[3], Aleksandra Ignjatović[4]

Abstract **–This paper shortly describes realization of the computer network, and SCADA software used for monitoring and peak power and energy distribution control in transformer station at Copper Tubes Factory (FBC) in Majdanpek. Some specific details of hardware and software (own made) implementation were specially noted.**

Keyword **– monitoring system, peak power, energy**

I. Introduction

Department of Industrial Informatics, at the Copper Institute in Bor has been designing real-time systems for monitoring and control of industrial processes since 1990. As a result, three generations of UMS (Universal Measuring Station) have been developed. UMS is industrial PLC (Programmable Logical Controller) fully designed at the Department of Industrial Informatics and it is a core of process control system. Main objectives of such systems are real-time data processing and displaying of results on dynamic synoptical schemes, real-time graphs or tables, database forming, and automatic process control.

The unified SCADA software [3] has been modified and used for monitoring of power consumption and peak power control at Copper Tubes Plant (FBC) in Majdanpek.

II. System configuration

The main reason for introduction of PC based control system at the plant transformer station is the reduction of the peak power, which is directly related to lower expences of electrical energy. The monitoring project predicts measuring of active and reactive power on each input and output cell at the plant transformer station, and automatic control of main power consumers (switching based on given criteria).

As the power consumers are dislocated, all signals, (analog signals from power transducers, and digital control signals) had to be concentrated at one place, were measuring station is placed.

All analog signals are inputted to UMS as standard current signals (0-20 mA). Monitoring computer (PC) is placed at plant transformer station control room. Monitoring workstation (standard PC) communicates with UMS over serial communication port (RS-232 interface connector), and the physical link is full-duplex line. To overcome the distance between PC and UMS, Line Amplifyers are used. The standard bit rate of data transfer is 9600 bps. Communication protocol is ASP (Asynchronous Serial Protocol) [1]. Monitoring computer is connected over Local Area Network to other PC's and acts as a data server. Server software, which drives communication with UMS, also has an interactive role and lets the user change parameters of the monitoring process, like sampling rate, number of measuring channels, parameter limits, type of data archiving etc. However, client software has only the ability to monitor process, without any interactivity. Diagram of realized monitoring system is given in Fig. 1.

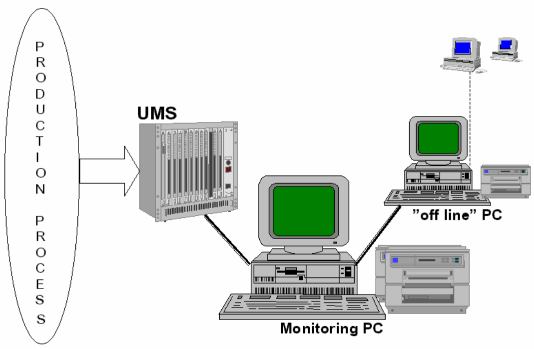


Fig. 1. Overall system configuration diagram

III. Hardware

UMS is a device for data acquisition based on Motorola 68HC11 microprocessor. Main characteristics of UMS (standard configuration) are:

- microcontroller Motorola 68HC11E
- intern eight channel, 8-bit A/D converter (conversion time less than 40 µs),
- 64 differential analog inputs,
- 64 + 64 digital state signals (input + output) with mutual point (or independent). All digital signals are galvanicly isolated.
- RS232 communication port,
- 48 (56) KB for data (RAM)
- 16 (8) KB for software (EPROM)

Local display and functional keyboard gives a possibility of device control, correction of the time, and start of measuring. UMS can work independent of PC, and can control the

[1] Vladimir Despotović is with the Copper Institute Bor, Zeleni bulevar 35, 19210 Bor, Serbia and Montenegro, e-mail: vlajcheck@ptt.yu,

[2] Viša Tasić is with the Copper Institute Bor, Zeleni bulevar 35, 19210 Bor, Serbia and Montenegro, e-mail: visa@ibb-bor.co.yu,

[3] Dragan R. Milivojević is with the Copper Institute Bor, Zeleni bulevar 35, 19210 Bor, Serbia and Montenegro, e-mail: dane@ibb-bor.co.yu,

[4] Aleksandra Ignjatović is with the EMS Beograd, Serbia and Montenegro, e-mail: a.ignjatovic@ems.co.yu

process itself. It can also work like data logger, and memorize about 2000 data messages in local RAM, and later, when connection with a PC is established, transfer them to PC. Minimal demands for monitoring computer are: Windows 98 OS, Intel Pentium 3 processor, at 450 MHz or higher, 128 MB RAM, SVGA graphics (800x600).

IV. SOFTWARE

EPROM of UMS is consist of executable versions of test, control, operational and communication software modules. These modules are optimized in space and time and they are global solution for standard configuration. Operational program module is responsible for measuring of analog channels and checking states of digital inputs. Type of measuring, sampling rate and other parameters can be changed using local keyboard, or commanded from a monitoring PC (more often case). The results of the measuring form the message, which is transferred to monitoring PC, or memorized in local RAM (if PC is disconnected), so it can be transferred later when the connection is established. UMS may work independently from monitoring computer, so local process control is also possible. If any parameter exceeds given limits, it causes alarm message, or even better, if any parameter shows trend of reaching limit value, it firstly can cause warning message, which is more effective, so the operator, or the system itself can react on time.

Appropriate process control application is installed on monitoring computer. This real-time based application is developed using Microsoft Visual C++ 6.0 [2,3,4], and it's main characteristics are: communication with UMS, data processing, data presentation, command and process control (interactive), data archiving and off-line analysis and interpretation of data.

Most important class of this software contains driver for communications with UMS as secondary network node (PC is a master). Communication protocol is ASP (Asynchronous Serial Protocol) [1]. Main objectives of this class are:

- request for communications
- connection establishment
- data transfer
- control of transfer quality
- state indication of network nodes

Member functions of this class control main parameters of UMS: sampling rate, type of measuring, number of analogue and digital channels, correction of time on UMS based on PC time etc.

Other classes are used for displaying of data on dynamic synoptical schemes, real-time graphs or in tables. They analyze and process data and display them in adequate manner. Fig. 2. shows the look of one synoptic screen and Fig. 3. shows measuring results presented in tabular form.

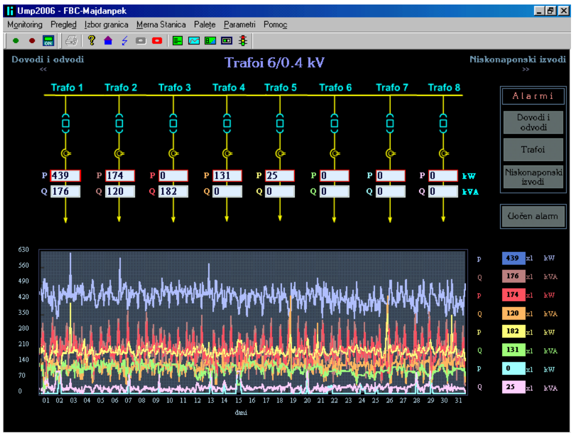


Fig. 2. Synoptic screen

Data are memorized in different buffers, until the moment of archiving. Further processing is consist of number of actions:

- data processing for graphic display on the screen,
- data processing for alert system,
- data processing for distribution over LAN,
- forming of daily archives and
- forming of monthly archives.

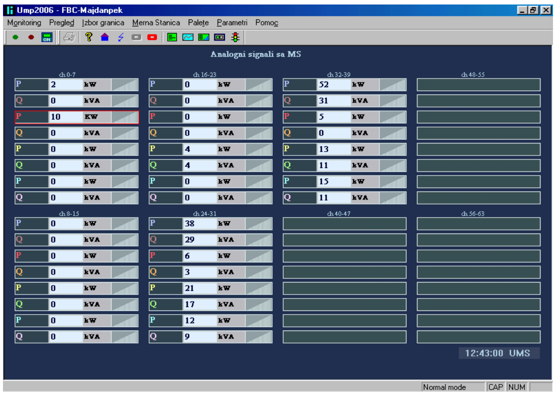


Fig. 3 Measuring results in tabular form

In order to display data in the user demanded way, independent of channel's order in a message, channel switch class has been created. Other classes enable additional facilities to the users like: changing the range of measuring, changing the alarm limits, scaling the axes at the real-time diagram etc. Software is user-friendly, and all this can be done easily using dialog boxes.

File management class includes functions for data archiving. Buffered messages are written on hard disk based on time criterion defined by user. Data are saved as text files and can be used for further processing (Excel, Access). Three groups of files are formed: daily reports, monthly reports and log files. Log files are used to display the history of a controlled process. Those files contain informations about every user intervention (operator) with the exact time and date, time of alerts etc., which can be very useful for later analyses.

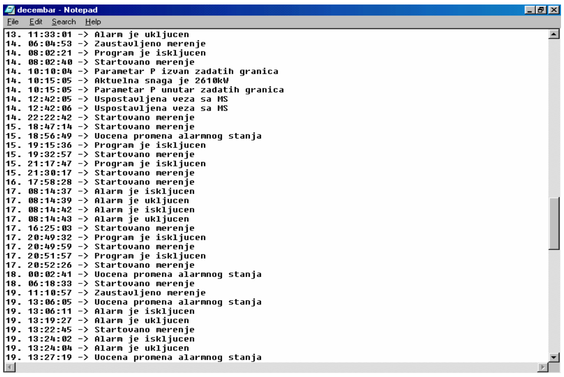


Fig. 4. Part of a log file

In order to get better performance, user can change process priority, comparing to other active applications on PC, from low to real-time priority. In high or real-time mode application performed very stable. Some of the functions are password protected.

V. OBJECTIVES AND PERFORMANCES

Main objective of the system is control of power consumption, intending to make significant financial effects by peak power reduction. System has the ability to predict trends of active power consumption based on 15-minute average values. User chooses the limits, and also the starting minute for power control (from 7^{th} to 14^{th}). From that moment, average power values are compared to given limits every minute, and up to eight consumers can be automatically turned off if the limit is exceeded, in order to reduce the power consumption.

As monitoring application is working in the multitasking OS, it is recommended to choose high or real-time priority to obtain stability. Communication between PC and UMS is based on master-slave principle, where PC is a master and requests communication, and UMS replays with the information about its state, quality of transfer and measuring data when they are present. PC gives a command to start measuring on UMS, transferring all parameters needed for its independent work, including real time from PC. UMS continuously measures defined parameters in a rhythm given by the user.

Monitoring PC receives the message, analyzes, displays and saves data. It also alerts if any parameter exceeded limits, both sound and visually. Operator must react on alert states, and exact time of alarm and his reaction are written in log file.

Each computer in Local Area Network (under Windows OS) can monitor the process using client application. This (client) version [2] of the software has only the ability to monitor process, without any interactivity.

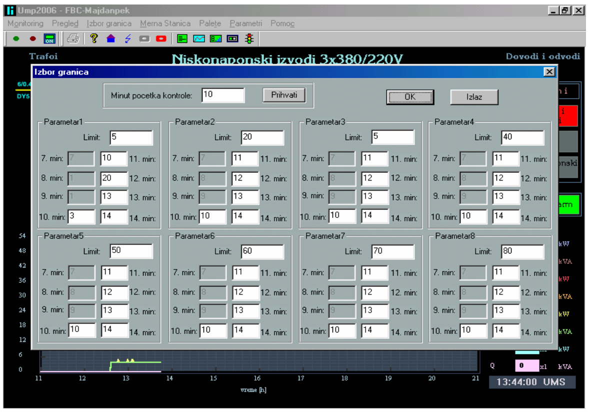


Fig. 5. Dialog box for changing parameters of process control (limits and starting minute of control)

VI. CONCLUSION

Electrical power consumers in FBC can be divided in 3 categories:

- High priority cosumers which must not be switched off by the process control system,
- Middle priority consumers - system alerts the operator both sound and visually about possible overload of peak power. Depending on operators decision, consumer should be switched off, or switching on of some other consumer should be postponed,
- Low priority consumers are automatically switched off if system predicts possible overload of peak power, without any bad effects on technological process.

Although system has been installed only few months ago, the financial effects are already visible. The effects are both technological and psychological. Technological effects are peak power reduction and rational use of electric energy. Psychological effects are reflected in better discipline at work due to operator's feeling of permanent and objective control of their behavior.

REFERENCES

[1]. D.Milivojevic, V.Tasic, D.Karabasevic, *Communication in realized real-time systems, ETRAN 2003, Conference Proceedings CD,* Herceg Novi, Serbia and Montenegro, 2003.

[2]. D.Karabasevic, D.Milivojevic, V.Tasic, *Universal measuring program, YUINFO 2004, Conference Proceedings CD,* Kopaonik, Serbia and Montenegro, 2004.

[3]. D.Milivojevic, B.Lazic and V.Tasic, "*Applied process control systems*", Journal Bakar, No. 1, pp.125-134, Copper Institute Bor, 2000.

[4]. D. Milicev, *Object oriented programming in C++*, Mikro Knjiga, Belgrade, 1995.

An Example of Influence on Reduction of Electrical Energy Costs

Dragan R.Milivojević[1], Viša Tasić[2], Vladimir Despotović[3], Aleksandra Ignjatović[4]

Abstract –**This paper shortly describes realized computer network used for monitoring of electrical power consumption in consuming transformer station at the copper mine and copper smelting plants in Bor and Majdanpek. Some effects on reduction of costs by control of peak power were specially noted.**

Keyword – **electrical power consumption, process control, peak power, monitoring system**

I. Introduction

Copper Mining and Smelting Complex Bor (RTB) is certainly among greatest electrical energy consumers in Serbia and Montenegro. About 80 thousands MWh of active electrical energy are consumed monthly in its production sites (mines, flotation plants, smelters and refineries and copper treatment in Bor and Majdanpek), what is about 1 000 000 US$, if the costs of maximum power (peak) are considered.

Certainly, it is not required to state the other reasons for permanent or periodical measures for reduction of mentioned costs.

Beside often normative and organizational actions, the decision was made to begin with the realization of a real time based monitoring system for control of electroenergetic system in RTB Bor. The known performers from the production sites of RTB Bor and Copper Institute were engaged in this work. Strong connection and cooperation of two basic groups of experts: technologists (energetic and process engineers) and automatic engineers (primary, experts for informatics) was required for successful realization of this complex project.

The pilot operation of the system has started from the middle 1990's in Bor's part of Complex, and later an identical system has also been developed at the Copper Mine Majdanpek (RBM).

II. Characteristics of a Control System

Due to a great complexity of electroenergetic network in RTB, a large number of transformer stations and their spatial dislocation, the first project stage had to include only the consuming transformer stations.

[1] Dragan R. Milivojević is with the Copper Institute Bor, Zeleni bulevar 35, 19210 Bor, Serbia and Montenegro, e-mail: dane@ibb-rtb.bor.co.yu

[2] Visa Tasić is with the Copper Institute Bor, Zeleni bulevar 35, 19210 Bor, Serbia and Montenegro, e-mail: visa@ibb-rtb.bor.co.yu,

[3] Vladimir Despotović is with the Copper Institute Bor, Zeleni bulevar 35, 19210 Bor, Serbia and Montenegro, e-mail: vlajcheck@ptt.yu,

[4] Aleksandra Ignjatović is with the EMS Beograd, Serbia and Montenegro, e-mail: a.ignjatovic@ems.co.yu

The preparation of transformer stations was carried out in such way that the measuring power converters, connected as it is shown in Fig. 1. were installed into the measuring circuit of secondary cells of all transformers (supply and for the all subdivide ones). Outputs from the measuring power converters are standard current signals (4 - 20 mA) proportional to the active and reactive power of given transformer cell. Transmitters (measuring converters) of active and reactive power, are the result of our own development, and their basic characteristics are: Aron's measuring methods, input voltage is 3x100 V symmetrically, output is 4 - 20 mA, unlinearity < 1% and overloading up to 20%.

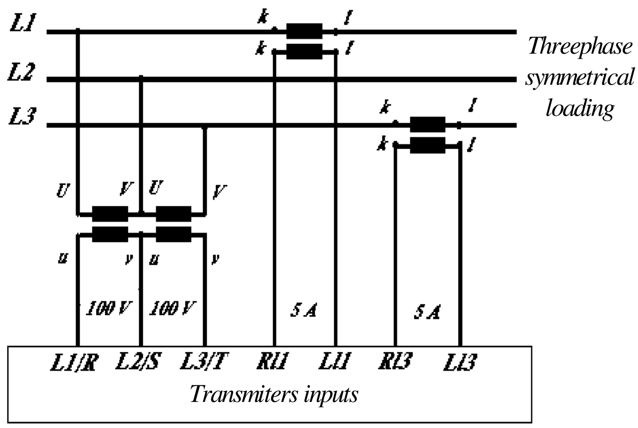


Fig. 1. Way of connection the P&Q transmitters

Signals from the P&Q transmitters are supplied to the inputs of universal measuring station (UMS), which carries out measuring (sampling). Measuring method includes A/D conversion of inputs and processing of the obtained numerical values. It was experimentally estimated that measuring at 200 ms gives a real view on changes of involved power. More frequent sampling does not give noticeable better results due to the system inertia. 8-bit A/D conversion is used and it is expected that conversion error is less than 1 LSB, i.e. lower than 0.5%.

Data processing starts with logic control. Possible saturation at inlet is checked disregarding a fact that it is a result of measuring converter error, or bad signal condition at inlet into measuring station. Negative value of active power also causes the error message. Since the measuring is carried out five times per second, then the average value of power represents the minute power. Every minute it is transferred to the PC workstation (at the control center). Since the 15-minute average power is calculated for the electrical power rate and payment, measuring stations calculate this power as a cumulative divided by the number of minutes in a range of given 15 minutes (depending on number of messages received). Such measuring method gives a possibility for prediction of overload of given (or maximum) power limit for any plant and each transformer in the transformer station. It is worth mentioning that the measuring station checks all conditions (switches and ground protection) at every 100 ms. The condition signalization is not done in some transformer stations, so this possibility is not useful, but in principle it is possible to have the information on condition change with time resolution of 1/10 seconds.

III. Operation method

All consumer transformer stations (two in Bor and three in Majdanpek) are measured. Regarding to their dislocation and distance from control center, it was necessary to realize the computer network that includes the measuring devices in transformer stations and personal computer in control centre. Fig.2. presents the hardware network structure in the Copper Mine Majdanpek (RBM).

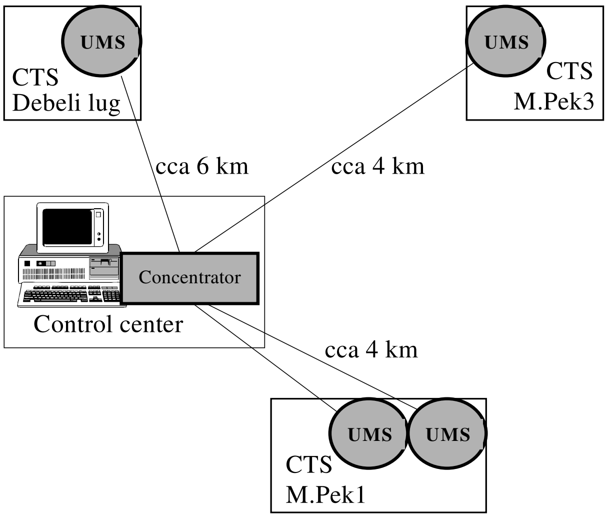


Fig. 2. Disposition of consumer TS (CTS) and UMS in RBM

Hierarchic type of computer network is a starlike structure with central control [5]. The main unit in network is the concentrator [1], and secondary nodes are PC and measuring stations. Transmission lines are leased telephone lines, and the base band modems are used for the synchronous transmission. The operation of all devices and network functions are carried out under the control of suitable software.

Executive versions of program modules are included at the measuring station's EPROM: test-control, operational and communication program module. The complex communication program enables concentrator to configure and maintain network in the operational regime.

PC (expanded with hardware communication module) in control center carries out the application program for real time monitoring of the electroenergetic system. This program [2-4], includes three basic modules: communication module for data receiving from measuring stations, operational module for real time data processing and their presentation on the screen, and module for offline data processing and archiving. Interaction with the process is obtained under control of operational program, and it is reflected in forced warning messages on a screen. Those messages inform on limit or maximum power overload (peak) at each transformer station, or some technological or organizational part of the company, regarding to a possibility of prediction of the involved power in a range of 15 minutes. As a reaction on warning message, some consumers could be switched off or switched on, if it is technically possible and economically justified.

After a period of pilot operation, realized system has turned into exploitation stage. It has shown out the satisfied stability and acceptable reliability from a technical point of view. Organizational side has not completely followed the technical possibilities of installed system. The resistance was understandable because the "objective" costs limitation would disturb the previous agreeable distribution, where unrational consumption of some plants was overturned to the rest of company complex. Besides all, at the beginning of its operation, the system had positive effects, because the greater consumers, aggregates with power greater than 300 KW were switched on whenever it was possible, with the agreement of control center (they needed some kind of permission).

The positive effects were both technological and psychological. Technological influence is realized in change of work organization and sufficiently rapid monitoring of involved power, with a possibility of influence on its reduction (switching off or blocking of switching on, i.e. warning on exceeding) whenever it is possible. Psychological influence is reflected in increase of technological discipline and responsibility of performer, due to knowledge of presence of objective measuring and supervision.

IV. Results review

It was already mentioned that the involved power and energy are monitored at the plant level. Program solution for analysis and data review for every plant and organizational part was developed. The greatest system contribution in costs optimization is an influence on decrease of maximum power. Since there is sufficient number of data for plant's consumption, it is possible to determine the limit power and give it as a parameter to the system, so that it could warn on approximation to this value, or alarm its exceeding. Program solution allows two levels of power limit and it could be dynamically changed.

The serious team work of technologists and energetic engineers, and experiments from the field of finances are required for realization of universal and qualified analysis of behavior of the electroenergetic system in RTB Bor and Majdanpek. Discussion of those results could not give a real illustration on rational consumption if production is not taken into consideration, i.e. if energetic condition is not calculated (quantity of consumed energy per product unit: tone of cathode copper), and if this data are not compared to equivalent world parameters. Report on monthly consumption and maximum power for consumer transformer station Bor 3 for 2004 and part of 2005 will be used for a superficial analysis of consumption. A part of standard report which also includes a limitation of consumed electric energy (active and reactive) per pay scales and shifts, including the financial indexes (price of active energy, reactive energy and maximum power), for RBM is presented in Table I.

	E_a (MWh)	P_{amax} (kWh)	E_a/P_{max}
2004			
January	26298	39160	0.6715
February	23987	37420	0.6410
March	27300	38620	0.7068
April	20400	36450	0.5596
May	21570	34720	0.6212
June	22620	38120	0.5933
July	31640	32800	0.6597
August	21460	32000	0.6706
September	21520	31230	0.6890
October	22507	38820	0.5797
November	22856	39380	0.6803
December	24965	39950	0.6249
2005			
January	24009	38.23	0.6280
February	21749	39.79	0.5466
March	21387	37.41	0.5717

Table I. Consumed active electric energy in consumer transformer station Bor3 per months in 2004. and 2005. year.

Due to a complexity of energetic network, some plants are supplied simultaneously from many transformer stations (Floating pump station) or they could be alternatively supplied from various sides. It is necessary to treat correctly every energetic configuration in program solution for analyzing of measured data in real time and data processing and preparation. The expert support of main energetic engineer and suitable assistants in the plant is required in this work. Data on time origin of maximum power for the whole region and also for all consumer transformer stations in RTB are reported by the energy supplier in the first half of month for the previous month. Regarding the existence of data on 15 minutes power, a measured results with given time are simply extracted from database, and a report on involved power per transformer stations and plants at the time of regional maximum is made. If it is known that this participation of maximum power (peak) in total costs is important (near 50%), the costs restriction based on this could force the plant's employees to be more careful in a process of greater consumers involvement.

By superficial analysis of Table II, it is noticed that monthly consumption of active energy is nearly the same, and a distribution according to the rate items is completely uniform and maintained.

A part of standard report is given in Fig. 3. as an illustration of participation in maximum power of RTB plants, supplied from consumer transformer station Bor 3 in February 2004.

Month	**Rate**	**E_a (MWh)**	**E_r (MVAh)**	**P (MW)**
February	High Low Total	5122 4999 10121	1153 1032 2185	Min. 7.61 Max. 20.47 Aver. 15.06
March	High Low Total	5955 5645 11600	1629 1255 2884	Min. 7.80 Max. 20.11 Aver. 15.60
April	High Low Total	5642 5465 11107	1628 1294 2922	Min. 6.98 Max. 20.65 Aver. 15.42
May	High Low Total	5057 5029 10086	1610 1501 3111	Min. 1.54 Max. 21.95 Aver. 13.56
June	High Low Total	5680 5625 11305	1393 1177 2570	Min. 3.68 Max. 20.99 Aver. 15.70
July	High Low Total	5296 5249 10545	1783 1619 3402	Min. 2.88 Max. 20.59 Aver. 14.65

Table II. Part of the report on consumed electric energy and power.

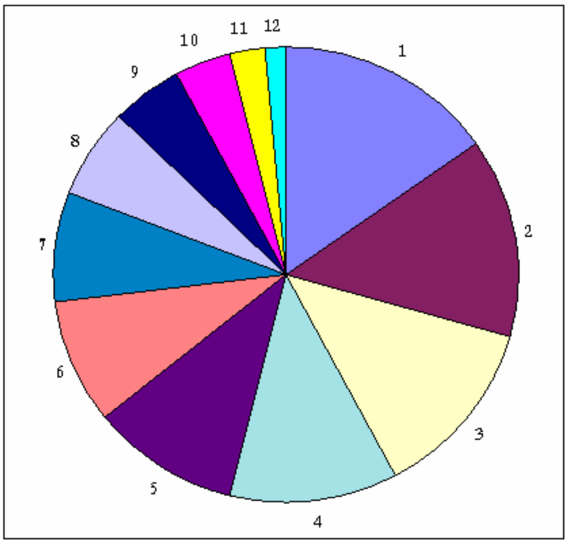


Fig. 3. Distribution of maximum power (peak) Legend: 1. New Flotation Plant (13 %); 2. Old Flotation Plant (9 %); 3. Haulage Shaft (4 %); 4. Contact (8 %); 5. Thermoelectric Power Plant (3 %); 6. Heating Plant (6 %); 7. Oxygen Plant (6 %); 8. Foundry Plant (5 %); 9. Sulphuric Acid Plant (0 %); 10. Electrolytic Refinery (12 %); 11. T100 (11 %); 12. T456 (23 %).

V. CONCLUSION

The process control system of electroenergetic system in RTB has justified its development and financial investments. As far as there is economical force for saving, it is possible to introduce an obligation of cooperation between all parts of plant and control centre in order to obtain more effective power consumption control and monitoring.

REFERENCES

[1] D.Milivojevic, V.Tasic, D.Karabasevic, *Communication in realized real-time systems, ETRAN 2003, Conference Proceedings CD,* Herceg Novi, Serbia and Montenegro, 2003.

[2] D.Karabasevic, D.Milivojevic, V.Tasic, *Universal measuring program, YUINFO 2004, Conference Proceedings CD,* Kopaonik, Serbia and Montenegro, 2004.

[3] K.H.Kim, *Object Structures for Real-Time Systems and Simulators,* Computer, pp. 62-70, California USA, August 1997.

[4] D. Milicev, *Object oriented programming in C++*, Mikro Knjiga, Belgrade, Serbia and Montenegro, 1995.

[5] A.S.Tanenbaum, *Computer Networks*, Prentice-Hall, International, Inc, 1989.

29 June – 1 July 2006, Sofia, Bulgaria

Effects of Measures for Energy Loss Reduction on Suburban Distribution Networks

Dragan S. Tasić[1] and Miodrag S. Stojanović[2]

Abstract – **The most often used measures for energy loss reduction in distribution networks are discussed in this paper: reconfiguration, voltage regulation, reactive power compensation, and conductors changing. Theoretical bases of the measures are given at the first. After that, effects of application of these measures on the real suburban distribution network are analyzed. It is shown that all measures should be analyzed simultaneously.**

Keywords – **Electric energy, losses, distribution network, reconfiguration, compensation.**

I. INTRODUCTION

Measures for energy loss reduction have received an ever increasing attention from the electric utility industry in recent years. However, unique recommendations for distribution utilities what measures should be applied still do not exist. This is consequence of different conditions in distribution utilities. Nevertheless, results obtained on one representative distribution network can be used on networks with similar configuration and similar profile of consumers.

Measures for energy loss reduction in distribution networks can be classified into two groups: organizational, and technical. Organizational measures are ones which do not require additional investments. These are: optimal load distribution, optimization of separation points of double feeding primary lines, voltage regulation, alignment of phase loadings in low voltage networks, and disconnection of distribution transformers in low load regimes in substations with two or more transformers. These measures we can realize in a short time and do not require any investment.

Technical measures are: reconstruction of network, application of capacitor banks, replacement of conductors or transformers, and automation of networks. Technical measures, on the difference of operational ones, require investments and have the long applying time.

Beside mentioned measures distribution utilities should perform measures for accurate reading of spend energy. These measures are periodical checking of accuracy of electric meters, and replacement of old electric meters with new ones which have better technical performances.

Efficiency of one measure for energy loss reduction can be estimated based on the effects of applying of this measure. For operational measures the effect is shown only by amount of

[1]Dragan S. Tasić is with the Faculty of Electronic Engineering, University of Niš, Beogradska 14, 18000 Niš, Yugoslavia, E-mail: dtasic@elfak.ni.ac.yu

[2]Miodrag S. Stojanović is with the Faculty of Electronic Engineering, University of Niš, Beogradska 14, 18000 Niš, Yugoslavia, E-mail: miodrag@elfak.ni.ac.yu

Efficiency of one measure for energy loss reduction can be estimated based on the effects of applying of this measure. For operational measures the effect is shown only by amount of energy loss reduction. When technical measures are analyzed, investment expenditures must be considered. In [1,2] efficiency of technical measures is evaluated by time of returning of the investments.

Analyses of effect that some of measures for energy loss reduction have on real medium voltage suburban networks is presented in this paper.

II. MEASURES FOR ENERGY LOSS REDUCTION

Reconfiguration

Reconfiguration of distribution network [3-6] is defined as changing of network topology by switching tie breakers and sectionalizing switches. Methods for solving of reconfiguration problem we can classify into three groups:

- methods based on optimization techniques,
- heuristic methods,
- methods based on artificial intelligence techniques.

The problem of finding the network configuration with minimum line losses, a mixed-integer, non-linear optimization problem, has been solved using branch-and-bound method. There is, however, no assurance of convergence when branch-and-bound is used for this type of optimization problem. Additionally, sometimes all relevant constraints can not be comprehended. Therefore, actual approaches for reconfiguration are based upon heuristic or artificial intelligence techniques (expert systems, fuzzy logic, and neuron networks). Heuristic methods are finding quested configuration following physics of problem. Two main approaches of these methods are:

- method of individual switch opening,
- switch exchange method.

The first heuristic approach is very efficient and determines the network configuration with minimum or near-minimum line losses based on special variation of branch-and-bound techniques. All network switches are initially closed converting a radial into meshed network. Network switches are then opened one at a time until a new radial configuration is reached. In this process, the switch to be opened at each stage is selected in order to minimize resistive line losses of resulting network.

The second heuristic approach achieves loss reduction by performing switch exchange operations. A switch exchange operation corresponds to the selection of a pair of switches, one for opening and the other for closing, so that resulting network has lower line losses while remaining connected and radial. Other network operating constraints, such as voltage limits and conductor ampacities are taken into account prior to or after performing switch exchange operation.

Expert systems, genetics algorithm, and neuron networks are also found application in solving of reconfiguration problem. Expert systems observe problem of finding optimal configuration as combination and heuristic problem. Application of genetics algorithm has shown as very efficient from the aspect of the time needed for solving of problem.

Voltage regulation

There are system and regional regulation of voltage levels. System regulations are performed by changing motion of synchronous generator. However, voltage regulation can be performed as regional by installing equipment for changing reactive power flows. Node voltages and reactive power flows are in correlation, therefore we can affect on voltage profiles by installing compensation accessories.

Regional voltage regulation can also be made using tap changing transformers. There are two types of these transformers on which:

- tap changing is possible only when transformer is unloaded,
- tap changing can be made under loading.

Obviously, only second type of transformer can be used for regulation of voltages. Voltage regulation in distribution networks by tap changing transformers are based on the principle of voltage drop compensation of unique line. This means that consumer installed into some node gives constant voltage level regardless of load changing or voltage changing on high voltage busses. This regulation is known as compounding voltage regulation. Real and reactive power of node i are change with voltage changing as:

$$P_i = P_{ri}\left(\frac{V_i}{V_r}\right)^{k_{pv}}, \tag{1}$$

$$Q_i = Q_{ri}\left(\frac{V_i}{V_r}\right)^{k_{qv}}, \tag{2}$$

where V_i is voltage of node i, V_r rated voltage of network, k_{pv} and k_{qv} static coefficients of real and reactive power change with voltage changing.

Reactive power compensation

Three principal problems occur on compensation of reactive power [7-9]:

- selection of compensation equipment type,
- optimal placement of equipments,
- optimal size of compensation equipments.

Solving of these problems is very difficult and requires considering of following factors: power and energy losses, costs of delivered power and energy, occupation of line and transformer capacitances, bed consequences due to decreasing of voltage quality. Economic valorisation of all these factors is difficult. Therefore, problem is solved considering only investment and costs of power and energy losses whence influence of other factors are neglected.

When we are finding optimal solution of compensation problem, optimal criterion must be presumed. This criterion has effect on selection of objective function. Mathematical model for compensation problem is made of mentioned objective function and constraints.

Optimal planning of compensation accessories depend on the nature of reactive loading. If reactive loads are “stilly” compensation only lead to better node voltages and power factors. If nonlinear character of loads is dominated or loading changes rapidly, it is possible to occur overvoltages or flicker. Then selection of compensation equipment is made in order to satisfy technical normative.

Extensive usage of static compensation equipments is caused by great progression of thiristor technique. Selection of optimal pleasing of capacitors and their size depend on the used mathematical model. This problem can be solved on the way that nodes for installing of capacitors are selected at the first, and then optimal size of capacitors is found. Second approach is solving of these two tasks simultaneously.

Replacement of conductors

When detail analysis of energy losses is made, conductors of lines that are identified as potential “neck” of network (large losses) can be replaced by ones with bigger cross section area in order to reduce energy losses. Selection of lines that will be replaced and optimal cross section areas of new conductors we should made comprehending all possible operational measures for losses reduction. Calculating annual costs before and after replacement for any distribution line, we can obtain loading on which replacement of conductor is reasonable.

III. TEST EXAMPLE

Analyses are made on two representative networks, and due to similar conclusions are obtained, in this paper are presented only results of 10 kV distribution network shown on the Fig. 1. This is one typical suburban network. Total length of 158 lines is 48 km (63 underground cables with total length 27.3 km, 95 overhead lines with total length 20.7 km), 72 distribution transformers with total rated power 30 MVA (Table I). The most frequently used are aluminium cables with cross section area of 150 mm^2, and ACSR overhead lined with 50 mm^2 and 70 mm^2 cross section areas. Consumers are dominantly individual residential houses. The houses are heating on coal, wood or electrical energy. Power factor in network under analyses is 0.95. This network is one typical suburban network in our country. Small circles on Fig. 1 represent substations 10/0.4 kV/kV.

TABLE I
RATED POWERS AND NUMBER OF DISTRIBUTION TRANSFORMERS FEEDED BY MEDIUM VOLTAGE NETWORK SHOWN IN FIGURE 1

Rated power (kVA)	100	160	250	400	630	1000
Number of transformers	4	4	34	6	16	8

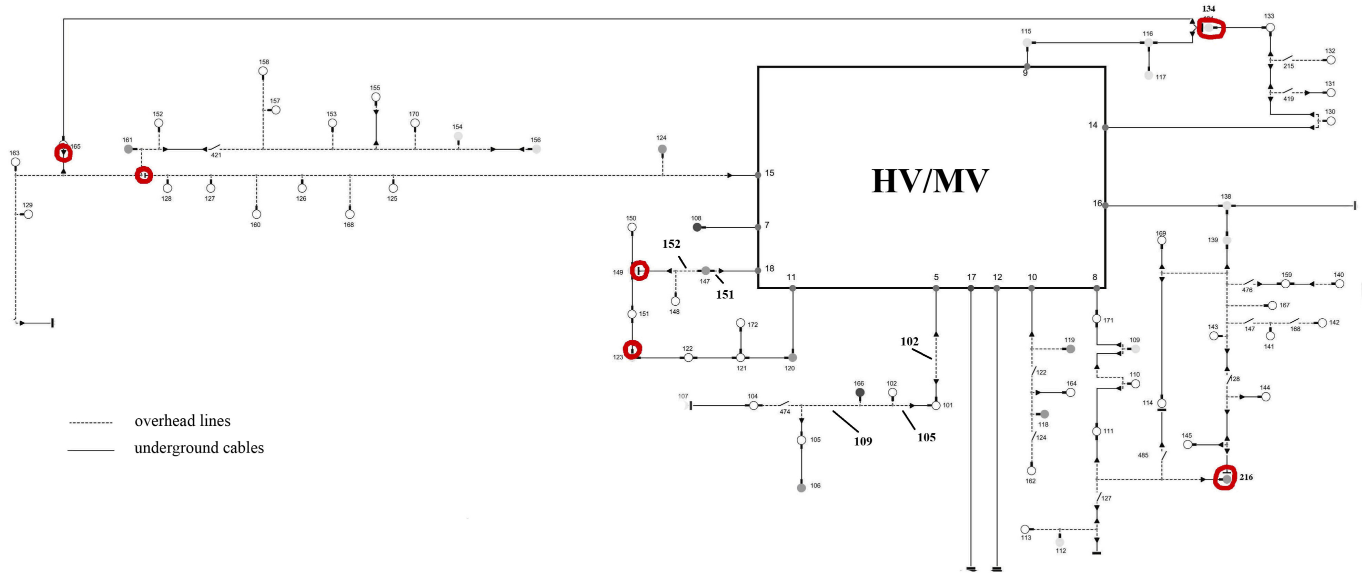


Fig 1. Medium voltage 10 kV suburban distribution network

TABLE II
INFLUENCE OF VOLTAGE REGULATION ON THE ENERGY LOSSES

		k_{pv}=2, k_{qv}=2		k_{pv}=1, k_{qv}=1		k_{pv}=0, k_{qv}=0		k_{pv}=1.7, k_{qv}=2.5	
U_{rn}	1	1.05	0.95	1.05	0.95	1.05	0.95	1.05	0.95
Delivered energy (MWh)	52207.5	54910.9	44949.7	53557.8	48397	52168.3	52325.6	54497.4	45965.6
Total energy losses (MWh)	1252.34	1318.76	1079.53	1266.75	1208.55	1213.87	1371.12	1306.54	1108.06
Percentage losses (%)	2.40	2.40	2.40	2.36	2.50	2.33	2.62	2.40	2.41
Line losses (MWh)	663.38	699.84	572.88	663.55	662.92	626.37	777.28	691.32	592.78
Transformer losses (MWh)	588.96	618.92	506.65	603.20	545.63	587.50	593.84	615.22	515.28

The data needed for analyses are collected for period from September 1 to March 31. Maximal value of real power of network during observed interval is 20.5 MW, and total delivered real energy is 52207.5 MWh. Based on these data at the first we made estimation of node loads during interval under consideration.

Based on results of load estimations and load flow calculations, for each moment of observed interval we calculated real and reactive node powers for four combinations of k_{pv} and k_{qv}. Results of analyses of voltage regulation (Table II) show that absolute value of energy losses as well as total delivered energy depend on the voltages as well as on the load category of consumers. Therefore, percentage energy losses should observed for concluding how regulation of voltage affect on losses. Results shown in Table II show that voltage of root node (U_{rn}) should be keep on maximal values whenever it is possible.

For state of network in the moment when total delivered power is maximal we analyzed reconfiguration for losses reduction. Results of calculation show that eight switching operations should be performed to obtain optimal configuration (four switches are opened and same number are closed). Switches that should be opened/closed are denoted on the Fig. 1. Table III shows energy loss calculations results before and after reconfiguration is made. For considered interval energy losses will be decreased for 154.8 MWh if reconfiguration is made.

TABLE III
DECREASING OF ENERGY LOSSES EFFECTED BY RECONFIGURATION

	Actual configuration	After reconfiguration	difference
Delivered energy (MWh)	52207.5	52052.7	154.8
Total energy losses (MWh)	1252.34	1105.21	147.13
Line losses (MWh)	663.38	516.22	147.16
Transformer losses (MWh)	588.96	588.99	-0.03

Analyse of optimal location and sizes of capacitors are made for actual configuration of network as well as for the state obtained after reconfiguration was made. Having in mind standard sizes of capacitors, and using optimization approach, we conclude that installation of capacitors in substations 10/0.4 kV/kV should be made on following way:

- if rated power of distribution transformer is 1000 kVA capacitors of 90 kVAr are installed,
- if rated power of distribution transformer is 630 kVA capacitors of 60 kVAr are installed.

This means that in analyzed network we should install 8 capacitors of 90 kVAr and 16 capacitors of 60 kVAr. Results of losses reduction due to compensation an actual configuration are shown in Table IV, whence results of compensation after reconfiguration is made are shown in Table V. Returning time of investments in capacitors is 2.3 years.

TABLE IV
RESULTS OF REACTIVE POWER COMPENSATION

	Without compensation	With compensation	difference
Delivered energy (MWh)	52207.5	52154.2	53.3
Total energy losses (MWh)	1252.34	1199.73	52.61
Line losses (MWh)	663.38	624.64	34.74
Transformer losses (MWh)	588.96	575.09	13.87

TABLE V
RESULTS OF REACTIVE POWER COMPENSATION AFTER RECONFIGURATION OF NETWORK IS MADE

	Without compensation	With compensation	difference
Delivered energy (MWh)	52052.7	52007.1	45.6
Total energy losses (MWh)	1105.21	1060.81	44.4
Line losses (MWh)	516.22	485.65	30.57
Transformer losses (MWh)	588.99	575.16	13.83

TABLE VI
LINE LOSSES BEFORE AND AFTER CHANGING OF CONDUCTORS

	Losses before changing (MWh)	Losses after changing (MWh)	Returning time of investments
102	38.00	18.8	2.66
105	15.81	10.97	10.56
109	15.14	10.51	16.56
151	102.6	71.22	4.07
152	87.58	60.80	4.77

Analysing structure of losses (losses distribution throughout the network elements) we identified five lines on which replacement of conductor may be reasonable. These are overhead lines denoted in the Fig.1 by numbers 102, 105, 109, 151, and 152. Cross section areas of these lines are 35 mm^2 and 50 mm^2. Energy losses in these lines for actual state of network is shown in the second column of Table VI. If mentioned lines are replaced with ACSR 70 mm^2 conductors, losses are decreased as shows third column of Table VI. Column four of Table VI contain data about returning times of investment on changing of conductors. Replacement of line 102 is obviously payable, whence for lines 151 and 152 we must consider long time forecasting of loads.

IV. CONCLUSION

Effects of measures for energy loss reduction are analyzed in this paper. Analyse is made on real suburban medium voltage distribution network. Results of calculations show that reconfiguration and voltage regulation should be applied whenever is possible. Additionally, it is shown that investment in installing of capacitor banks is reasonable regardless organisation measures are previously applied or not. Changing of some conductors (increasing of cross section area) also may have very short returning time of investment and therefore must be considered in analyse.

REFERENCES

[1] В. Э. Воротницкий, Ю. С. Железко, В. Н. Казанцев, В. Г. Пекелис, Д. Л. Файбисович, Потери электроэнергии в электрических сетях энергосистем, Энергоатомиздат, Москва, 1983.

[2] Ю. С, Железко, Выбор мероприятий по снижению потери электроэнергии в электрических сетях, Энергоатомиздат, Москва, 1989.

[3] S. Cilanvar, J. J. Grainger, H. Yin, S. S. H. Lee, "Distribution Feeder Reconfiguration for Loss Reduction", IEEE Trans. on Power Delivery, Vol. 3. No. 3, July 1988, pp. 1217-1223.

[4] D. Shirmohammadi, H. W. Hong, "Reconfiguration of Electric Distribution Networks for Resistive Line Losses Reduction", IEEE Trans. on Power Delivery, Vol. 4. No. 2, April 1989, pp. 1492-1498.

[5] M. E. Baran, F. F. Wu, "Network Reconfiguration in Distribution Systems for Loss Reduction and Load Balancing", IEEE Trans. on Power Delivery, Vol. 4. No. 2, April 1989, pp. 1401-1407.

[6] S. K. Goswami, S. K. Basu, "A New Algorithm for the Reconfiguration of Distribution Feeders for Loss Minimization, IEEE Trans. on Power Delivery, Vol. 7. No. 3, April 1992, pp. 1484-1491.

[7] T. J. E. Miller, *Reactive Power Control in Electric Systems,* John Wiley and Sons, New York, 1972.

[8] D. Tasic, "Selection of optimal locations and sizes of capacitors in electric networks", Elektrodistribucija, No 9-10, 1991, pp. 348-353 (in Serbian).

[9] D. Tasic, N. Rajakovic, Z. Stajic, "Linearized model for an Optimal Exploitation of Reactive Power Resources in Industrial Networks", Proceeding of 35th Universities Power Engineering Conference UPEC 2000, Belfast, September 2000.

[10] T. Gönen, *Electric Power Distribution System Engineering,* Mc Grow-Hill, New York, 1986.

Analysis of the normative documents for the electrical energy quality

Atanas Genchev

Abstract – **Ensuring and supporting of the electrical energy quality is a basic duty of the electric supplying companies. The consumers have every right to require and to get qualitative electrical energy. The operative standards of the electrical energy quality are studied in the paper.**

Keywords – **electrical energy, quality, voltage**

I. INTRODUCTION

The combination of such characteristics of the supplying system, in which consumers of electrical energy can execute functions deposited in them, is defined by the general term *electrical energy quality*. The electrical energy quality has two basic components – continuity and voltage level.

Often the concept of electrical energy quality is used to describe the specific characteristics of the supply voltage. The most important of them are: harmonics and voltage deviation; voltage fluctuations, especially these caused flicker; voltage sags and short interruptions; voltage asymmetry in three phase systems; transitory overvoltages.

Today, the complication of the technological processes imposed the wide using of variable semiconductor drives, power excavators, arc welding furnaces, arc welding aggregates and electric trucks. Characteristic feature of these loads is that they affect on the electrical energy quality of the supplying network. The normal work of the electric equipments depends on electrical energy quality of the power supplying network. The mutually influence of the electrical equipment and the supplying system is defined by the term *electromagnetic compatibility*.

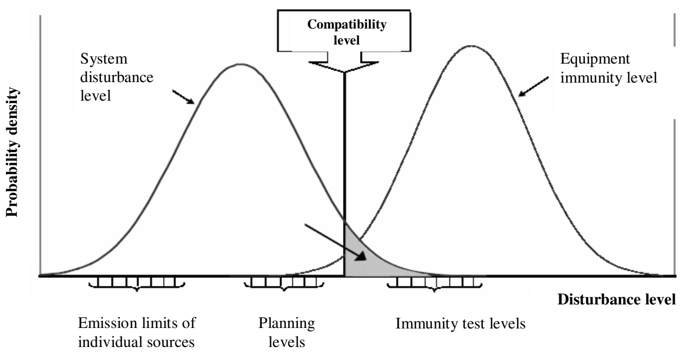


Fig. 1. Probability distribution of the disturbance level and immunity level of public low voltage power system

Fig. 1 shows probability distribution of the disturbance level and immunity level of public low voltage power system [1].

The most probable causes for worsening electrical energy quality are given in Table I.

TABLE I
THE MOST PROBABLE CAUSES FOR WORSENING ELECTRICAL ENERGY QUALITY

Electrical energy quality indexes	**The most probable causes**
Voltage deviation	Electric supplying company
Voltage fluctuation	Users with non-linear load
Harmonics	Users with non-linear load
Voltage asymmetry in three phase systems	Users with non-linear load
Frequency deviation	Electric supplying company
Voltage sags	Electric supplying company
Pulse voltage	Electric supplying company
Temporary overvoltage	Electric supplying company

II. OPERATIVE STANDARDS AND NORMATIVE DOCUMENTS FOR ELECTRICAL ENERGY QUALITY

European parliament and Council directive 96/92/EC is accepted in 1996. This directive defines the frame of producing, transfer, distribution and supplying of electric energy to the end users. It determines rules of an organization, functioning of the united energy market. The electric energy production and wholesale market are efficiently opened for competition in most European countries, but the retail market is liberated to a different extent in each country [2].

There are two base European Union directives about electromagnetical compatibility. Directive 85/374 since July 1995 describes the electrical energy as a product, which has to posses the obligatory requiring indexes for quality. Directive 89/336 since May 1989 describes details of electromagnetical compatibility and radiation from the electrical power systems. This is the one of so - called new approach directives. It is typical for this approach to be given only in the most general and important demands along with administrative principles. The detailed technical demands are not included here, which gives an opportunity for their developing and changing without revising the directive itself.

Numbers of harmonized European standards for electromagnetical compatibility are applied to directive 89/336: EN 50065, EN 50081, EN 50082, EN 50083, EN 50091, EN 61000.

In 1999, European organization for standardization in the electromagnetic field approved European standard EN 50160, which reflects the theoretical level, level of the measuring techniques and exploitation practice in the electrical energy quality field during the past years. It includes the following basic indexes of electrical energy quality [3]: frequency

Atanas Genchev is PhD student in the Faculty of Electrical Engineering, Technical University - Sofia, Kliment Ohridski 8, 1000 Sofia, Bulgaria, E-mail: asgen@lycos.com

TABLE II

NORMS OF THE INDEXES TO ELECTRICAL ENERGY QUALITY FOR LOW AND MIDDLE VOLTAGE ELECTRICAL DISTRIBUTION NETWORKS ACCORDING TO EN 50160

Characterization	**Low voltage networks**	**Middle voltage networks**
Frequency	49,5-50,5 Hz (for 99,5% from 1 year period) or 47-52 Hz (whole year)	49,5-50,5 Hz (за 99,5% from 1 year period) or 47-52 Hz (whole year)
Voltage deviation	U_H ±10% (for every period from one week, 95% from the average effective voltage value per 10 min); U_H +10/-15% (for every period from 1 week, all average effective voltage values per 10 min)	U_H ±10% (for every period from one week, 95% from the average effective voltage value per 10 min); U_H +10/-15% (for every period from 1 week, all average effective voltage values per 10 min)
Fast voltage fluctuations	Less than 5% U_H; fluctuations up to 10% U_H with short duration may advent few times per day in some conditions. Flicker: $P_{lt} \leq 1$ (for 95% from period from 1 week)	Less than 4% U_H; fluctuations up to 6% U_H with short duration may advent few times per day in some conditions. Flicker: $P_{lt} \leq 1$ (for 95% from period from 1 week)
Voltage asymmetry	95% from the average effective voltage value with back sequence per 10 min must to be in limits from 0 to 2% U_H from the right sequence for every period of 1 week. In some power network areas may have values up to 3% U_H	95% from the average effective voltage value with back sequence per 10 min must to be in limits from 0 to 2% U_H from the right sequence for every period of 1 week. In some power network areas may have values up to 3% U_H
Harmonics	95% on the average effective voltage value of each harmonic formation of voltage per 10 min for every period of one week must be: $U_3 \leq 5\%$, $U_5 \leq 6\%$, $U_7 \leq 5\%$, $U_{11} \leq 3.5\%$, $U_{13} \leq 3\%$; total harmonic distortion ≤ 8%	95% on the average effective voltage value of each harmonic formation of voltage per 10 min for every period of one week must be: $U_3 \leq 5\%$, $U_5 \leq 6\%$, $U_7 \leq 5\%$, $U_{11} \leq 3.5\%$, $U_{13} \leq 3\%$; total harmonic distortion ≤ 8%
Voltage dip	Expected number may be from few scores to one thousand per one-year period.	Expected number may to be from few scores to one thousand for a year period.
Short-time interruptions	Values: from a few scores to several hundred	Values: from a few scores to several hundred
Long-time interrupting	Values: (interrupting over 3 min) annual frequency from 10 to 50, depending on area	Values: (interrupting over 3 min) annual frequency from 10 to 50, depending on area

deviation; voltage deviation; fast voltage fluctuations; voltage asymmetry; harmonics; short-time and long-time interrupttions; voltage dips. The norms of electrical energy quality indexes for low and middle voltage electric distribution networks are given in Table II according to EN 50160.

The basic indexes for the electrical energy quality are legalized in Russian state standard for electrical energy quality GOST 13109-97 [4]: voltage deviation; flicker; harmonics; voltage asymmetry; frequency deviation; voltage dip; pulse voltage; momentary overvoltage.

BSS 10694-80 continuing to be operative for user's indexes electrical energy quality in Bulgaria [5]. This standard and all state standards have advisable statute in virtue of the National standardization act (become operative since May 05, 2006). Moreover it contains smaller indexes than in the European EN 50160. In this standard are legalized basic indexes for electrical energy quality: frequency deviation in normal work condition, when the system works independently; frequency fluctuation in normal work condition; voltage deviation at user's terminals; average statistic value of the voltage deviation for receivers supplied from power network for general-purpose; voltage value with back consistency; voltage value with zero consistency in networks with mono-phase electrical lamps and home receivers; harmonics. The norms of electrical energy quality, according to BSS10964-80, are shown in Table III.

In conformity with the general tendency in Bulgaria, the number of standards was accepted in the field of electromagnetic compatibility. These standards determine the electrical energy quality.

The State Energy and Water Regulatory Commission accepted and published the quality indexes of the electrical energy [6], which are identical to these in EN 50160, in June 2004.

III. ANALYSIS OF THE NORMATIVE DOCUMENTS FOR THE ELECTRICAL ENERGY QUALITY

Ensuring and supporting electrical energy quality is a basic duty of the electric supplying companies. On the other hand, the consumers have right to require and receive qualitative electrical energy. But they should to have obligations, with their consumers and their work regimes, not to worsen the indexes of electrical energy quality in electric supplying systems.

In the last years, in Bulgaria begin to pay attention on the activities for ensuring and control of the indexes of the electrical energy quality. Not only electric supplying companies are interested in this, but consumers who more often express their claims, determined by realizing their rights or by introducing modern equipments. Their functioning is connected with requirements to the quality of the electrical energy.

From the differences between EN 50160 и BSS 10694-80, we will separate special attention to these, which refer to interruptions of the electric supply.

TABLE III
NORMS FOR ELECTRICAL ENERGY QUALITY ACCORDING TO BSS 10964-80

Indexes	Standard values
1. Frequency deviation in normal regime, when system works independently.	± 0,1 Hz; Difference between the astronomic and synchronous time < 2 minutes. The system is allowed to work temporary with ± 0,2 Hz
2. Frequency fluctuations in normal regime.	Less than 0,2 Hz, bigger from the deviation in point 1. Frequency deviations at the points above are not valid for after-failure recovery power system period.
3. Voltage deviation at user's terminals except for given below: 3.1. Voltage deviation at the terminals of electric motors and start stop devices.	Not bigger than ± 5 %U_H In limits between – 5 to 10 %
4. Average statistical value of voltage fluctuation for receivers, supplied from power network for general purpose, except for given below: 4.1. Average statistical value of the voltage fluctuation: - Of the electrical lamps and radio-electronic device terminals; - Of the radio apparatus terminals, dispatcher and audio system, control-measuring devices and computers; - In the receivers with AC supply.	Not bigger than 2,5 %U_H (by unlimited frequency fluctuations) -Values smaller than given in a graph: fluctuation size in % in a function of frequency fluctuation in Hz; - Not bigger than 1,5%U_H (in not limited frequency of the fluctuations); -Not bigger than 5 %U_H (in frequency up to 25 Hz)
5. Voltage value with back consistency: - In three phase networks with electrical lamps and public receivers; - Of the terminals in each three phase symmetrical receiver.	- Voltage deviation have to be in the limits given in points 3 and 3.1; - Not bigger than 2 %U_H.
6. Voltage value with zero consistency in networks with mono-phase electrical lamps and public receivers.	Voltage deviation must be in the limits given in points 3 and 3.1;
7. Voltage harmonics of all receivers - Of the electrical lamps and public receivers terminals.	The effective value of all high harmonics must be lower than 5% from the effective value of the voltage with main frequency: - Value not bigger than this in which deviation limits are given in points 3 and 3.1, attended to the effective value of the voltage with main frequency.

The following basic indexes for interruptions of the electric supply are determined:
- Total number of the electric supply interruptions for definite period (month, three months, a year);
- Total continuity of the electrical supply interruptions for definite period (month, three months, a year);
- Average continuity of one interruption.

In the most European countries are adopted the electrical supply continuity indexes using in the practice of the USA electric supplying companies. These indexes have statistical character and they bear a proportion of the interruptions and the duration of the interruptions to the total number of connected consumers or to the number of interrupted consumers.

In the standard EN 50160 interruptions are determined as:
- Planed, for which the consumers are informed in advance from the electric supplying company;
- Accidental (non-planed), which are caused by stable or transitive failures of equipments, incorrect manipulations etc., and for which on the whole is not possible to inform consumers in advanced.

The accidental interruptions are separated to short and long-time interruptions.

In the standard EN 50160, short-time interruptions are those with duration less than 3 minutes, and long-time interruptions – with the bigger duration than 3 minutes. For the different European countries and for the electric supplying companies this limit is accepted from 1 to 5 minutes and depends on the level of automation and telecontrol in the substations and the distribution networks.

At the present stage, in Bulgaria can be accepted a limit of 10 minutes, which depends on network modernization and taking into the automate control in the future could be reduced.

The accidental interruptions indexes are following:
- System Average Interruption Frequency Index, number/year:

$$\mathrm{SAIFI} = \frac{\text{Total number of interruptions}}{\text{Total number connected consumers}}, \qquad (1)$$

- System Average Interruption Duration Index, minutes:

$$\mathrm{SAIDI} = \frac{\text{Total duration of the interruptions}}{\text{Total number suppyed consumers}}, \qquad (2)$$

- Customer Average Interruption Frequency Index, number/ year:

$$\mathrm{CAIFI} = \frac{\text{Total number interruptions}}{\text{Total number interrupted consumers}}, \qquad (3)$$

- Customer Average Interruption Duration Index, minutes:

$$\mathrm{CAIDI} = \frac{\text{Total duration of the interruptions}}{\text{Total number interruptions}}. \qquad (4)$$

For events, which have accidental character is not always possible to use simple certain indexes. Because of this in the practice of many countries there are not normative electric supply interruption indexes (accidental interruptions), observation of which is controlled or took in consideration customer demands. The definite numerical values of these indexes are used for comparison valuations and analysis.

The imperfections of BSS 10694-80 and taking into work of EN 50160, approved to be accomplished series of activities as: acquainting with the wide circle of specialists with the requirements and prescriptions in the European standard, and adopting respective methods and equipment for studying electrical energy quality indexes in a comparatively short period. Not in the last place, there would be established and estimated the real index quality parameters in our networks, to plan the concrete steps for supplying an appropriation with the requirements and norms, put in EN 50160.

IV. Conclusion

From the analysis and comparison of international standards and normative documents for electrical energy quality with the working standards and normative documents in Bulgaria, could be made the following conclusions:

- In the last years is observed a tendency of growing numbers and capacity of the consumers, but also these who made increasing requirement to the quality.
- Corresponding with the preparation activities for the future joining of our country to the European Union and the adopted policy for harmonization of the Bulgarian standards to these of the European Union, many European standards are accepted in our country in the field of electromagnetic compatibility. These standards determined also indexes which are direct or indirect connected with the quality of supplying electrical energy, as with voltage sags, voltage fluctuations, voltage harmonics, flicker and etc.
- The present operative electrical energy quality standard in Bulgaria BSS 10694-80 is outdated and consisted less indexes than the European standard for electrical energy quality EN 50160.
- For ensuring electrical energy quality it is necessary the indexes to be in correspondence with EN 50160 and electrical supplying indexes adopted by Resolution №P-3 since June 30, 2004 of the State Energy and Water Regulatory Commission.

Literature:

[1] Schlabbach J., D. Blume and T. Stephanblome, Voltage Quality in Electrical Power Systems, UK, 2001.

[2] Tzanev T., S. Zaanov, B. Kibarow, Power Quality – Element from the Relations between the Electric Supplying Company and Consumers, Energy Forum'2004, Varna, 2004.

[3] EN 50160:2000 „Voltage Characteristics of the Electrical Energy Delivered from the Public Supplier".

[4] GOST 1309-97 "Norms for Electrical Energy Quality in Electric Supplying Systems".

[5] BSS 10694-80 „Norms for the Indexes for Electrical Energy Quality".

[6] "Indexes for Electric Supplying Quality", State Energy and Water Regulatory Commission, June 2004.

Wireless Technology in Body Area Networks for Multiplex Sclerosis Patient Monitoring

A. Peulic[1], S. Randjic[2], A. Dostanic[3], M. Acovic[4]

Abstract – **In this moment, measuring of physiological data is performed at hospitals or laboratories where patients are trapped with many electrodes attached on the body. Wireless networking will liberate people from such confinement and enable continuous real time monitoring of physiological data, which is vital for medical care. To achieve this goal, it is necessary research and development wireless body area network for signal processing, wireless transmission, and databases for these vital data. This paper presents arm moving wireless monitoring by accelerate sensors for patients with sclerosis multiplex**

Keywords – **Wireless Technology, Body Area Network, Intelligent Sensors, Health Care Information System, sclerosis multiplex**

I. Introduction

To provide human healthcare support with a better quality, we should be able to collect a very large amount of people's vital signs and monitor it efficiently. Current welfare system is based on medical doctor regular consultation, on behalf of our own feeling. The idea is not to replace the current system, but to augment it by an environment using wireless networking, to provide continuous monitoring of one's physiological information, perform simple diagnosis and communicate all that with medical institutions.

The interest for wearable systems originates from the need for monitoring patients over extensive periods of time. This case arises when physicians want to monitor individuals whose chronic condition includes risk of sudden acute events or individuals for whom interventions need to be assessed in the home and outdoor environment. If observations over one or two days are satisfactory, ambulatory systems can be utilized to gather physiological data. An obvious example is the use of ambulatory systems for ECG monitoring, which has been part of the routine evaluation of cardiovascular patients for almost three decades. However, ambulatory systems are not suitable when monitoring has to be accomplished over periods of several weeks or months, as is desirable in a number of clinical applications.

Wearable systems are devices that allow physicians to overcome the limitations of ambulatory technology and provide a response to the need for monitoring individuals over weeks or even months. They typically rely on wireless, miniature sensors enclosed in patches or bandages, or in items that can be worn, such as a ring or a shirt.

[1]Aleksandar Peulic is with the Tehnical Faculty, Svetog Save 65, 32000 Cacak, Serbia, E-mail: apeulic@tfc.kg.ac.yu
[2]Sinisa Randjic is with the Tehnical Faculty, Svetog Save 65, 32000 Cacak, Serbia, E-mail: rasin@tfc.kg.ac.yu
[3]Adam Dostanic is with the Tehnical Faculty, Svetog Save 65, 32000 Cacak, Serbia, E-mail: adamdostanic@tfc.kg.ac.yu
[4]Marko Acovic is with the Tehnical Faculty, Svetog Save 65, 32000 Cacak, Serbia, E-mail: markoacovic@tfc.kg.ac.yu

We propose a system using wearable micro sensors, to monitor patient's physiological information such as blood-flow, pulse, temperature, EKG, EMG, EEG, etc. On a wearable controller, a software environment allows accumulating data from physiological sensors, recording them into a local database, operating basic data manipulations, and communicating data with the database at medical institutions. These elements compose health remote control systems, which concept is shown on Fig. 1.

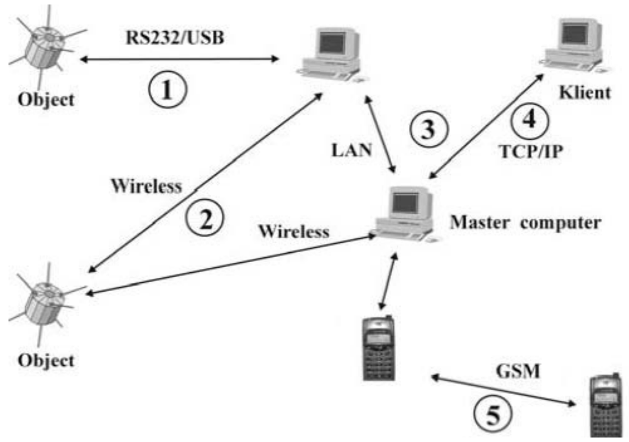


Fig. 1. Health remote control system

It consists of several measurement modules, a local wireless body area network, a local wire network and GSM network.

II. Methods

Slave board Fig. 2 is based on 868 MHz RF TI TRF6900 Transceiver. The RF transceiver data and controls lines are connected to low-power Texas Instruments microcontroller MPS430F149. The output of signal conditions circuit is connected to microcontroller's 12-bit AD converter. The proposed communication algorithm is implemented in microcontroler memory.

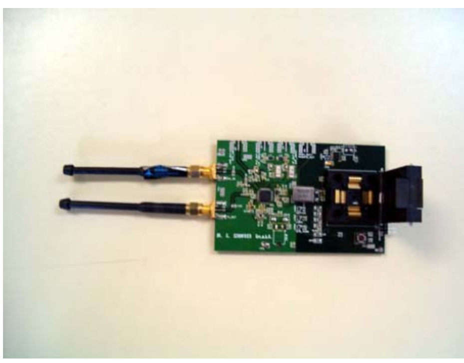


Fig. 2. Wireless module

Wireless Body Area Network is realized like cyclic network, Fig. 3.

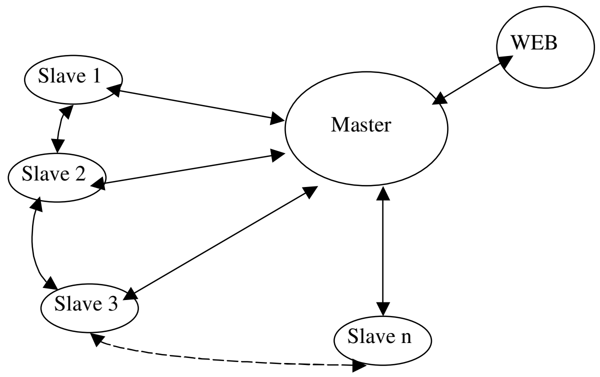


Fig. 3. Wireless Body Area Network

Proposed system is designed for up to 256 modules. Any slave sends own address, two bytes, measured signals information, up to 16 bytes, 8 measured signals, and two bytes of next slave address. Last slave expect first if there is any new slave at network, and if there is not new slave, sends address of the first slave. In master module is implemented agent algorithm. It is programs procedure for case that there is not answer from any slave. Agent scans slaves and reprograms network. One byte is transmitted for 8*13.02μs and for whole slave information there is need 2187.36μs. For all slaves with 8-channel need is about 600ms. USB 2 connects the master module to server PC. 0.

III. Discussion

Wireless intelligent sensors have made possible a new generation of non-invasive, unobtrusive personal medical monitors applicable during normal activity. Sensor intelligence allows implementation of real-time processing and sophisticated encryption algorithms. On sensor data processing decreases the amount of energy spent on communication and allows implementation of power-efficient communication protocols. Decreased power consumption will significantly increase battery life and even enable externally powered intelligent sensors. The current technological trend will allow wider use of wireless intelligent sensors, lower power consumption, and smaller sensor sizes.

In case of medical emergency master module can send an SMS message to the centre medical doctor.

Daily monitoring of patient's vital signals will make possible a broad range of healthcare application services. It allows enhancing accuracy of diagnosis, networked system allows several doctors to monitor the same patient and detect pathological change in its early stage, evaluate behaviour of elderly people with dementia, quantitatively estimate human physiological condition in daily life, and support interactive home care.

Developed system is based on wearable sensors to enable continuous monitoring of patient's physiological information. A wearable controller collects wirelessly sensor data, integrate it into a database, which allow the exchange with medical institutions where a system manages the database for each patient's vital data. Improvements of the physiological information integration system concern simultaneous sensor communication management method, and algorithm evaluation to detect simple abnormalities. Simultaneous use of several wearable sensors allowed performing correlation analysis, which is the source of sensor system simplification (information redundancy). The determination of minimum sensing environment, size and energy consumption improvements are critical to adapt the system to practical use.

Acceleration sensor is placed at patients arm or leg and connected to RF transmitter, Fig. 4.

Fig. 4. The sensor place at the patient arm

The ADXL311 is a low cost, low power, complete dual-axis accelerometer with signal conditioned voltage outputs, all on a single monolithic IC. The ADXL311 is built using the same proven iMEMS® process used in over 100 million Analog Devices accelerometers shipped to date, with demonstrated 1 FIT reliability (1 failure per 1 billion device operating hours). The ADXL311 will measure acceleration with a full-scale range of ±2 g. The ADXL311 can measure both dynamic acceleration (e.g., vibration) and static acceleration (e.g., gravity). The outputs are analog voltages proportional to acceleration.

The typical noise floor is 300 μ g/$\sqrt{Hz}$ allowing signals below 2 m g (0.1° of inclination) to be resolved in tilt sensing applications using narrow bandwidths (10 Hz). The user selects the bandwidth of the accelerometer using capacitors CX and CY at the XFILT and YFILT pins. Bandwidths of 1 Hz to 2 kHz may be selected to suit the application.

IV. Experimental Results

Measured and transmitted accelerate signals are presented at Figure 5, X axis and Figure 6, Y axis.

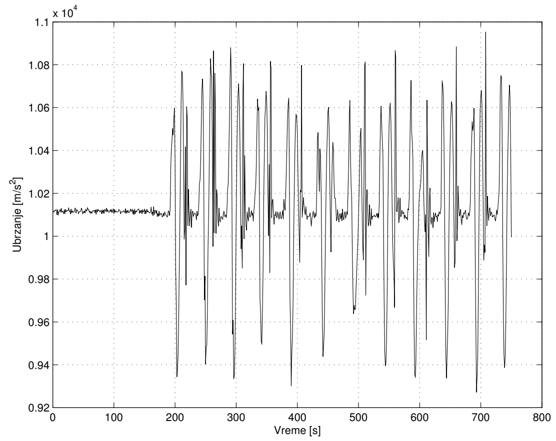


Fig. 5. The transmitted accelerate signal X axis

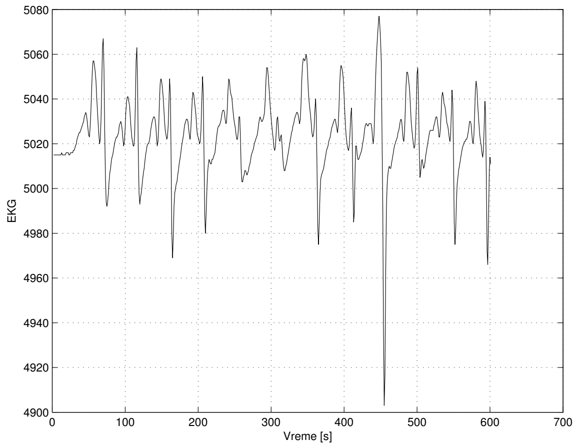


Fig. 6. The transmitted accelerate signal Y axis

When the accelerometer is oriented so both its X-axis and Y-axis are parallel to the earth's surface, it can be used as a two axis tilt sensor with a roll axis and a pitch axis. Once the output signal from the accelerometer has been converted to an acceleration that varies between –1 g and +1 g, the output tilt in degrees is calculated as follows:

$$tilt = a\sin(X/1g) \quad (1)$$

$$rool = a\sin(Y/1g) \quad (2)$$

Figure 7 represents right arm moving angle in XY plane measured in scale of about 300 sec, and figure 8 left arm moving angle.

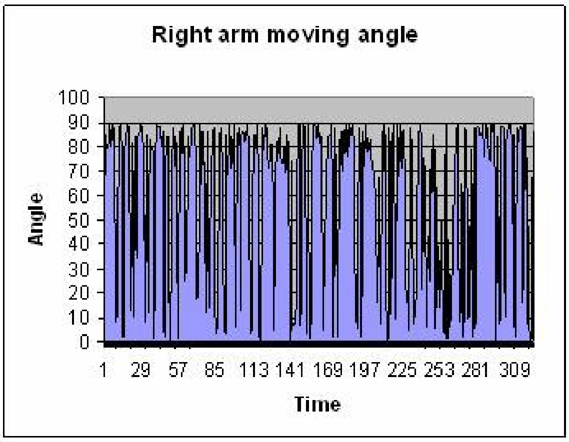


Fig. 7. Right arm moving angle

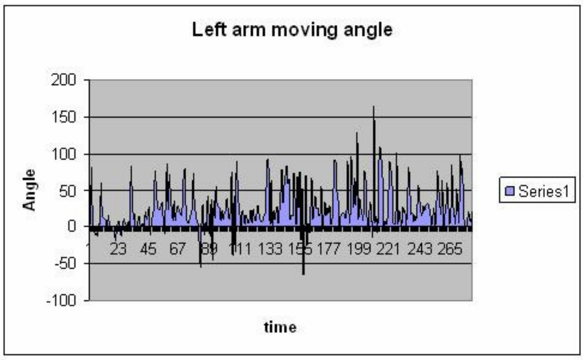


Fig. 8. Left arm moving angle

V. REFERENCES

[1] Ž Čučej, A Peulić, S Randjić: Wireless monitoring system of biomedical data: security issues, IWSSIP_05, Chalkida, Greece, Sep. 20005

[2] A. Peulic, N. Zdravkovic: Integrating radio frequency network and information technologies for wireless collection of measured physiological data, BioMED 2005, February 16 to February 18, 2005, at Innsbruck, Austria.

[3] Peulic, A., Randjic, S.: 'Radio Frequency Measured Data Transmit', ICEST 01.-04. Oct. 2002, Nis, Yugoslavia, CD zbornik, Measurement Technique

[4] Peulic, A., Randjic, S.: 'Computer Based High Radio Frequency Systems Contro', TELSIKS 01.-03. October 2003, Nis, Yugoslavia, Proceedings is on CD

[5] Peulic, A., Randjic, S.: 'Computer Based Remote Measuring and Acquisition of Dynamic Data', ICEST 16.-18. October 2003, Sofia, Bulgaria

[6] Jovanov, E., Price, J., Raskovic, D., Kavi K, Martin, T., 'Adhami R.: Wireless Personal Area Networks in Telemedical Environment' *Proceedings of Third International Conference on Information technology in Biomedicine, (ITAB-ITIS2000), pp. 22-27, 2000.*

[7] Jovanov, E., Price J., Raskovic, D, Moore, A., Chapman, J., Krishnamurthy A.: 'Patient Monitoring Using Personal Area Networks of Wireless Intelligent Sensors', *Biomedical Sciences Instrumentation Vol. 37, in Proc. 38th Annual Rocky Mountain Bioengineering Symposium, RMBS 2001, April 2001, Copper Mountain Conference Center, pp. 373-378*

A Wavelet Based GUI for NM Images Filtering Using Variance-Stabilizing Transformation

Cvetko D. Mitrovski[1] and Mitko B. Kostov[2]

***Abstract* – In this paper we present an application and its interface for pre-processing of chest region NM images. It works in the Matlab environment and uses Matlab user defined functions. The software processes the images in the discrete wavelet transform domain. It uses QMF filters with linear phase designed to achieve both good image decomposition and near perfect reconstruction. Due to the fact that the noise in NM images is highly dependent on the image space distribution and signal intensity, it applies the Anscombe variance-stabilizing transformation over original images. As a result, the transformed images are characterized with a white Gaussian noise model and Donoho's level dependent threshold is used.**

***Keywords* – Nuclear medicine image, wavelet-domain filtering, autocorrelation, denoising.**

I. Introduction

Nuclear Medicine (NM) images are diagnostic digital images, which provide both anatomical and functional information abut a patient. They represent the projection of the distribution of the radioisotopes in the patient body after injection of certain dose of radioisotopes. The raw NM images are based directly on the total number of photon counts detected over fixed observation period by computerized gamma cameras. They are very noisy due to the nature of the gamma ray emission process and the operational characteristics of the gamma cameras (low count levels, scatter, attenuation, and electronic noises in the detector/camera). The noise is highly dependent on the image space distribution and signal intensity. Therefore, a suitable image pre-processing must precede the NM images analysis in order to provide an accurate recognition of the anatomical data of the patient (the boundaries of the various objects – organs). This process of separating signal from noise is a rather difficult and much diversified task that should be adjusted to the organs and tissues, which physiology is to be investigated.

This paper presents an application and its graphic user interface (GUI) for pre-processing of chest-region NM images. It reduces the noise in the discrete wavelet domain. The images are decomposed with QMF filters with linear phase. The filters are designed to achieve both good image decomposition and near perfect reconstruction. Due to the signal-dependence of the noise, the Anscombe variance-stabilizing transformation is applied over the original images. As a result, the transformed images are characterized with a white Gaussian noise model and Donoho's level dependent threshold is used.

[1]Cvetko D. Mitrovski is with the Faculty of Technical Sciences, I.L.Ribar bb, 7000 Bitola, Macedonia, E-mail: cvetko.mitrovski@uklo.edu.mk

[2]Mitko B. Kostov is with the Faculty of Technical Sciences, I.L.Ribar bb, 7000 Bitola, Macedonia, E-mail: mitko.kostov@uklo.edu.mk

The paper is organized as follows. The software for processing NM images and its GUI are shown in the Section II. The Section III discloses the algorithm that is utilized by the software. Ideas for improving the efficiency of the software are considered in the Section IV, while the Section V concludes the paper.

II. The GUI for Processing NM Images

The Fig. 1 shows the GUI we have developed for processsing the NM images. It works in the Matlab environment and uses Matlab user defined functions. It can be used both for displaying raw images and processed images.

The main window on its left side contains two frames (labels 1 and 2 in Fig. 1) where the raw image and the processed image are displayed. By clicking the GUI button 'OPEN IMAGE' (4) a dialog window appears and user can choose a file that has the Matlab extension 'mat'. The set of images that are contained in the selected mat file is shown in the top-right part of the GUI (6). An image selected from this set appears in the top-left frame (1). By clicking the button 'FILTER' (5) this image is processed and the result is shown in the bottom-left frame (2). It is possible to remove the shadow around the object by moving the slider or typing a value in the adequate textbox (8). In addition, the user can choose a colormap from the combobox (7) with a list of defined Matlab colormaps (Fig. 3).

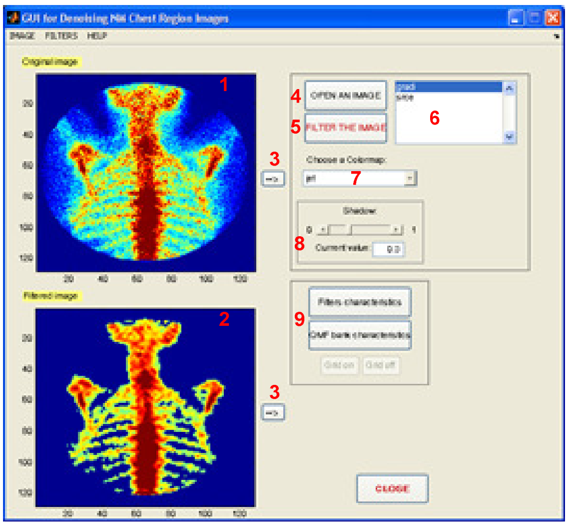


Fig. 1 The main window

1. Top-left frame for the original image; 2. Bottom-left frame for the result; 3. Buttons for exporting the images; 4. Button for reading an image; 5. Button for filtering; 6. The images that are contained in the selected 'mat' file; 7. Colormap combo box; 8. Slider for removing the shadow; 9. QMF bank and filter characteristics

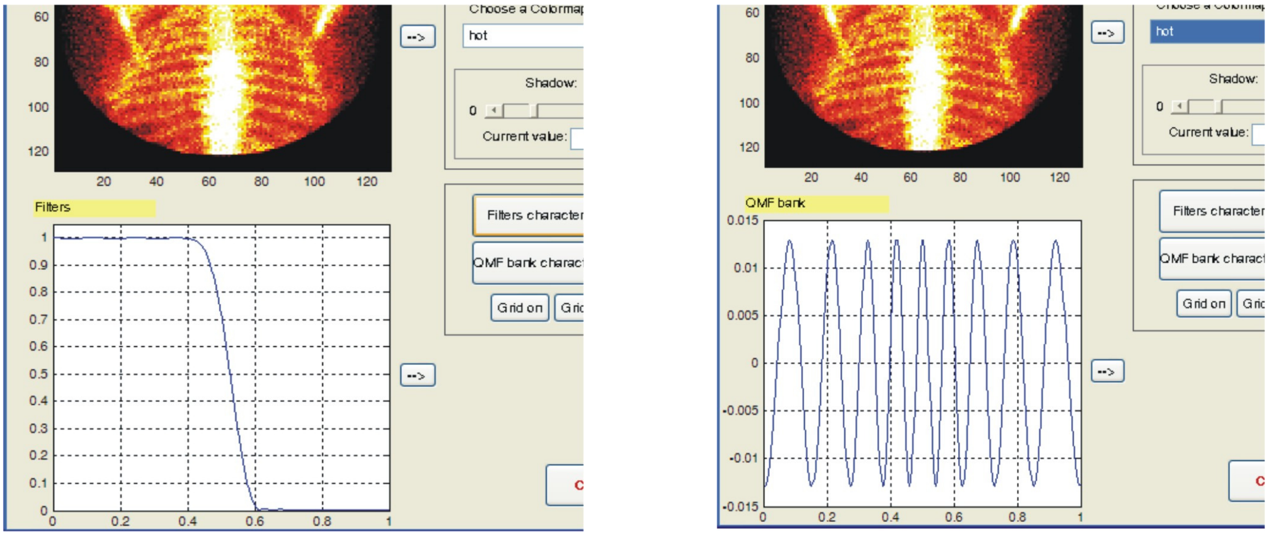


Fig. 2. The QMF bank and prototype filter characteristics

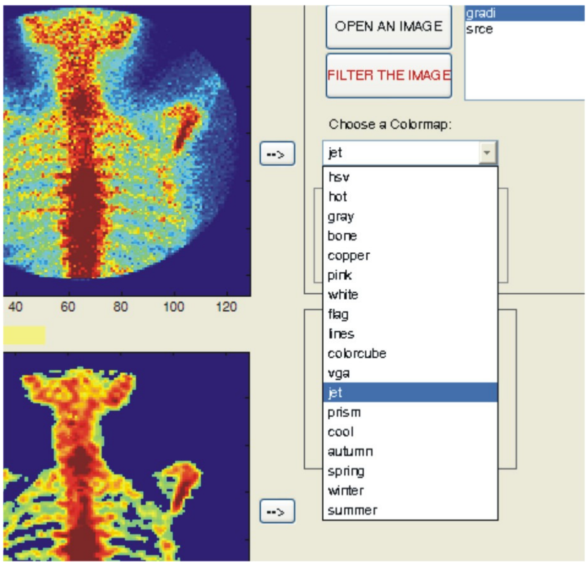


Fig. 3. Choosing a colormap

By clicking the GUI buttons labelled 'FILTER CHARACTERISTICS' or 'QMF BANK CHARACTERISTICS' (9), the prototype filter characteristics or QMF bank characteristics appear in the bottom-left frame for results, respectively. This is shown in Fig. 2. By clicking the buttons labelled with arrow (3), the images are exported in a separated window for further processing or using in the Matlab command window or other application.

III. THE ALGORITHM THAT WORKS BEHIND

Chest region images contain quantum noise, which is highly dependent on the underlying light intensity pattern being imaged. For denoising purposes, it is often advantageous instead of working in the spatial (pixel) domain to work in a transform domain. One possible choice for image transform is the discrete wavelet transform (DWT) domain.

The Discrete Wavelet Transform (DWT) decomposes a signal into a set of orthogonal components describing the signal variation across the scale [1]. The orthogonal components are generated by dilations and translations of a prototype function ψ called *mother wavelet.*

$$\psi_{i,k}(t) = 2^{-i/2}\psi(t/2^i - k), \quad k,i \in Z \tag{1}$$

In analogy with other function expansions, a function f may be written for each discrete coordinate t as sum of a wavelet expansion up to certain scale J plus a residual term, that is:

$$f(t) = \sum_{j=1}^{J} \sum_{k=1}^{2^{-j}M} d_{jk}\psi_{jk}(t) + \sum_{k=1}^{2^{-J}M} c_{Jk}\phi_{Jk}(t) \tag{2}$$

Motivated by the DWT tendency to concentrate the energy of a signal into a small number of coefficients, while a large number of coefficients have low SNR, we modify the wavelet coefficients w_i according to:

$$\hat{w}_i = w_i \cdot h_i \tag{5}$$

where the shrinkage filter h_i corresponds to the "soft threshold" nonlinearity

$$h_i^{(\text{soft})} = \begin{cases} 1 - \dfrac{\tau \operatorname{sgn}(w_i)}{w_i}, & \text{if } |w_i| \ge \tau \\ 0, & \text{if } |w_i| < \tau \end{cases} \tag{7}$$

with τ a user-specified threshold level.

But, because the noise in the NM images is not additive white Gaussian, the noise level is not uniform throughout the image and hence uniform across all the wavelet coefficients. Therefore, a simple global noise threshold can not be determined independently on the signal and the wavelet-domain filtering based on a global threshold is inappropriate. Hence, for denoising this type of images we use the Anscombe (1948) variance-stabilizing transformation:

$$y_{i_1,i_2} = \sqrt{N_{i_1,i_2}} \quad (8)$$

As a result, the coefficients of the transformed images are characterized with a white Gaussian noise model [2] and we calculate the Donoho's level dependent threshold:

$$\tau_j = \sigma\sqrt{2\log N} \cdot 2^{(j-J)/2}, \quad j = 0,...,J \quad (9)$$

where σ is the noise standard deviation (to be estimated), J is the number of decomposition levels, j is the scale level and N is the total number of image pixels [2].

In addition, by applying nonlinear filtering to the wavelet coefficients, we discard the wavelet coefficients with a low SNR. Therefore the signal perfect reconstruction (PR) is not possible. Hence we decompose the images by using a near perfect reconstruction (NPR) QMF bank instead of using standard wavelet filters. With this we provide better frequency response and increase the design flexibility since the filter cut-off frequency, filter length and overall reconstruction error of the designed QMF bank are design parameters which has to be adjusted to each image. Software can display the prototype filter and QMF bank characteristics when buttons labelled with 9 in the GUI (Fig. 1) are clicked. In the algorithm the prototype filter is designed according to the method proposed in [3]

The whole algorithm utilized by the software is completely described in [5].

IV. Future Work

NM images are a specific kind of images and a process of separating signal from noise is a difficult task. In this Section we consider our future work to improve the efficiency of the software.

A NPR-QMF bank allows certain flexibility in the designing of its filters. So, the prototype filter cut-off frequency as a design parameter can be chosen taking into consideration the images 2D power spectrum characteristics [6]. The power spectrum of all the chest region NM images is concentrated on low frequencies in an ellipse like region. This means, the filter cut-off frequency always can be chosen such, so the images can be decomposed in one level only.

Further, the software can be upgraded to be able to process a set of dynamic NM images [4]. Since the dynamic images are consecutive, it can be expected that each image contains certain information that can be used for further pre-processing of its time neighbouring images [7]. This information is contained in the vectors formed by joining all the pixels from all the images that are located at same position. The vectors have a length equal to the total number of images in the images set. The vectors obtained from real NM images form a shape that contains isolated peaks as a result of the noise presence in the images.

In addition, the threshold for filtering the wavelet coefficients can be adjusted to the signal intensity since the noise is signal dependent. This means, instead using a global threshold, the bigger signal intensity the bigger threshold and vice versa.

V. Conclusion

We present a GUI for pre-processing chest region NM images. It works in the Matlab environment and uses Matlab user defined functions. The software processes the images in the DWT domain by using NPR-QMF filters with linear phase. The Anscombe variance-stabilizing transformation is applied over original images and as a result the transformed images are characterized with a white Gaussian noise model and Donoho's level dependent threshold is used.

References

[1] Gilbert Strang and Truong Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, 1996;

[2] D. L. Donoho, "Wavelet Thresholding and W.V.D.: A 10-minute Tour", *Int. Conf. on Wavelets and Applications*, Toulouse, France, Jun. 1992;

[3] Sofija Bogdanova, Mitko Kostov, and Momcilo Bogdanov, “Design of QMF Banks with Reduced Number of Iterations”, *IEEE Int. Conf. on Signal Processing, Application and Technology, ICSPAT ’99,* Orlando, USA, Nov. 1999;

[4] Cvetko D. Mitrovski and Mitko B. Kostov, “On the Preprocessing of Dynamic Nuclear Medicine Images”, *International Scientific Conference on Information, Communication and Energy Systems and Technologies ICEST 2002*, Nis, Serbia and Montenegro, 2002;

[5] Mitko B. Kostov and Cvetko D. Mitrovski, “On The Denoising Of Nuclear Medicine Chest Region Images”, *International Scientific And Applied Science Conference Electronics ET’2004*, Sozopol, Bulgaria, 22-24 Sept. 2004;

[6] Cvetko D. Mitrovski and Mitko B. Kostov, “NM Images Filtering Using NPR QMF Filters Dependent on the Images Spectrum”, *7th Int. Conf. on Telecommunications in Modern Satellite, Cable and Broadcasting Services*, Nis, Serbia and Montenegro, Sept. 2005;

[7] Cvetko D. Mitrovski and Mitko B. Kostov, “A Wavelet Approach on NM Images Filtering Using Adjacent Images Information”, *VII National Conference ETAI 2005*, Ohrid, Macedonia, Sept. 2005.

Denoising of ECG Signal Based on HHT

Lixin Song[1], Qi Wang[2] and Yujing Wang[3]

Abstract – **This paper introduces a new ECG signal denoising based on Hilbert-Huang transformation (HHT). In order to ensure the estimated level of Gaussian white noise based on threshold denoising is more valid, present the verification of the white noise using probability and statistics. About not white noise intrinsic mode function (IMF), the white noise level can be obtained by the product of the mean frequency and the energy density of each IMF. Point out instantaneous frequency of HHT is an important parameter for removing powerline interference and baseline wander from the ECG.**

Keyword – **HHT, ECG signal, Denoising, Verification of the white noise, Threshold.**

I. INTRODUCTION

Recording of electrocardiograms (ECGs) is an important basis for biomedical diagnosis. ECG signal is a weak low frequency non-stationary signal. When acquiring the data of ECG, because of many random interference such as electromyographical interference, powerline interference and baseline wander, ECG signal is embedded in the noise. The noise causes ECG signal to distort, and influences the validity of medicinal diagnosis. In order to obtain exact signal parameters, identify waveforms and diagnose diseases, we must restrain the noise, and improve the Signal-to-Noise (SNR). ECG signal denoising methods are mostly low pass filtering[1] and wavelet denoising[2]. Low pass filtering can filter high frequency interference, but simultaneously cause ECG signal to lose some information, moreover, it is difficult to restrain additive white noise whose bandwidth is the same as ECG signal. Wavelet transform threshold denoising that proposed by Donoho et al. is an effective method to remove the white noise[3][4]. But, there usually is the selection of wavelet base function when processing material signal. Furthermore, how to choose valid threshold is a problem.

Empirical mode decomposition (EMD) and HHT is a new method pioneered by Huang et al. in 1998 for non-linear and non-stationary time series analysis[5]. Signal is adaptively decomposed into several IMFs, and the IMFs have well-behaved Hilbert transforms. The input data is presented in an energy-frequency-time quantitative distribution by using this method, wavelet transform can't. Threshold denoising method, as described in wavelet analysis, is presented in some signal denoising articles based on EMD[6]. The key of threshold denoising is to choose valid threshold. The threshold must be just larger than the highest level of the noise exactly. So, some effective threshold methods for removing the white noise generate. When the white noise and useful signal component are mixed, with the useful signal component enhancing, existing threshold methods become useless in evidence. So, it is a practical significant subject to study the verification of the white noise series and to choose valid threshold when useful signal component is dominant.

At first, we study the verification of the white noise using probability and statistics. The choosing of the noise level about not white noise IMF is discussed here. Then based on HHT, confirm quantitatively the relation of instantaneous frequency and instantaneous amplitude, present a feasible denoising method by prior information of ECG signal frequency range, and carry this method into ECG signal denoising.

II. HHT

In the HHT, at first, we decompose a complex signal into a series of IMFs based on EMD. Each IMF must satisfy two requirements: (1) the number of extrema and the number of zero crossings are either equal or differ at most by one; (2) at any point, the mean value of the upper envelope defined by the local maxima and the lower envelope that defined by the local minima is zero. The sifting process of the signal involves the following steps: calculate the mean value m (t) of the upper and the lower envelope by using the local maxima and the local minima, then the difference h between x(t) and m(t) is:

$$\mathrm{h} = x(t) - m(t) \tag{1}$$

Set h as new x(t) and repeat above step until h satisfies the condition of IMF, written as:

$$c_1 = h \tag{2}$$

Define c_1 as the 1th IMF, and do:

$$x(t) - c_1 = \mathrm{r} \tag{3}$$

Set r as new x(t), repeat above steps and then obtain the 2th IMF c_2 ,the 3th IMF c_3 …, until c_n or r satisfies the stopping criterion. Then x(t) can be decomposed into:

$$x(t) = \sum_{i=1}^{N} c_i + r \tag{4}$$

[1]Lixin Song and [2]Qi Wang are with Department of Automation Measurement and Control, Harbin Institute of Technology, Harbin, China, Email: lixins99@yahoo.com.cn
[3]Yujing Wang is with College of Electrical and Electronic Engineering, Harbin university of science and technology, Harbin, China,
Email: mirrorwyj@sina.com

For each IMF, the Hilbert transform is defined as:

$$\mathrm{H[c(t)]} = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{c(\tau)}{t-\tau}\mathrm{d}\tau \tag{5}$$

Construct the analytic signal:

$$z(t) = c(t) + jH[c(t)] = a(t)e^{j\Phi(t)} \tag{6}$$

The amplitude function:

$$a(t) = \sqrt{c^2(t) + H^2[c(t)]} \tag{7}$$

And the phase function:

$$\Phi(t) = \arctan\frac{H[c(t)]}{c(t)} \tag{8}$$

Instantaneous frequency of the IMF component is defined as:

$$\mathrm{f}(t) = \frac{1}{2\pi}\frac{\mathrm{d}\theta(\mathrm{t})}{\mathrm{dt}} \tag{9}$$

After applying Hilbert transform to each IMF:

$$s(t) = \mathrm{Re}\sum_{i=1}^{N} a_i(t)e^{j\Phi_i(t)} = \mathrm{Re}\sum_{i=1}^{N} a_i(t)e^{j\int\omega_i(t)dt} \tag{10}$$

The Hilbert spectrum is defined as:

$$H(\omega,t) = \mathrm{Re}\sum_{i=1}^{N} a_i(t)e^{j\int\omega(t)dt_i} \tag{11}$$

The marginal spectrum is defined as:

$$h(\omega) = \int_0^T H(\omega,t)dt \tag{12}$$

III. ECG SIGNAL DENOISING USING HHT

The frequency range of ECG signal is 0.05~100Hz, but 90% of ECG spectrum energy concentrates on 0.25~45Hz[7]. P, Q, R, S, and T waveforms are mainly characters of ECG signal analysis. The interference signals are: powerline interference and opposite harmonic interference; baseline wander whose frequency range is 0.15~0.3Hz; electro-myographical interference whose frequency range is 5~2KHz and spectrum characteristics obeys approximate zero-mean Gaussian white noise.

ECG signal is decomposed into high-low frequency IMFs based on HHT. We can obtain instantaneous frequency and amplitude of each IMF, then, the quantitative energy-frequency-time distribution of the signal is obvious. Formally, signal Hilbert-Huang marginal spectrum is very similar as Fourier transform, so, ECG signal dynamic denoising becomes possible. Fig. 1 is the contrast of ECG signal marginal spectrum and Fourier transform. The marginal spectrum hasn't the problem of leakage of spectrum as Fourier transform, and its principal component spectrum is more prominent.

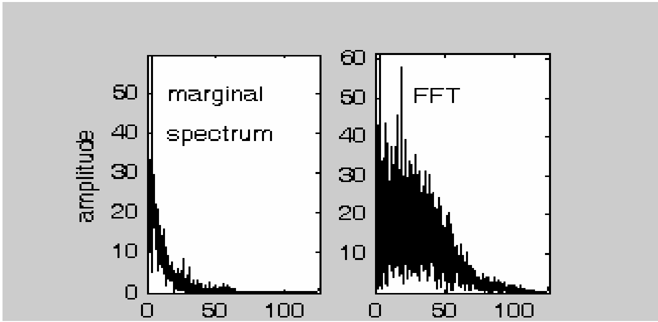


Fig. 1. Marginal spectrum and Fourier spectrum

A. Problems of EMD threshold Denoising

EMD-based Threshold denoising[6] is proposed. Its basic idea is decomposing a given signal x(t) into N IMFs based on EMD, Choose a appropriate threshold for each IMF, and cut c_i to $\hat{c}_i$ by this threshold, then reconstruct the signal by using EMD.

$$\hat{x}(t) = \sum_{i=1}^{N}\hat{c}_i + r \tag{13}$$

The threshold, aiming to removing Gaussian white noise in wavelet denoising, proposed by Donoho et al. is defined as:

$$\tau_i = \hat{\sigma}_i\sqrt{2\log(N)} \tag{14}$$

$$\hat{\sigma}_i = MAD_i / 0.6745 \tag{15}$$

Where $\hat{\sigma}_i$ is the noise level of the ith IMF. MAD_i represents the absolute median deviation of the ith IMF and is defined as:

$$MAD_i = Median\{|c_i(t) - Median\{c_i(t)\}|\} \tag{16}$$

the estimated c_i of the ith IMF is:

$$\hat{c}_i(t) = \begin{cases} sign(c_i(t))\ (|c_i(t) - \tau_i|) & \text{if } |c_i(t)| \ge \tau_i \\ 0 & \text{if } |c_i(t)| < \tau_i \end{cases} \tag{17}$$

Because the noise and useful signal component are mixed, it becomes difficult to separate the noise from each IMF component. Actually, for high frequency IMFs, including less signal component, consequently, Eq. (15) can be applied to them for calculating noise variance. Because the threshold selection method above is more suitable for Gaussian white noise, in order to ensure the result of Eq. (16) is valid, it is necessary to verify whether the series is the white noise.

Decomposing Gaussian white noise using EMD, each IMF also is Gaussian distribution, and the energy will reduce following N's increase. In theory, EMD decomposition is orthogonal, and the white noise should still be the white noise after orthogonal decomposition. In fact, due to the problem of calculating the mean value of the upper and lower envelopes, EMD is only orthogonal decomposition approximately. The trend of the white noise after approximate orthogonal decomposition is the white noise. All of these offer basis for the white noise verification of IMF.

B. Verification of the White Noise [8]

Verifying to ensure whether the series satisfies the null hypothesis, namely:

The null hypothesis $H_0:\{x(i)\}$ is independent white noise, otherwise hypothesis H_1 is correlative series.

Assume the estimated value of self-correlation coefficient as $\{\hat{\rho}_k$,k=1,2,...,$m\}$,when N is large enough, the distribution of $\sqrt{N}(\hat{\rho}_1,\hat{\rho}_2,\cdots,\hat{\rho}_m)$ obeys approximately m-dimensional standard normal distribution .

$N(\hat{\rho}_1{}^2+\hat{\rho}_2{}^2+\cdots+\hat{\rho}_m{}^2)$ obeys approximately χ^2 distribution .Since under the null hypothesis H_0

$$\hat{\rho}_1{}^2+\hat{\rho}_2{}^2+\cdots+\hat{\rho}_m{}^2=0 \tag{18}$$

Under the important level α =0.01, judging whether

$$N(\rho_1^2+\rho_2^2+\cdots\rho_m^2)<\lambda_a \tag{19}$$

is tenable. Where λ_a is critical value of χ^2 distribution with the prominent level α and m degrees of freedom.

C. Estimation of the White Noise Energy in EMD

While IMF component is dominated by useful signal, Eq. (15) can not be applied to estimating the noise energy directly. Gaussian white noise energy is defined as[9]:

$$E_i\overline{T}_i=const \tag{20}$$

Where E_i is the noise energy density of the ith IMF, $\overline{T}_i$ is the mean period of the ith IMF. The mean period can be obtained by averaging the instantaneous frequency of Hilbert transform of current IMF. Based on the noise energy and the mean period that estimated by high frequency IMF, we can estimate the white noise energy of current jth IMF:

$$E_j=const/\overline{T}_j \tag{21}$$

D. Removal of Powerline Interference and Baseline Wander

HHT can obtain instantaneous frequency of each IMF, then, before reconstructing, set signal amplitude whose frequency is close to 50Hz to zero. So, powerline interference can be removed easily by HHT.

About baseline wander, HHT is also an effective method, the low frequency IMF, whose instantaneous frequency is less than 0.3Hz, isn't involved in signal reconstruction process.

IV. EXPERIMENT RESULTS AND ANALYSIS

In order to test the validity of the denoising method above. In experiments, we add Gaussian noise (electromyographical interference), powerline interference, and baseline wander to free noise ECG signal. The correlation coefficients of noisy signal and original signal, denoised signal and original signal are given respectively in Table I.

TABLE I
DENOISING RESULTS

SNR (original signal power and the noise)	5dB	10dB
Correlation Coefficient (noisy signal and original signal)	0.278	0.279
Correlation coefficient (denoised signal and original signal)	0.949	0.958

In Table I, SNR is the ratio of original signal power and the noise (Gaussian noise and powerline interference). Moreover, we experiment many times using real signals from standard MIT/BIH ECG signal database. Experiment results prove that the noise influence of ECG signal can be well removed based on HHT above. The waveform contrast of original signal, noisy signal, and denoised signal using HHT above is given in Fig. 2 (SNR is 5dB).

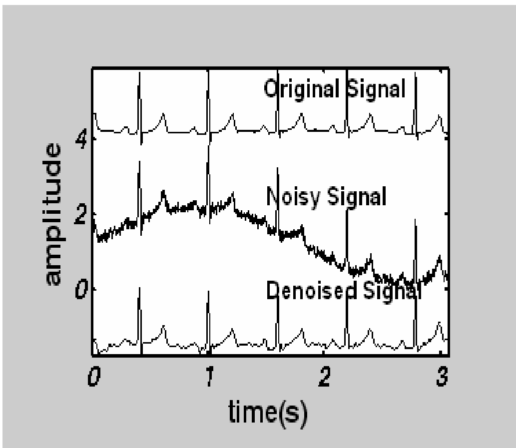


Fig. 2. Contrast of original signal, noisy signal and denoised signal

V. CONCLUSIONS

Through theoretic analysis and large numbers of experiments, conclusions are as follows:

A. Verification of the white noise is necessary for IMFs. It is the base of estimating the white noise variance correctly.
B. About not white noise IMFs, the white noise level can be obtained by the product, which obtained by the energy and mean period of high frequency IMF, and the mean period of current IMF.
C. Instantaneous frequency of each IMF is dependable basis for ECG denoising. HHT is an effective method for ECG signal denoising.

REFERENCES

[1] Van Alste, A. and Schilden, T.S. "Removal of baseline wander and powerline interference from the ECG by an efficient FIR filter with reduced number of taps". *IEEE Trans. on Biomed.Eng.,vol.* 32, no.12**,** pp. 1052-1060, 1985.

[2] S. Nibhanupudi, R. Youssif, and C. Purdy **"**Data-specific Signal Denoising Using Wavelets, with Applications to ECG Data" The 47h IEEE International Midwest Symposium on *Circuits* and Systems vol.3, pp. 219-222.

[3] S.Mallat. "A wavelet Tour of Signal Processing " Academic Press 1998.

[4] D.L. Donoho, "De-noising by soft-thresholding". IEEE Trans. Inform. *Theory, vol. 41, no. 3, pp. 613-627, 1995.*

[5] Huang, N. E, Shen, Z, Long, S. R, Wu, M. C, Shih, E. H, Zheng, Q., Yen, N.-C., Tung, C. C.& Liu, H. H. "The empirical mode decomposition method and the Hilbert spectrum for non-stationary time series analysis". *Proc. R. Soc. Lond.* A454, pp.903–995, 1998.

[6] A.O.Boudraa, J.C.Cexus,and Z.Saidi "EMD-Based Signal Noise Reduction". International Journal of Signal Processing. *Vol. 1, pp.33~36* . 2004.

[7] Xiwu Cao and Qinkal Deng "Frequency analysis on the ECG waveform". Chinese journal of medical physics Vol.18, no.1, pp.46-48. 2001.

[8] Shujing Zhang and Lixin Qi "Analysis of Time Series" Beijing. TsingHua 2003.

[9] Zhaohua Wu and Norden E. Huang "A study of the characteristics of white noise using the empirical mode decomposition method". *Proc. R. Soc. Lond. A 454, pp. 1597–1611, 2004.*

An investigation on the experimental measurement of conduction velocity of the nerve signals in the arm

Dimiter Tz. Dimitrov

Abstrac - **It's well clear that our nervous system is composed by excitable cells called neurons.These cells produce bioelectrical potentials as a result of electrochemical activity between the inside and outside of the cell membrane, thus transmitting the electrical signals and allowing the nervous impulses to go through our body. An investigation on the experimental measurement of conduction velocity of the nerve signals in the arm is presented in the paper. The obtained results are very important not only for the process of medical diagnostic, but for process of design of robot system.**

Keyword – **nervous signals, conduction velocity, experimental measurement**

I. INTRODUCTION

When a single nerve cell is stimulated, the amplitude of the signal does not affect the response. However, when the current applied to the nerve is increased, a higher number of neurons are stimulated, and so a larger number of muscle fibers are activated [1]. The way in which the number of pulses affects the response can be explained by two factors: the threshold and the refractory period [2].

Because in order to have an action potential the excitation value must exceed a threshold value of about 15 mV higher than the resting potential, for any value above the threshold level, the action potential will have the same value[1],[3].

Also the cells have a refractory period,that limits the frequency of consecutive action potentials, so no matter how strong the applied stimulus is, there will only be one response.

So for example if a stimulus of 150 mA in just one pulse is provided, the cells where only activated once with that signal. But when the same current is applied in five pulses, because the time interval between the pulses was higher than the refractory period of the cells, they were activated five times, thus making the nervous response stronger [4].

II. EXPERIMENTAL MEASUREMENT OF CONDUCTION VELOCITY OF THE NERVE SIGNALS IN THE ARM

The idea is to detect the onset of the stimulation on the stimulation point, and measure the time for the response at the target muscle. This is done for two different stimulation points, as will be described below. The stimulated nerve is the ulnar nerve. Two stimulation points would be designated, and the electrodes (AgCl) would be placed. The first stimulation point would be just below the elbow, the nerve passes beside the bone, and can easily be felt by the uncomfortable feeling when its pressed. The second point is near the wrist, again by the bone. The measurement electrodes would be attached to the palm, below the little finger, so that the activity on the 4th and 5th fingers can be detected. When the stimulation pulse is given to the skin, it would be conducted between electrodes as shown in Fig. 1.

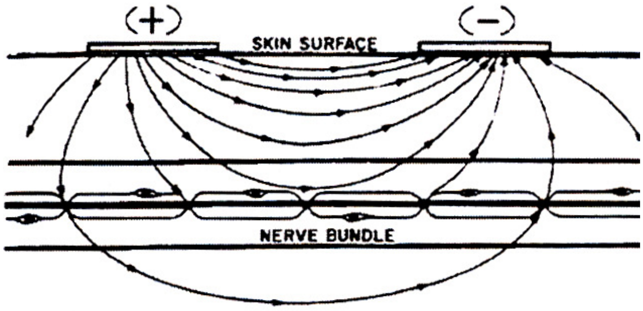


Fig. 1

So, the higher the stimulation is, the better response we would get (too high stimulation may be painful, care is needed).The Daq(data acquisition) channel to be triggered can be connected to electrodes nearby the stimulating electrodes, to recognize the stimulation pulse. The Daq program would be activated, and while it would be waiting for a trigger to occur at the stimulating side. The pulse(s) would be given by manually triggering the Transcranial Stimulator. The Daqboard (data acquisition board) would automatically record 300 values at a frequency of 10kHz for each channel and graph them.

The stimulation and response can be seen on the same graph, amplitude on the Y axis, and time on the x axis. By using these graph curves, the times between the first rise of the stimulation, and the beginning of the first distinguishable muscle activity can be recorded. If two different stimulation regions are used, it would be possible to obtain two graphs, thus giving us two different time measurements. These recorded times, say Atl and At2 can be subtracted from each other. The resulting time would give the time the pulse takes to travel between the two designated stimulation areas. The distance between these areas can be measured as the distance between the midpoints of the electrodes, namely, x. As the formula says: $v = x/t$, it's possible to get the nerve conduction velocity. There are two possibilities to measure the conduction velocity. Either measuring the action potential of the nerve, or to detecting the excitation of the target muscle. It is not easy to measure the potential of the nerve who is isolated from the skin surface, so it would be better to use the second method. The second method is much easier since the contracting muscles cause a detectable current flow in the surface of the

Dimiter Tz. Dimitrov is from Faculty of Communication Technique and Technologies, Technical University of Sofia, Kliment Ohridsky 8,1000 Sofia, Bulgaria, E-mail: dcd@tu-sofia.bg

skin. If only one measurement is used, it would measure the time difference between the activation and the contraction of the muscle. This time difference would not just include the propagation of the nerve pulse, but also other phenomena, related to the contraction of the muscle. If the measurement is from to different points it would be possible to eliminate the time delay due those phenomena. A sketch of the measurement points can be found on Fig 2.

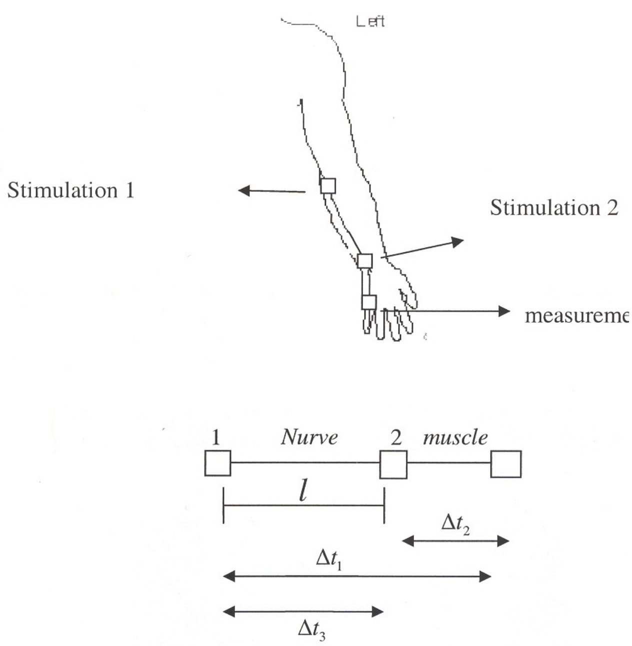


Fig. 2

If the first time interval is measured and afterwards time interval it's possible to find time interval by subtracting:

$$\Delta t_3 = \Delta t_1 - \Delta t_2 \quad (1)$$

It's possible to find the time Al3 that is needed for the nerve pulse to propagate from electrode one to electrode two. In this way only the nerve propagation is taken in account.

The nerve pulse velocity can be easily calculate using the distance between the electrodes one and two:

$$v = \frac{l}{\Delta t_3} \quad (2)$$

In practice a pulse train of 10 kHz is used, this means a 100µs per pulse. Results or given by:

$$\Delta t_1 = 13{,}2ms \quad (3)$$

$$\Delta t_2 = 10ms \quad (4)$$

This gives for the velocity:

$$v = \frac{0{,}21m}{3{,}2.10^{-3}s} = 65{,}6ms^{-1} \quad (5)$$

According to the reference the average nerve velocity is about 55 to 70 meter per second [4]. So, the obtained experimental result is acceptable.

III. Experimental System for Measurement of Conduction Velocity of the Nerve Signals in the Arm

It's clear that a preliminary muscle stimulation should be provided before the above mentioned measurement of impulse velocity. Apart from thinking and reflexive movements, the muscles can also be stimulated by external magnetic and/or electric fields, as these simulate what happens when the brain sends a signal commanding the muscles to move. The measurement schematics (Fig.3) includes: Electrical Stimulator, Bioamplifier (LNA), DaqBoard and PC.

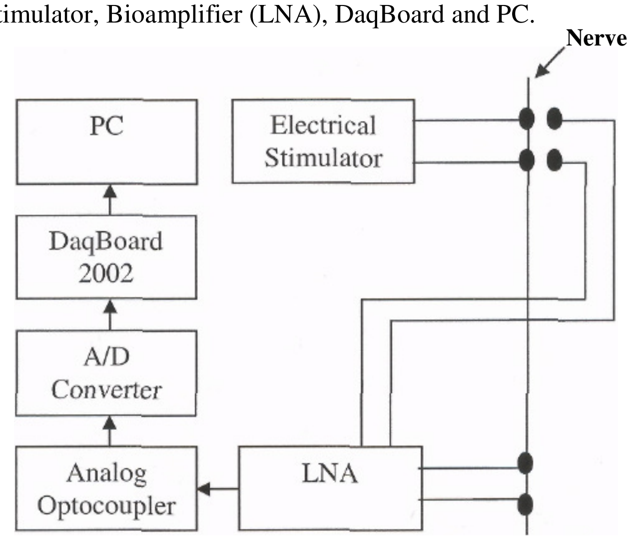


Fig. 3

A *Electrical Stimulator*

The stimulator produces a 50 /is square pulse at variable amplitude. The stimulation can also be done as a pulse train, having up to 5 pulses in one sequence. The current amplitude can be changed from 30 to 990 mA, in incremental steps of 30, by using the rotational knob.

To avoid confusion: the amplitude of the stimulation does not affect the response. That is true when stimulating single nerve fibers. When stimulating an arm, as we are doing, a higher current enables the stimulation of a higher number of neurons, thus affecting a larger number of muscle fibers. The additive effect of stimulation also plays a role, so that: A one pulse with 90 mA stimulation has more effect than 1 pulse, 30 mA stimulation. A one pulse with 90 mA stimulation has less effect than 5 pulses, 90 mA stimulation. All these parameters can be adjusted using the control buttons, and manual triggering is used in the experiment.

B. *Bioamplifier (LNA)*

The bioelectrical signals to be measured range between 20 /iV to 300 /xV, which are relatively low. The difficulty in measuring these values lies in the system's susceptibility to noise, artifacts and the 50 Hz interference. The accuracy of the measurement depends on the susceptibility to noise, so it's necessary to take care on the signal to noise ratio. To obtain the optimum measurement, the amplifier used is a differential input, high Common Mode Rejection Ratio (CMRR), variable gain amplifier with a frequency response set for the ECG measurements, which is adequate for the purpose. The differential input type is needed since the signals measured are the potential differences between two points of the body. The amplifier is battery operated, at +/- 15 V and +5V. Two channels of the amplifier are used. They are connected to an isolation amplifier utilizing optical isolation. The gain can be adjusted using the parallel port of the PC, two channels at one. The parallel port has 8 output pins, which can be used two represent a 2 digit hexadecimal number.

C. *DaqBoard (Data acquisition Board)*

DaqBoard is located between the amplifier and PC (it is installed to a PCI bus). The measurements coming from the amplifier are fed into the DaqBoard – data acquisition board. This card is installed to a PCI bus in the PC. The card allows Windows to sample data at a high rate without loss, in real time.

D. *PC*

The package MATLAB is used to plot the stimulation and the answer (x axis - time in samples, y axes - amplitude).

E. *Common Mode Rejection Ratio*

The measured signals have very low amplitudes. So, it is necessary to amplify these signals, before to continue with the signal processing. This is done by a normal instrumentation amplifier. This has the advantage (to other amplifiers) that many influences from other sources (e.g. power supply, mobile phones ...) were filtered out (the instrumentation - should - amplify only the differential potential between the two wires from the bipolar electrodes). The second advantage is the very high input resistance and which provides a very good sensitivity.

In reality it isn't true, that the amplifier amplifies only the differential signal but this amplification (which is called common mode gain) should be as small as possible.

More important is the relation between the differential gain and the common mode gain. This relation is described by the CMRR - Common Mode Rejection Ratio which is defined as:

CMRR = Adiff / Acorn[dB]

Isolation of amplifier's output

The Low Noise Amplifier is fully isolated from the PC (no electrical connection to the DaqBoard or to the ground-potential). So it is impossible to get impressed any current from outside, which could be dangerous for the test person, especially when the power supply from the PC has a connection to the electrodes. The isolation is realized by 2 components:

First: The Low Noise Amplifier has his own power supply (+-15V, +5V), which is provided by normal 9V batteries and a Voltage stabilization circuit.

Second: The Signals are transmitted with optocouplers, which separates it completely from the potential of the PC.

F. *A/D Conversion*

The A/D converter is a very important and very sensitive part of the whole measurement. It is located after the optocouplers There are many different types of A/D-converters, every whit his advantages and disadvantages. The conversion time and of course the prize are important for the described investigation. One type makes the conversion in one step (2 ^n comparators are required) or step by step (successive approximation, using only one comparator). It's necessary to have a reference voltage of very high quality. It's clear that the ideal converter isn't able to convert the signal without a failure, because the signal is quantified (in the case of resolution of 8 bits it's only possible to get 256 different solutions for the input signals, but the input signals can have infinite different values. Therefore the resolution should be flexible.

G. *Patient's electrodes*

The most important things required for patient's electrode using are: symmetric stimulation in order the electrode does not get polarized, decrease the motion artefact, and good Connection between the skin and the electrode. Usually the physicians stretch the skin or they clean it with alcohol, or they use some gel in order to decrease the impedance and the noise.

IV. CONCLUSION

1. A simple experimental method and system for measurement of conduction velocity of the nerve signals in the arm is described in the paper. The method and system can be used for measurement of conduction velocity of nerve signals in other parts of the human body.
2. The obtained experimental results on the base of described method and system can be used successfully in the process of medical diagnostic and in the process of engineering design of medical information systems, robot systems, etc.
3. The described experimental method and system can be used easy in the process of practical engineering education, also

REFERENCES

[1] http://www.medicine.mcgill.ca/phvsio/vlabonline/cnslab/cnstheory.htm

[2] http://www.leleemg.com/ .

[3] http://www.surrey.ac.uk/MME/Research/BioMed/de page/fesintr/

[4] http ://www.adinstruments.com/cust/teaching/phvsioex .html

An Electrical Model of Propagation of Signals in the Nerve Axon

Dimiter Tz. Dimitrov

Abstract - **Nowadays with the projects going on in prosthesis technology becomes more important. It's necessary to use Electromyogram (EMG) in the process of replacing a lost limb. The nerve impulses to be sent to the ghost of that limb can be detected and this then transferred to the prosthesis.**

The present lab work is designed to give the student knowledge about the nerve system and the conduction process. This knowledge will be utilized during the measurement of nerve propagation speed, and the display and evaluation of the resulting data. While doing the lab, the student will also get practice in various aspects of general signal acquisition.

INTRODUCTION

Since the experiment focuses on detecting nerve impulses, it would be convenient to explain the function of the nerve system, and later, go through the single nerve impulse. The nerve cell may be divided on the basis of its structure and function into three
main parts[1]:

- the cell "body", also called the soma
- numerous short processes of the "soma", called the "dendrites"
- the single long nerve fiber, the "axon"

The short processes of the cell body, the dendrites, receive impulses from other cells and transfer them to the cell body. The long nerve fiber, the "axon", transfers the signal from the cell body to another nerve or to a muscle cell. Mammalian axons are usually about 1-20 µm in diameter. Some axons in larger animals may be several meters in length. The axon may be covered with an insulating layer called the "myelin sheath", which is formed by "Schwann cells" (Fig. 1).

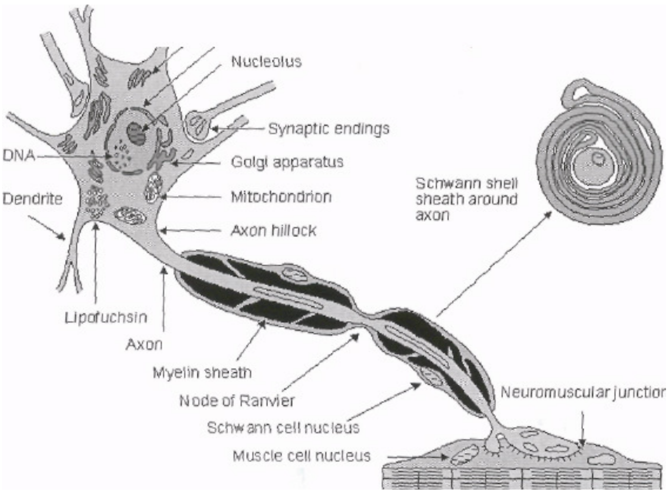


Fig. 1 - The major components of neuron

Dimiter Tz. Dimitrov is from Faculty of Communication Technique and Technologies, Technical University of Sofia, Kliment Ohridsky 8,1000 Sofia, Bulgaria, E-mail: dcd@tu-sofia.bg

The membrane voltage Vm(transmembrane voltage) of an excitable cell is defined as the potential at the inner surface C_J relative to that at the outer C_i surface of the membrane, i.e. $Vm = Cj - C_i$. If a nerve cell is stimulated, the transmembrane voltage necessarily changes. After stimulation the membrane voltage returns to its original resting value. If the excitatory stimulus is strong enough, the transmembrane potential reaches the threshold, and the membrane produces a characteristic electric impulse, the nerve impulse. This potential response follows a characteristic form regardless of the strength of the transthreshold stimulus. The response of the membrane to various stimuli of changing strength can be find on Fig. 2

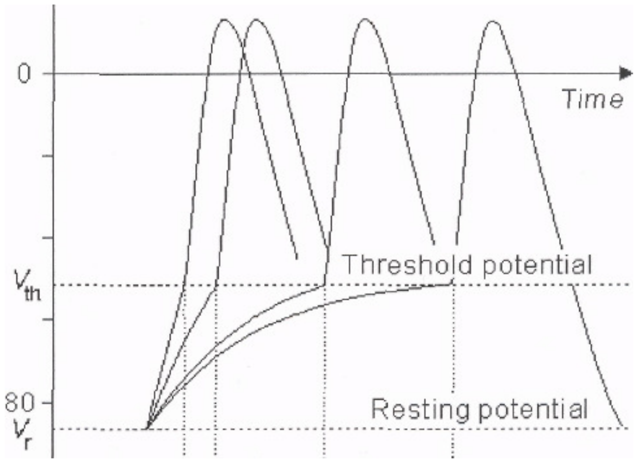


Fig. 2 - The response of the membrane to various stimuli of changing strength

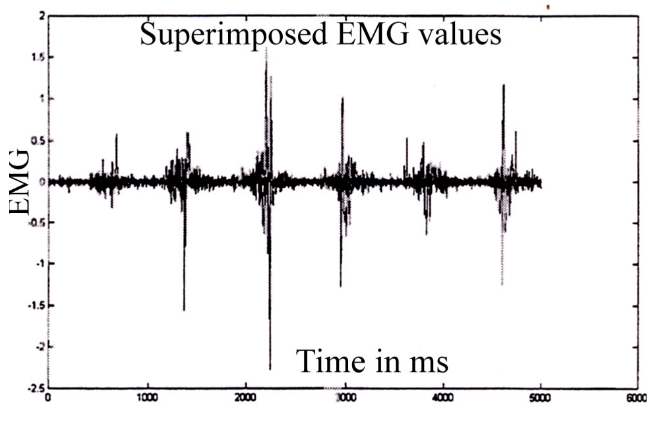


Fig. 3 EMG of a resting biceps, stimulated occasionally, and the contraction moments.

When measuring, if there is no forced contraction of the muscle, a signal resembling white noise is seen. When the target muscle is stimulated, in this case with the use of a electrical stimulator, the seen signal should be more distinctive[3].

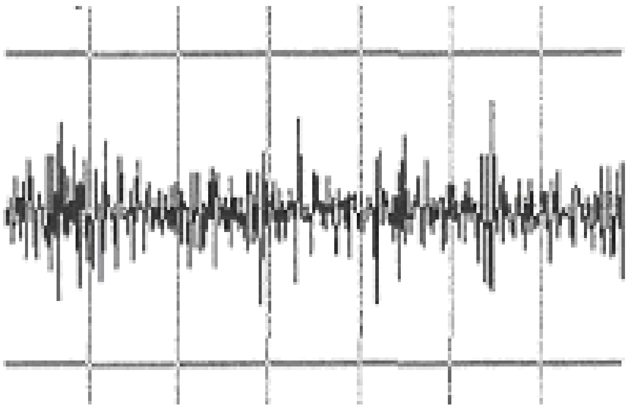


Fig. 4a

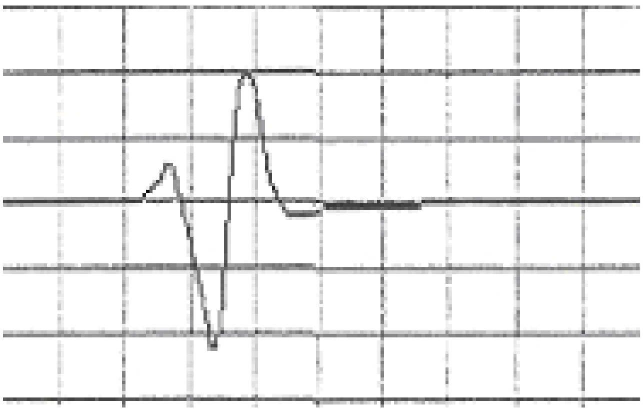


Fig. 4b

Fig. 4. A muscle group at constant contraction (4a), and the EMG of one contraction of a muscle fiber (4b).

The spikes occur when the right arm makes a contraction can be seen on Fig.3. The activity in the muscle can easily be seen, compared to the times of rest. The first one (Fig.4.a) is a contracted muscle, which seems like nose, having many components in different frequencies, seemingly random[4]. The picture (Fig.4b) shows the EMG taken from a single fiber, where a clear signal can be obtained.fig.4a Fig.4b It’s well known[2] that the concentration of sodium ions (Na+) is about 10 times higher outside the membrane than inside, whereas the concentration of the potassium (K+) ions is about 30 times higher inside as compared to outside. When the membrane is stimulated so that the transmembrane potential rises about 20 mV and reaches the threshold - that is, when the membrane voltage changes from -70 mV to about -50 mV (these are illustrative and common numerical values) - the sodium and potassium ionic permeabilities of the membrane change. The sodium ion permeability increases very rapidly at first, allowing sodium ions to flow from outside to inside, making the inside more positive. The inside reaches a potential of about +20 mV. After that, the more slowly increasing potassium ion permeability allows potassium ions to flow from inside to outside, thus returning the intracellular potential to its resting value. The maximum excursion of the membrane voltage during activation is about

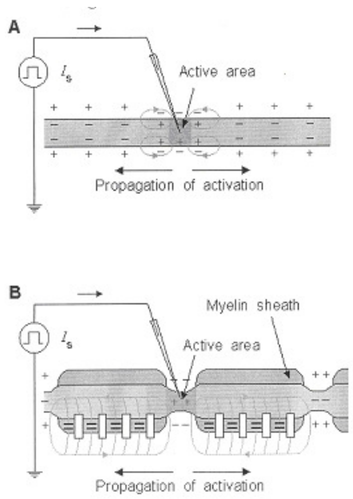


Fig.5
A - Continuous conduction of a nerve impulse in a nerve axon in an unmyelinated axon
B - Saltatory conduction of a nerve impulse in a nerve axon in a myelinated axon.

9

100 mV; the duration of the nerve impulse is around 1 ms. While at rest, following activation, the Na-K pump restores the ion concentrations inside and outside the membrane to their original values. The activation propagates in an axon as anunattenuated nerve impulse.[3]. The potential difference between excited and unexcited regions of an axon would cause small currents, now called “local circuit currents”, to flow between them in such a direction that they stimulate the unexcited region. Although excitatory inputs may be seen in the dendrites and/or soma, activation originates normally only in the soma. Activation in the form of the nerve impulse (action potential) is first seen in the root of the axon - the initial segment of the axon, often called the “axon hillock”. From there it propagates along the axon. If excitation is initiated artificially somewhere along the axon, propagation then takes place in both directions from the stimulus site. The conduction velocity depends on the electric properties and the geometry of the axon. An important physical property of the membrane is the change in sodium conductance due to activation.[4] The higher the maximum value achieved by the sodium conductance, the higher the maximum value of the sodium ion current and the higher the rate of change in the membrane voltage. The result is a higher gradient of voltage, increased local currents, faster excitation, and increased conduction velocity. The decrease in the threshold potential facilitates the triggering of the activation process. The velocity also depends on the resistivity of the medium inside and outside the membrane since these also affect the depolarization time constant [4]. The smaller the resistance, the smaller the time constant and the faster the conduction velocity. The temperature greatly affects the time constant of the sodium conductance; a decrease in temperature decreases the conduction velocity. A myelinated axon (surrounded by the myelin sheath) can produce a nerve impulse only at the nodes of Ranvier. The fig.5 shows propagation of signals. In these axons the nerve impulse propagates from one node to another, as illustrated in Fig. 5B. Such a propagation is called saltatory conduction In unmyelinated axo the conduction is continuous and slower on Fig. 5A.

The membrane capacitance per unit length of a myelinated axon is much smaller than in an unmyelinated axon. Therefore, the myelin sheath increases the conduction velocity. The resistance of the axoplasm per unit length is inversely proportional to the cross-sectional area of the axon and thus to the square of the diameter. The membrane capacitance per unit length is directly proportional to the diameter. Because the time constant formed from the product controls the nodal transmembrane potential, It is reasonable to suppose that the velocity would be inversely proportional to the time constant. On this basis the conduction velocity of the myelinated axon should be directly proportional to the diameter of the axon. Some experimental investigations were done in the cases of different diameters of the axons (different parts of the bodies and different animals) The experimental results can be find on the fig.6 as points. The conduction velocity in myelinated axon has the approximate value shown:

$$v = 6d$$

where

v= velocity of the nerve impulse [m/s]

d= axon diameter [μm]

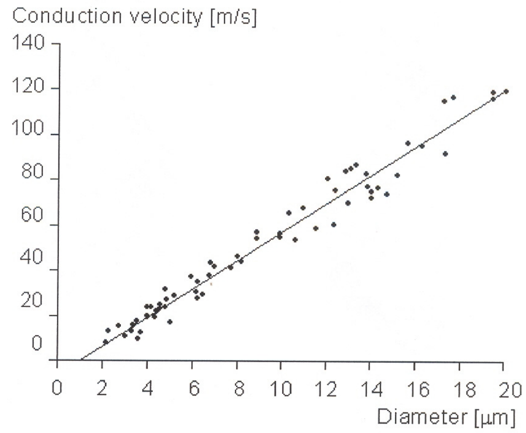


Figure 6 - Experimentally determined conduction velocity of a nerve ir in a mammalian myelinated axon as a function of the diameter.

Fig. 6. Experimentally determined conduction velocity of a nerve impulse in a mammalian myelinated axon as a function of the diameter.

CONCLUSION

1.The method for investigation of EMG, presented in the paper, is designed to give the student know-ledge about the nerve pulses and the conduction process.

2. An electrical model of propagation of signals in the nerve axon is described in the paper.

3. Experimentally determined conduction velocity of a nerve impulse in a mammalian myelinated axon as a function of the diameter was obtained in the process of investigation.

REFERENCES:

[1] http://www.lucs.lu.se/arthand/,

[2] http://uc.rutgers.edu/news/science/arthand2.html

[3] rhttp://www.medicine.mcgill.ca/physio/vlabonline/cnslab/cnstheory.htm1

[4] rhttp://web.media. mit.edu/~marrin/CIM.htm1

Digital Matched Filter

Slavcho D. Lishkov[1], Veska M. Georgieva[2]

Abstract – **In the paper are presented the principles of building digital matched filters (DMF), which are used by signal processing with background noise. Their perspective using in signal processing is appropriated by high selectivity, stability, precision of the characteristics and the reliability of the filters.**
On the base of the theory of digital matched filters, the students can made a simulation model for analyze and synthesis of digital matched filter. Some important characteristics can be analyzed: the spectrum, impulse response, correlation function, the influence of the sampling on the level of the sideband of the output signal and on the level of quantizing. The paper can be used in engineering education in studying these filters.

Keywords – **Digital matched filters, ambient noise, correlation functions, detection and identifying of signals, computer simulation.**

I. INTRODUCTION

It's described in the paper an idea of exercises for the course on Signals and Systems of students from faculty of communications, from faculty of computer systems and from faculty of electronics.
The digital matched filters are used by signal processing with background noise. Their perspective using in signal processing is appropriated by high selectivity, stability, precision of the characteristics and the reliability of the filters[1].
On the base of the theory of digital matched filters, the students can made a simulation model for analyze and synthesis of digital matched filter. The synthesis of the DMF can be realized on the base of the digital filter by giving the Amplitude/frequency characteristic and Phase/frequency characteristics. The DMF can be realized as FIR or IIR filters[2]. Some important characteristics can be analyzed: the spectrum, impulse response, correlation function, the influence of the sampling on the level of the sideband of the output signal and on the level of quantizing.

II. PROBLEM FORMULATION

The principles of building digital matched filter are not differed from the usual digital filter. The matched filter is linear filter, which provides maximal signal-to-noise ratio (SNR). Its impulse response can be presented with the following mathematical description, given in Eq.1.

$$g(l\Delta t) = k.U\left[(n_c - l)\Delta t\right] = k.U\left[(n_0 - l)\Delta t\right] \qquad (1)$$

where k is a constant; $n_0 = n_c = n_{\max}$ is a max delay of the signal in the output of the filter, which is needed to its physical realize; $n_c = \tau_c / \Delta t$; τ_c is the time of the signal; $l = (0,1,2,.........n_c - 1)$; Δt is the sampling time.

The frequency response of the digital matched filter is given in Eq.2.

$$H(j\omega) = \exp(-j\omega n_0)\sum_{l=0}^{\infty} g(l\Delta t)\exp(-j\omega l) = \\ = U(-j\omega)\exp(-j\omega n_0) \qquad (2)$$

According to principles of the matched filtration (Eq.3) [3]:

$$\left|H(j\omega)\right| = k\left|U(-j\omega)\right| = k\left|S^*(j\omega)\right| \qquad (3)$$

where $S^*(j\omega)$ is the complex conjugate $\Delta f = -1/2\Delta t \le \Delta f \le 1/2\Delta t$ ed spectrum of the signal.
The noise of the input of the digital matched filter is a background noise can be study as fortuitously stationary sequence an even distribution of the spectrum density $N_0/2$ in the frequency band $\Delta f = -1/2\Delta t \le \Delta f \le 1/2\Delta t$, dependent on sampling frequency. The correlation function of the digital noise in the output of the DMF is coincided with accuracy of constant multiplier on form of the output signal $y(k\Delta t)$, given in Eq.4.

$$\Psi_N(m\Delta t) = N_0/2 . y\left[(n_c - m)\Delta t\right] \qquad (4)$$

The Energy SNR is $q^2 = \dfrac{2E_0}{\sigma_N^2}$ and $\sigma_N^2 = \Psi(0) = N_0/2\, E_U$ is the mean power noise.

III. SYNTHESIS OF THE DMF

The synthesis of the DMF can be realized on the base of the digital filter by giving the Amplitude/frequency characteristic and Phase/frequency characteristics. These can be determinate from the amplitude and phase spectrum of the signal.
The simple case is on the base of the FIR filters, which can be defined in Eq.5 [4].

$$y(k\Delta t) = \sum_{l=0}^{n_c - 1} g(l\Delta t).U\left[(k-l)\Delta t\right] = \sum_{l=0}^{n_c - 1} C_l U(l) \qquad (5)$$

where $C_l = g(l\Delta t)$ is the digital impulse response.

[1]Slavcho D. Lishkov is with the Faculty of Communication, TU-Sofia, Kl.Ohridsky str.8, Sofia, Bulgaria.

[2]Veska M. Georgieva is with the Faculty of Communication, TU-Sofia, Kl.Ohridsky str.8, Sofia, Bulgaria, E-mail: vesg@tu-sofia.bg

To preserve the information about the amplitude and phase of the signal and to decrease the sampling frequency can be used the quadrature processing. The signal $u(t)$ can be presented in the complex form, given in Eq.6.

$$\dot{u}(t) = \dot{U}(t)\exp[j\omega t+\varphi_0] = \dot{U}(t)\cos[\omega t+\varphi_0] + j\dot{U}\ (t)\sin[\omega t+\varphi_0] \tag{6}$$

where $\dot{U}(t) = U_1(t)\cos\omega t + jU_2(t)\sin\omega t$;

$U(t) = \sqrt{U_1^2(t)+U_2^2(t)}$; $tg\varphi_0 = {U_2(t)}/{U_1(t)}$,

$U_1(t)$ and $U_2(t)$ are respectively the equiphase and the quadrature components of the signal. The signal can be decomposed to the quadrature components using the phase comparators.

The digital quadrature components of the signal can be marked respectively $U_1(k)$ and $U_2(k)$. The complex shape of the input signal and the complex shape of the impulse response can be described in Eq.7.

$$\dot{U}[k] = U_1[k]-U_2[k]\ ;\ \dot{g}[k] = g_1[k]+g_2[k] \tag{7}$$

The output signal of the DMF with the accuracy of constant multiplier ${\Delta t}/{2}$ is the following (Eq.8):

$$\dot{Y}[k] = y_1[k]+jy_2[k] = \sum_{l=0}^{n-1_c}\{U_1[k-l]-jU_2[k-l]\}\times\{g_1[n_c-l]+jg_2[n_c-l]\} \tag{8}$$

When $n_c-l=i$, for the quadrature components can be obtained (Eq.9):

$$y_1[k] = \sum_{i=1}^{n_c} g_1(i)U_1[k-(n_c-i)] + \sum_{i=1}^{n_c} g_2(i)U_2[k-(n_c-i)] = y_{11}(k)+y_{22}(k)$$

$$y_2[k] = \sum_{i=1}^{n_c} g_2(i)U_1[k-(n_c-i)] - \sum_{i=1}^{n_c} g_1(i)U_2[k-(n_c-i)] = y_{21}(k)+y_{12}(k) \tag{9}$$

This algorithm can be realized by using the following structural diagram (Fig. 1):

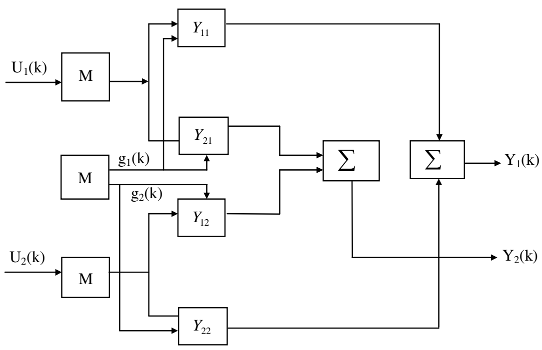


Fig. 1. Structural diagram of the algorithm

The quadrature channels can be realized on the base of the IIR filters. In this case the coefficients and the time of the processing are the same. The advantage of the qudrature processing is the decreasing of the sampling frequency. To realize the DMF on the base of the IIR filters, we can use the following correlation, given in Eq.10.

$$\vec{U} = \vec{U}_1+\vec{U}_2\ ;\ \vec{C} = \vec{C}_1+\vec{C}_2 \tag{10}$$

where $\vec{U}$ and $\vec{C}$ are respectively the input signal and the coefficients of the filter. For the output signal can be obtained (Eq.11):

$$\vec{y} = \vec{C}.\vec{U} = \vec{y}_1 + j\,\vec{y}_2\ ;$$

$$\vec{y}_1 = \vec{U}_1.\vec{C}_1 - \vec{U}_2.\vec{C}_2\ ;\ \vec{y}_2 = \vec{U}_2.\vec{C}_1 + \vec{U}_1.\vec{C}_2 \tag{11}$$

By the matched filtration can be realized the operation convolution of the signals in the time domain, what corresponds to the multiplication of their spectrums in the frequency domain. The spectrums can be obtained by Fast Fourier Transform (FFT).

In Fig.2 is given the DMF for processing in frequency domain.

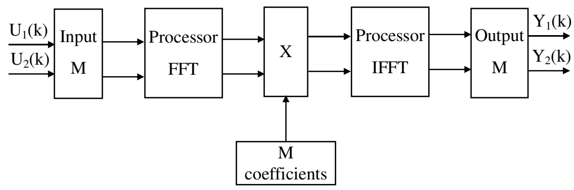


Fig. 2. Schema of DMF for processing in frequency domain

The effectiveness of the filtration is dependent on the sampling frequency and the level of the quantizing. It's known, that the sampling frequency $f_s = 2\Delta f_{SB}$, where Δf_{SB} is the sideband of the signal.

On the base of the theory the students can do the following tasks:

- To made a synthesis of the DMF as FIR or IIR filter
- To analyze the influence of the sampling frequency on the output signal
- To analyze the influence of the level of quantizing on the output signal

IV. CONCLUSION

In the paper are presented the principles of building digital matched filters (DMF), which are used by signal processing with background noise.

On the base of the theory of digital matched filters, the students can create a simulation model for analyze and synthesis of digital matched filter. As variant for simulation can be used MATLAB environment, respective the SIMULINK TOOLBOX. The synthesis of the DMF can be realized on the base of the digital filter, by giving the Amplitude/frequency characteristic and Phase/frequency characteristics. The DMF can be realized as FIR or IIR filters. The signal can be processed in time as in frequency domain.

To preserve the information about the amplitude and phase of the signal and to decrease the sampling frequency can be used the quadrature processing.

REFERENCES

[1] Тургин Г. Введение в теорию цифровых согласованных филтров, ТИНЕЗр тл64, №7, 1974

[2] Proakis John G., Digital Communications, Cole Publishing Company, 2000

[3] Тихонов В. Н. Статистический анализи синтез радиотехнических устройств и систем, М., Радиосвезь, 1991

[4] Иванов Р. Цифрова обработка на едномерни сигнали, Г., Алмаматер, 1999

Combined use of Meteorological Data from "METEOSAT" Satellite and Meteorological Radars

Uliana T. Petrova[1]

Abstract – **Radar images often show precipitation echoes in areas where precipitation does not occur (ground clutter). Until there will be better radar techniques, radar results must be verified with other techniques. One way is to compare them with METEOSAT data. A cloud mask can be calculated using IR image and a reference image of surface temperature. Radar echoes can only be present in areas where algorithm identifies clouds.**

Keywords – **meteorological satellite images, radar meteorological images, educational satellite data, ground cutter map.**

I. Introduction

Especially for agricultural and aeronautical meteorology, the knowledge on the cloud top highs (CTH) is very important. To get the CTH from METEOSAT image, it is necessary to use additional information like radio-probe measurements or model output. For this case, a method was developed to combine the cloud top temperature (METEOSAT B-format) with temperature-high data from the EM-Model. Since equipment cost for the detection of satellite and radar meteorological transmission is becoming affordable to universities, some educational institutions have been involved in applying of this method, which opens ways for more efficient interactions between students and computers.

Radar meteorological images are usually tied to a single local observation. In the same time this observation is part of corresponding satellite observation. The two images can be obtained together using appropriate computer systems. Such presentations are very important not only for research, but also for training students and professionals. Several meteorological workstations are set up in the major meteorological office of the Ministry of Agriculture and in some regional offices. The main office is in one of the research laboratories for students training at the Technical University in Sofia.

Several meteorological program packages are provided on these workstations to construct surface weather charts, significant weather maps and to display model output or satellite images.

Research is conducted in cooperation with the Department of Medientechnik of the Technical University in Ilmenau, Germany.

The satellite image processing program is capable of showing composite radar images from the corresponding regional radar stations, to combine radar and satellite images and to present loops with satellite images as well as radar images.

Radar composites are constructed from 5 radar stations and are available every 15 minutes at the spatial resolution of 4 km.

II. Verification of Radar Images with METEOSAT Images

A. Ground clutters

In some cases radar images show echoes where precipitation activity does not occur. Echoes are caused by ground clutters. Such misleading ground clutter echoes occur most frequently under clear sky situation in the morning hours when well– pronounced surface inversions exist (Fig. 1 shows an example). There are several possibilities to remove these effects:

a) Modify the "Ground Clutter Map" when radar information is preprocessed;

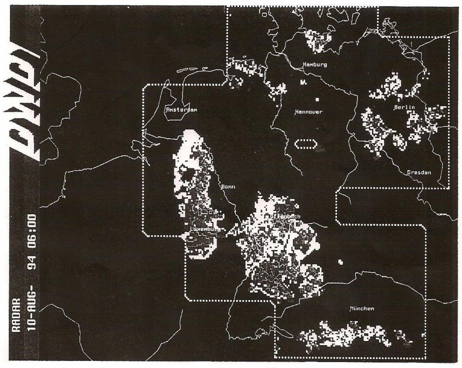
b) Use special procedure which analyses regions with ground clutters (in development);

c) Verify composite radar images against METEOSAT data.

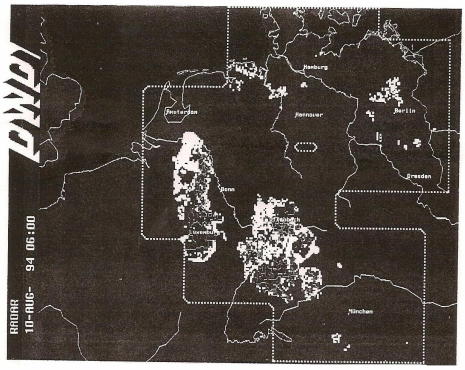
B. *Verification of radar composite images with METEOSAT data*

One way to correct radar composites is to compare radar data with METEOSAT images. A cloud mask from the IR data is calculated in order to perform automatic cloud detection. The cloud mask needs a thermal surface reference image. It is calculated on a daily basis out of the last ten 0400-UTC-IR images[1]. Each pixel of these reference images shows the minimum temperature at 4 UTC for this location. Alls pixels in the IR image, which are colder than the reference pixel, are considered a cloud. It is possible to calculate a cloud mask for every incoming IR – METEOSAT image. The user can change interactively parts of the reference images with a satellite image program. This is necessary when the last night happens to be colder than the previous nights thus locally warming up the reference. In this case too many pixels are identified as being cloudy. The cloud mask is also needed for the generation of specially processed images for TV presentation in the case of distance education. After composite radar picture enters the workstation, it is compared with the actual cloud mask (15 minutes max). Radar echoes are only displayed in those areas where the cloud mask detects clouds. When the cloud top temperature is above 0^0C, echoes intensity is reduced to the lowest

[1] Uliana T. Petrova is with the Faculty of Communications and Communications Technologies, Technical University, Kliment Ohridski 8, 1000, Sofia, Bulgaria, e-mail:uliana5@abv.bg



▲ Fig. 1: Radar composit without correction ▼ Fig. 2: Radar composit with correction



precipitation rate as more intensive precipitation activity can only be produced by clouds below freezing point (Fig. 2). Especially for TV presentation, the radar composite pictures ombined with clouds from METEOSAT images are available, also.

II. Calculation of CTH with Data from the European Model (EM)

A. *The method*

Information about the cloud top temperature can be taken from the METEOSAT's IR channel. For several applications, especially for agriculture and for aviation purposes, the high of clouds must be known. Cloud height can be determined if the surface temperature of the cloud and the current vertical temperature of the atmosphere are known. The first piece of information can be derived from METEOSAT satellite measurements and the second one from radio-probe observations, vertical satellite soundings, or from numerical weather prediction models. Only the last option offers widespread information on grid points which cover the whole satellite image area. The meso-scale "Europa Model" is a suitable tool to solve the task.

B. *Model data*

The "Europa Model" (EM) is part of a weather predicttion system, which also contains a global spectral model (GM) for large-scale predictions and a high-resolution meso-scale model [2].

The EM is working in the synoptic and meso-scale and provides more detailed forecasts of weather parameters close to the ground. For higher elevations, atmosphere is split into 20 layers of increasing thickness in a hybrid coordinate system. Eight layers are used for the lowest 2 km; four layers are placed around the tropopause and in the stratosphere. The integration of EM covers the North Atlantic and Europe with a net of 181*129 grid points (Mesh size 0.5^0). The true distance between grid points varies from 47,1 to 55,6 km.

C. *Horizontal Interpolation and Vertical Assignment*

METEOSAT AND- end EM-data are of different horizontal resolution. Due to the worse resolution of the model data, four grid points close to the METEOSAT pixel are chosen and data are bi-linearly interpolated to the pixel.

Then the level of best fit between the cloud top temperature from METEOSAT and the model temperature-level is searched in the vertical direction. The model level is interpolated to the level which corresponds to the cloud top temperature.

D. *Semi-transparent Cirrus*

The problem of semi-transparent cirrus is a well known problem: in the IR-image clouds appear as warmer clouds assuming to be in the lower atmosphere [3]. To overcome misinterpretations, the following assumption is made: cirrus clouds exist in regions of atmosphere where water vapor content is rather high. Using the METEOSAT WV image, which provides humidity information from higher levels of atmosphere, low clouds with high level of water vapor content above them are marked with a certain color (Fig. 3).

Not all cloud pixels are automatically marked, but the user can be warned when such suspicious clouds are detected.

E. *Results*

Cloud top charts are calculated four times a day by the Research and Training Laboratory at the Technical University in Sofia, Bulgaria and by the Department of "Medientechnik" at the Technical University in Ilmenau, Germany. Cloud level is evaluated using 6- or 12-hour forecasts from the EM. Cloud top charts are displayed within the interactive Graphical System (IGS). Figure 3 shows an example. At the left side, the color legend for the height assignment (in km) is displayed; for example heights between 0 and 12.5 km are in brown. In Fig.1 semi-transparent clouds are marked in a bright blue color; they cover a stripe from the coast of Norway to the Baltic Sea. A very narrow stripe, caused by an aircraft, can be seen over the North Sea for example. A similar strip can be seen over the North Sea just off the south coast of Norway. The highest tops result from thunderstorms near the east coast of Italy with heights of up to 11 km. From time to time, cloud top heights are compared with radio-probe measurements and show good agreement. Figure 4 displays a local radar image from Bulgaria. There are not any principal differences between this one and the above mentioned images without correction (Fig. 5).

F. *Problems*

Due to the different systems of the high assignment method, i.e. satellite radiation measurement and model simulation, the following problems may arise:

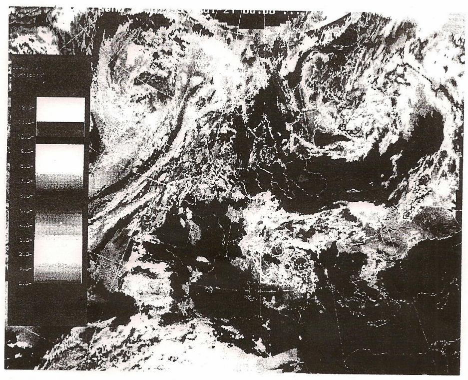


Fig. 3: METEOSAT CTH, 1994-JUL-21, 06 UTC semi-transparent cirrus is marked

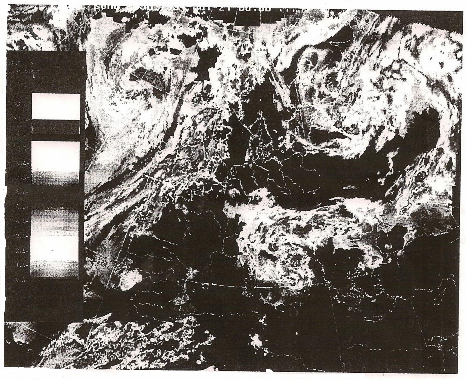


Fig. 4: Same situation: cirrus not marked, ground effects corrected

a) model forecast is incorrect: fictive clouds would occur; b) model forecast is correct but surface model temperature is higher than the surface layer sensed by satellite: fictive clouds would also occur.

These and other problems have to be considered and resolved by the assignment method.

III. Conclusion

Radar images verification against METEOSAT data is a simple method to present a radar image with less ground clutter echoes. This method does not work if ground clutters appear simultaneously with clouds without rain. Another method is to analyze regions with high probability risk of ground clutters with numerical algorithms (this method is under development) or even better, with a new generation of radar stations like Doppler radar.

First results show that the method is capable of assigning correct cloud top highs with the aid of METEOSAT cloud top temperature and EM temperature forecasts. Users are warned of suspicious cloud top levels which are flagged by a warning color [4].

The operator (teacher or student) can retrieve satellite and radar data together from the hard disk for display using a mouse controlled menu. On top of the storage capability, facilities allow the user to: do series of radar and satellite images, use multicolor density slice feature, estimate temperature and contrast stretch, improve image definition. Radar images verification with the meteorological images is possible using appropriate software.

References:

[1] Dimitrov D., U. Petrova. Computer presentation of images from METEOSAT satellite. CAEE'95, Bratislava, 1995, pp 191-195

[2] Mayewski D. The Europa model. ECMWF Seminar of numerical methods in atmospheric models. 1991, vol.2, pp 147-191.

[3] Schmetz J. Operational Cloudmotion Winds from Meteosat Infrared Images. Journal of Applied Meteorology, Vol.32, № 7, July 1993, pp 1220 -1225.

[4] htpp://www.eumetsat.de/

Basic properties of the real signals

Petre Tzvetanov Petrov[1]

Abstract **– Definitions of the basic parameters of the real signals are given. The disaccord between the engineering world with the real signals and the world of mathematicians with the oversimplified mathematical models for these signals is obvious especially in the following theoretical points: the classical sampling theorem, Fourier series, Dirichlet conditions, Gibbs phenomenon, Dirac function (delta function), Dirac comb, unit step function, staircase functions (based on rectangles) and theirs derivatives. These models are declared misrepresentatives and attempts to reject, modify and replace them were made. The paper is developed for educational and research purposes.**

Keywords **– real signals, simplest signals approximations**

I. Introduction

Although during the twentieth century the signal sampling (and reconstruction) theory (SST) and the digital signal processing (DSP) did significant technological progress there are several basic principles and simplified mathematical models (SMM) which could be questioned in order to be rectified or replaced with something more accurate when it is possible. The SMM mention in the paper and widely used in the SST do not represent the real signals (RS) because of the following common errors:

1/ oversimplification and overgeneralization of the RS.

2/ introduction of artificial and misrepresentative *theoretical concepts* which are irrelevant to the engineering practice (they cannot be observed, used or proved practically).

3/ not clear definitions of the RS and their parameters

4/ wrong models of error evaluations of the RS approximations.

Consequently, the SMM in the SST (discussed in this paper) with these drawbacks do not have the same natures as the RS, do not represent it and should be reevaluated. They are oversimplified mathematical models (OSMM).

The disaccord between the RS and the OSMM is obvious especially in the following theoretical points: the classical sampling theorem (CST), Fourier series (FS), Dirichlet conditions (DIC), Gibbs phenomenon (GP), Dirac function or delta function (DEF), Dirac comb (DCM), sin(x)/x function (SXF), unit step function (USF), unit step pulse (USP), rectangular staircase function (RSF) and theirs derivatives.

The present paper is based on the ideas developed in [1, 2]. The notion of the "signal sampling factor" (SSF) $N=F_d/F_s$ was introduced and the maximal amplitude error E_{max} (due to the non sampling the sinusoidal or cosinusoidal signal (SS or CS) into its maximum) during the sampling process was calculated [1,2].

The fundamentals of the classical SST could be found in many sources e.g. [3, 4, 5]. This paper is concentrating on the well-known and widely published fundamentals of the SST because the oversimplification and the overgeneralization have made harm there and are influencing the concepts in the higher levels of the theory.

II. Basis properties of the real signals (BPRS)

The RS (the physical signals) are generated, transmitted and processed by material objects. In order to analyze and use the RS simplified but representative mathematical models (MM) are needed. The difference between a representative mathematical model (RMM) and a misrepresentative (non representative) mathematical model (MRM) is that the RMM is tightly related to the RS and has the same "nature". The MRM and the RS have different nature. An MRM cannot be reproduced into material world. More precisely RMM has the same principal parameters as the RS and the differences (the errors) between the RMM and the RS are clearly defined and calculable. Usually, an error < 1% (0.1 dB) is a reason to accept the model as a accurate RMM. An error in one or more of the parameters >10% (1dB) and especially > 30% (3dB) is making the MM misrepresentative.

Unfortunately due to the widely spread oversimplification, overgeneralization and insufficient error definitions a lot of the basic models in SST are oversimplified MM (OSMM) or inappropriate applications of RM. We will mention a few examples: GP, "ideal rectangular pulse (signal)" (IRP), "ideal saw tooth signal" (ISTS), CST, DEF, DCM, SXF, USF and FS. The present paper is calling for a revision, correction and/or abandoning of these basic SMM in the SST.

The RS could be very different and frequently are defined as "fast", "slow", "audio", "video", "wide band", "narrow band", "with one, two or multiple frequency (signal components)", "with one two or many sinusoidal components", "digital", "pulsed" etc. but they have common basic properties, as stated below. During the definition of these properties it is taken into consideration that the inevitable parasitic resistors, capacitors and inductances are involved into generation, transmission and processing of the RS.

Definition 1: The simplest idealized band wide signal (SBLS) is given with one of the following formulas:

(1) $A = A_m\sin(2\pi F+\varphi)+C$

(2) $B = B_m\cos(2\pi F+\varphi)+C$

where Am and Bm are the maximal values, F is the linear frequency, φ is the phase and C is the offset or DC. The noise component of this signal is neglected.

All MM should be tested with SBLS in order to verify their representativeness.

[1]Petre Tzvetanov Petrov is with Microengineering-Sofia, Bulgaria. Emails: ptzvp@yahoo.fr and ppetre@caramail.com.

Definition 2: The postulate for the basic properties of the real signals (PBPRS) is stating that the basic properties of the RS an the RM are as follow:

* Finite spectrum because the RS is produced in the material world
* Finite slew rate (SR) in every parameter due to the reason stated before.
* Finite power and finite amplitude due to the finite among of energy, which could be produced, transmitted and processed by the RS.
* Direct current (DC) component and phase component (FC). (Frequently these two parameters are accepted to be zero or neglected, but they are always present and should be taken into consideration.)
* Noise component (NC). Sometimes it could be predicted or neglected in the simplest RMM.
* Every RS could be represented as a sum of SBLSs with individual DC and FC

Several more notes could be added as comments to the basic properties of the RS and to the SBLS:

1. It is important to repeat that the slew rate (SR) of every RS is always less than infinity and never should be accepted to be infinity. Consequently the SMM with infinite SR as the IRP, RSF, ISTS, DEF and USF are too idealized for the world of engineers.

2. The first derivative of the RS always exists and is less than infinity. Replacing a finite value with infinity is creating an MRM (e.g. GP, DCM, DIC, DEF, etc).

3. The spectrum of the RS is always finite and there is no use to replace it with infinite spectrum (e.g. FS).

4. The spectrum of RS is containing a finite number of spectral lines which could have any value between zero Hertz (or DC) and one maximal frequency F_{smax} determined by the material in which the RS is produced, transmitted or processed. There is no use to represent the RS as an infinity sum of signals (e.g. FS).

5. There is no use to represent obligatory a RS with sum of harmonically related SS and CS. But it is useful to represent RS with an algebraic sum of SS and CS, which could be harmonically or not harmonically related and with individual phase and individual DC component (e.g. FS).

6. If the RS is interrupted in point where the momentum power is equal to zero ($P_{mom} = 0$) this does not change its spectrum. That means that a half of sinusoidal wave, sequence of half waves and a full sinusoidal wave (all of them with no DC component) do have the same FC into theirs spectrums but the amplitude of the FC are different (e.g. FS).

7. The SBLS is containing two spectrum lines. The first is the DC line and the second is any frequency $F_s>0$ approximated with SS or CS. SBLS is used as a basic test signal everywhere when it is possible and as a basic component into the sum for approximation of any RS.

8. The RS does not have ideal “angles”. The “angles” of the RS are always “with curve” or “rounded”.

The RS could be very complex and in order to analyze it a lot of simplified mathematical models (SMM) are used. There are at least several points which are not respected in the SMM used into classical SST:

* The model and the RS have differences in all parameters. These differences should be evaluated and expressed in clear and verifiable form with the corresponding errors. This is not the case with CST, USF, FS, etc.
* There are always limits and if we cross them the SMM and the RS does not have the same nature and the model does not represent the RS. One of the limits is given with the finite value of SR.
* The oversimplification and the overgeneralization applied to the RS leads to wrong conclusions, (e.g. CST).

The points mention before are the main sources of errors between the RS and the SMM in the following cases: the CST, FS, DIC, GP, DEF, DCM, the approximation of the RS with set of DEF or Dirac-like functions and USP. These models have the following common error: “Replacing a RS with too idealized MM with infinite slew rate and ideal angles is leading to the non existing effects and wrong conclusions”. A good example of this error is the replacement of the real “trapezoidal pulse with rounded angles” with “trapezoidal pulse with ideal angles” and even worse with IRP. This replacement is leading to the non-existing phenolmenon called “GP”. Table 1 is giving an idea about which of the basic properties on the RS are not respected into the SMM listed above.

TABLE 1
THE OVERSIMPLIFICATION OF THE REAL SIGNALS INTO SOME THEORETICAL MODELS

#	Basic property of the RS	Frequently oversimplified parameter or signal	Theoretical model in the SST based on this oversimplification
1	Finite amplitude and finite peak to peak amplitude	Infinite amplitude	Delta function, Dirac function
2	Finite power in every moment	Infinite power during the transitions, rectangular signal, saw-tooth signal	Delta function, Dirac function, unit function, rectangular staircase function , unit step function
3	Finite spectrum, finite number of spectral lines	Infinite sum	Fourier series, Infinite polynomial approximation
4	Finite slew rate	Infinite SR, rectangular signal, saw-tooth signal	Gibbs phenomenon, unit step function, unit pulse, RSF
5	Individual DC component	Omission of the DC current component	Classical sampling theorem, Fourier series
6	Individual FC	Omission of the FC	Classical sampling theorem, Fourier series

III. FOURIER SERIES (FS)

From all of the “classical” theoretical concepts mention in the paper perhaps only the FS is still partially useful. Although the FS has several weak points as an attempt to approximate every periodical function (signal) with sum of SS and CS with harmonic frequencies, it is still interesting in the SST because the FS and RS has same nature. The FS is a simple and natural way to approximate some of the RS because it is relatively easy to use set of sinusoidal and cosinusoidal generators working on harmonic frequencies, to forward theirs outputs into analog summing amplifier and to produce many useful RS. But the standard FS has the following disadvantages:

TABLE 2
COMPARISONS BETWEEN THE BASIC PROPERTIES OF THE RS, FS, DIC, GP AND THE PRSS

#	Real signal (RS)	Furrier series (FS)	Dirih-let conditions (DIC)	Gibbs phenomenon (GP)	Postu-late (PBRSS)
1	Always finite spectrum	Not necessarily (NS)	NS	Infinite (*)	Yes
2	Always finite number of spectral lines (of SC)	NS	NS	Infinite (*)	Yes
3	Finite slew rate	NS	NS	Infinite (*)	Yes
4	Finite amplitude (of RS and of all of its SC)	Yes	Yes	Yes	Yes
5	Existing phase	No	No	No	Yes
6	Separate DC for each SC	Common DC	Not discu-ssed	Not discu-ssed	Yes
7	Separate FC for each SC	No sepa-rate FC for SC	No separate FC for SC	No separate FC for each SC	Yes

Notes: SC = Signal component. NS = Not necessarily. The SC could be SS or CS. (*) – Approximation with infinite SR.

* The individual phase of the SS / CS components in the sum is neglected.
* The individual DC component of the SS/CS components in the sum is neglected.
* Sometimes it is simpler to represent a RS as sum of SS/CS signals, which are not harmonically related.
* The sum of SS and CS for approximation of a RS should be finite. The reason is that the RS has always a finite spectrum and finite number of spectral lines in every moment.
* One more problem is that the FS is suffering from wrong applications from the part of the mathematicians. The wrong application of the sum of SS and CS (e.g. to the unreal signals as IRP and ISTS) leads to the DIC and GP, which in fact have no engineering value.

These are the raisons to extend the FS. A more general expression for RS approximation is

1/ with finite sum of non-harmonically related signals

2/ every member is given with formulas (1) or (2).

IV. DUAL-TONE AND MULTI-TONE MULTI FREQUENCY SIGNALS

Dual tone multi frequency signals (DTMF) and multi tone multi frequency signals (MTMF) are used for message transmition, canal testing and testing of the equipment and components. In some cases the frequencies of the tones are not harmonically related and the tones could be not periodical because they are produced and transmitted during short intervals. At least three principal points exist:

1. The phase relations and the DC components could be important. Consequently the FS, which contains harmonically related components without individual phase and DC components, is not always the best or the simplest solution.
2. There is no use to represent a sum of two non-harmonically related SS/CS with an expression, which contains more that two harmonically related SS/CS.
3. There is no use to express a RS with obviously short (finite) spectrum as too long (or infinite) sum of harmonically related SS/CS. Due to the way DTMF and MTMF signals are produced it is known that they have into the spectrum a finite number of SS and CS components with finite parameters.

Any DTMF r(t) could be presented as an algebraic sum of the two signals (SS and/or CS) with the formula below:

(3) $r(t)= s_1(t)+ s_2(t),$

where

(4) $s_1(t) = (A_{0k} + A_m\sin(2\pi F_k t +\varphi_k))$

and

(5) $s_2(t) = (B_{0k} + B_m\sin(2\pi F_l +\psi_l))$

or the expression for the DTMF is

(6) $r(t) = (A_{0k} + A_m\sin(2\pi F_k t +\varphi_k)) + (B_{0k} + B_m\sin(2\pi F_l +\psi_l))$

DTMF and MTMF have properties not covered by FS:
* Always finite sum of SS and CS
* Individual phase for each signal component (SC)
* Individual DC component for each SC
* Non-harmonic SS and CS into the sum in order to use shorter and simpler mathematical expression.

Keeping SS and CS into the sum is important. Making the difference between the SS and CS is also important and one of them cannot be omitted. Making a difference between the SS and CS is equivalent of making a difference between the four quadrants and the four principal directions into two axe coordinating system (x,y). The classical FS series (which could be found in [3, 5] is stating that every periodical function (in particular a signal) s(t) with period T could be represented as a sum of SS and CS signals with periods n/T (n = 1, 2 , 3…) plus possibly a constant C and the following general expression is applicable [e.g. 5]:

(7) $s(t) = C + \sum (A_{mk}\sin(2\pi F_k t)) + \sum (B_{ml}\cos(2\pi F_l))$

where k and l are integer number which could be infinity.

V. POSTULATE OF RS SPECTRUM PRESENTATION (PRSSP)

The postulate was formulated in order to:
1) Approximate every RS with basic properties in Table 1.
2) Describe in a better way any DTMF and MTMF with non-harmonic frequencies of the SCs
3) Describe RS used during the testing channels, equipment and components.

Definition 3. Postulate PRSSP: Every RS r(t) with basic properties given in Table 1 could be represented as a sum of SS and CS signals components each of them with four basic components (A_0, A_m, F and φ) and given with one of the two formulas below:
a/ for every sinusoidal SC $s_{ss}(t)$

(8) $s_{ss}(t)= A_0 + A_m\sin(2\pi F_{ss}t +\varphi))$

b/ for every cosinusoidal SC $s_{cs}(t)$

(9) $s_{cs}(t)= B_0 + B_m\cos(2\pi F_{cs}t + \psi)$

where
A_0, (B_0), A_m (B_m), F_{ss} (F_{cs}) and φ (ψ) are respectively the DC part, amplitude, frequency and the phase of the SC.

The signal r(t) could de given with the formula below

(10) $r(t) = r_{ss}(t) + r_{cs}(t)$

where
$r_{ss}(t)$ is the sum of all sinusoidal components in the spectrum of RS
$r_{cs}(t)$ is the sum of all cosinusoidal components in the spectrum of RS

The sum of all SS components $r_{ss}(t)$ is given with the formula

(11) $$r_{ss}(t) = \sum_{0}^{p} (A_{0k} + A_{mk}\sin(2\pi F_k t +\varphi_k)) ,$$

k= 0, 1, 2...p

The sum of all CS components $r_{cs}(t)$ is given with the formula

(12) $$r_{cs}(t) = \sum_{0}^{q} (B_{0l} + B_{ml}\cos (2\pi F_l +\psi_l)),$$

l = 0, 1, …q

Consequently the main task of the signal conversion and reconstruction is to convert and reconstruct RS with known error in each signal component. Testing the spectrum of the reconstructed signal should be done separately with analog and digital filters and the results should be compared.

VI. CONCLUSION

Basic properties of the RS are defined. A postulate of the presentation of every RS as a sum of SS and CS components each with four basic parameters is given the postulate is an extension of the standard Fourier series. Example with DTMF and MTMF signals is given. This topic is continued in [6].

The paper is intended to help students and researchers to evaluate the existing MM and to replace them with more representative models when it is necessarily. New ideas and examples in the field are given.

VII. ABBREVIATIONS IN THE PAPER

BRPS – basic properties of the real signal
BLS – band limited signal
SS/CS – sinusoidal/cosinusoidal signal
CST – classical sampling theorem
DC - Direct current
DCM – Dirac comb
DEF – delta function, Dirac function
DIC – Dirichlet conditions
DTMF - Dual tone multi frequency signals
MTMF - Multi tone multi frequency signals
FC - phase component
FS - Fourier series
IRP - ideal rectangular pulse (signal)
ISTS - ideal saw tooth signal
GP – Gibbs Phenomenon
MM – mathematical model
MRM - misrepresentative mathematical model
NS – Not necessarily
NC - Noise component
OSMM – over-simplified mathematical model
PBPRS - postulate for the basic properties of the real signals
PRSSP – postulate for real signal spectrum presentation
RSF - rectangular staircase function
RMM – representative mathematical model
RS – real signal
RSF – rectangular staircase function
TPIA – trapezoidal pulse (signal) with ideal angles
SBLS – simplest band wide signal $A= A_m\sin(2\pi F+\varphi)+C$
SC – signal component
SST - signal sampling theory
SSF – signal sampling factor N=Fd/Fs
SST – signal sampling theory
SR – slew rate
SSM – simplified mathematical model
USF – unit step function

VIII. REFERENCES

[1] Petre Tzv. Petrov, Sampling analog signals (abstract with formulas, tables and figures), Brochure printed in Technical University–Sofia, February 2004, Bulgaria, ISBN 954.438.385.9
[2] Petre Tzv. Petrov , Method of calculating the sampling factor N=Fd/Fs for analog signal conversion when the maximum amplitude error Emax is known and vice versa. Journal “Success of Modern Electronics”, vol. 7, 2004/Foreign radio-electronics 2/2004, Moscow, Russia.
[3] F. Manneville, J. Esquieu, Théorie du signal et composants, Dunod, Bordas, Paris, 1989. ISBN 2-04-018915-7. France.
[4] C. E. Shannon. A mathematical theory of communications. The Bell Systems Technical Journal, pp 379-423, 623-656, July, October. 1948. USA.
[5] Openheim A. V., A.S. Wilsky, I. T. Young. Signals and systems, Prentice Hall Inc. 1983. USA
[6] Petre Tzv. Petrov, “Mathematical models” versus “real signals”, ICEST-2006, Sofia, Bulgaria.

„Mathematical models“ versus „real signals“

Petre Tzvetanov Petrov[1]

***Abstract* – The disaccord between the engineering world with the real signals and the world of mathematicians with the mathematical models for these signals is obvious especially in the following theoretical points: the classical sampling theorem, Fourier series, Dirichlet conditions, Gibbs phenomenon, Dirac function (delta function), Dirac comb, sin(x)/x function, unit step function, staircase functions (based on rectangles) and theirs derivatives. These models are declared misrepresentatives and attempts to reject, modify and replace them were made. The paper is developed for educational and research purposes.**

***Keywords* - real signals definition, simplest signals approximations**

I. INTRODUCTION

New ideas in the signal sampling and reconstruction theory (SST) are given in [1,2,3]. But the classical SST is still using a lot of mathematical models (MM) which do not represent real signals (RS) [4, 5, 6]. The students and the researchers in the engineering and mathematical fields are having difficulties to understand and apply these models. The problem is discussed in this paper and solutions are proposed.

A clear difference between the term “mathematical model” (MM) and the term “real signal” (RS) should be made. These two terms have different nature. For example it is acceptable from mathematical point of view to imagine a MM with the following properties:
1/ infinite number of minimums and/or maximums
2/ ideal angles (45 deg, 90 deg. etc)
3/ infinite slew rate (SR)
but obviously it is not possible to produce RS with these parameters. Consequently these MM are not applicable to the SST and to the engineering (physical) world. They are misrepresentatives MM (MRM). The primary objective into the engineering world is to develop and utilize representative MM (RMM).

When a MM of the RS is developed, an evaluation of the errors in each parameter (amplitude, phase, spectrum, frequency, slew rate (SR), power, etc) should be done. This is not an easy process but without it there is a danger to develop a MRM and to draw wrong conclusions. In fact the level or the representativeness of the MM should be evaluated and stated in each particular case.

A lot of differences between the technical process (TP) (e.g. the real signal) and the MM exists. A bidirectional relation should be created and tested as follow:
1. The TP should be replaced with closest and simplest possible RMM. The nature of the TP should be respected and exaggerated simplification and overgeneralization should be avoided.

[1]Petre Tzvetanov Petrov is with Microengineering-Sofia, Bulgaria.
Emails: ptzvp@yahoo.fr and ppetre@caramail.com.

3. A difference should be made between RS and the MM with theirs mathematical functions (expressions). Not every mathematical function (MF) could be converted into RS but every RS could be represented with mathematical expressions**.**

II. DIRICHLET CONDITIONS (DIC)

The attempt of applying the Fourier series (FS) to “every periodical function” is suffering from overgeneralization, oversimplification and non-respect of the basic rule that “the “original” and “the approximation” (or the RS and its SMM) should have to have the same nature and the same basic properties in order to make parts of a representative couple “signal-model”. The FS could be very useful when applied to a lot of periodical RS with neglected noise components but not to “every periodical mathematical function or RS”.

The DIC are result of attempt to rectify some weak points in the FS but they are suffering from the same weak points. due to the oversimplification, overgeneralization and are misrepresentatives. Perhaps from the mathematical point of view the DIC are acceptable. But if the FS is used to represent a RS from the engineering world the DIC are useless because every RS is satisfying the conditions stronger than DIC For example, a RS with infinite number of minimums and maximums cannot be produced with technical devices Obviously the DIC conditions are not related to the RS and to the SST. A comparison between these conditions and the basic properties of the RS is stated in [3]. It is obvious that the RS are always satisfying the DIC. Consequently there is no use to mention DIC in relation to the FS and RS into SST.

Moreover we could mention the following against the DIC, delta function (DEF), Dirac comb (DCM), GP and unit step function (USF):
* The tension on the capacitor could not be discontinued. The circuit and signal source without parasitic (inherent) capacitor(s) could not exist. Consequently the SR of every signal source is limited and the “angles” of the produced signal are never ideal. How the “angles” will be approximated depends on the application. For example they can be approximated with exponential or sinusoidal functions.
* The same reason as above but applicable to the current traversing the inductance internal to the source and the inductance(s) of the signal connecting and processing elements.
* The electric chargers and waves have finite speed of movement. Consequently the SR of every signal source is limited and depends on the material where the signals is produced, transmitted and processed.
* Replacing a finite RS with infinity is unacceptable source of error in most of the cases and creates non-existing effects and infinity spectrum. The GP is an example of an artificial effect in the engineering world. (Perhaps it has some value in the world of mathematics.)

III. Classical Sampling Theorem (CST)

The classical sampling theorem (CST) or so called "sampling theorem of Kotelnikov - Shannon- Whittaker -Nyquist-etc" is based on DEF, DCM, FS and function sin(x)/x (SXF). CST is pretending in general "to represent or reconstruct exactly any band limited signal (BLS) with maximal frequentcy F_{smax} under the condition $Fd>=2F_{smax}$". The CST has many weak points [1, 2] and we will mention:
1/ lack of errors evaluations,
2/ possible exception (sometimes the signal could be reconstructed with signal sampling factor (SSF) N <2),
3/ possible lost of the signal with SSF N=2,
4/ wrong MM and
5/ possible strong and not desirable phase modulation.

The lack of errors evaluation is very important disadvantage of the CST because once converted into digital form the RS is reconstructed with differences (errors) in every parameter and these differences should be evaluated. Or the ADC is leading to the irreversible lost of information and this lost should be evaluated. Obviously CST is a result of the overgeneralization and oversimplification. If we compare the CST with another fundamental theorem (e.g. with the theorem of Pythagoras $c^2 = a^2 + b^2$) we will understand the differences:
* in the definitions (CST is not well defined),
* precision (lack of precision in the CST) and
* usefulness (CST is not useful in practice)
between the two theorems.

In many cases a set of RS with known parameters is used. This is giving the possibility to use the method "a particular solution for a particular case" which means that representative sampling theorems (RST) and RMM could be developed for each case (e.g. for each signal or group of signals) and under appropriate conditions (simplifications).

Several practical examples could be given concerning sampling the basic RS e.g. direct current (DC), sinusoidal signal(SS), co sinusoidal signal (CS), sum of SS and CS, triangular-like signals and trapezoidal-like signals with different form etc. When it is possible and necessary several particular RST and RMM could be put in one common RMM or RST under proper conditions. A good example for this approach is the sampling of SS and CS signals discussed in [1, 2].

Obviously, new RMM should be developed. Definitions of the parameters of RS, evaluations of the existing MMs and proposing new and better (more accurate and closer to the RS) models are necessary.

IV. Gibbs Phenomenon (GP) versus "Ringing" and "Approximation Error"

The GP is frequently discussed in the manuals for engineers [e.g. 6] and in software dealing with "signals and systems" such as Matlab. The usual example is an approximation of "ideal rectangular pulse" (IRP) with FS. This example is obviously wrong. In fact the GP is born due to the common error repeated widely in the SST: "Due to the oversimplification (overgeneralization) the SMM (in this case the IRP) does not represent the RS". The nature of the RS is not respected and an artificial (non-existing) and not useful "phenomenon" is "observed". The GP does not exist for the "trapezoidal-like signals" or "square-like signals" (signals with finite SR, with "rounded angles" which resembles to the IRP but have different nature). In fact the "real rectangular pulse" (RRP) is always a "trapezoidal-like signal with rounded angles"! The conclusions are:
1/ the GP and the DCI are not related to the RS and the SST
2/ the error of the approximation of an RS with the FS cannot be considered as a GP.

A clear difference between the process of "ringing" [6] and the GP should be done. The effect of "ringing" or the effects of "overshoots" and "undershoots" are real effects when a RRP is applied to the channel with inappropriate load, narrower frequency band or with not enough equalized frequency band. But the "overshoots" and the "undershoots" have nothing in common with the artificial GP because at least of the following reasons:
1. The IRP used to illustrate the GP is too idealized and does not have a real value. As we said before: "The infinite SR and the ideal angles cannot be generated, transmitted and processed in a material world and there are two mistakes:
a/ replacing a finite number with infinity
b/ replacing a continuous function with discontinuous function."
2. Trying to approximate an ideal signal with ideal "angles" (IRP, ISTS etc.) with set of real functions with different nature (e.g. SS and CS) is wrong because of the different nature of the signal and its approximation. Also, there is a difference between the GP (an inappropriate approximation) and the "approximation error due to appropriate but not exact approximation".
3. There is no sense to approximate a RS with obviously finite spectrum (and with finite number of spectrum lines) with infinity sum of SS and/or CS (e.g. FS).
4. The nature of the process of "ringing" due to the inappropriate loading is different from the process of GP. In fact a RRP passing through a real channel (circuit) with inappropriate load or frequency bandwidth is producing the real effect of "ringing" which is important and informative but this is not a GP.

The conclusion is that "the ringing effect exists but the GP does not exist in the engineering world or more exactly it is not applicable to the SST. Replacing an IRP with FS is like an attempt to replace an imaginary number with infinite sum of real numbers and to discover that there is a "problem" or an "error".

Also, the following main differences between the process of "ringing" and the GP exist:
1. The "ringing" is applicable for the transitions of the signal. It is not applicable for the horizontal parts of the pulse. The GP is applicable for both parts.
2. The "ringing" is not always symmetrical for the positive and the negative transition and frequently is applicable only for the end of the transition. The GP is symmetrical for the both transitions and exists for the beginning and the end of the transitions.
3. The "ringing" is a natural process due to the disaccord between two real objects (the "signal" and the "channel" or

the “load”). The GP is an artificial effect due to the dissacord between a RS (sum of SS and/or CS) and imaginary signal (e.g. the IRS and ISTS).

Also the GP is not equivalent to the normal “approximation error” between a RS and its RM. As it was mention before the RS and its RM has the same nature and the same basic properties and one or more approximation error(s) could be defined. The GP is the result of attempt to approximate an imaginary signal with algebraic sum of real signals.

V. Examples for oversimplification

Figure 1 is showing examples of the oversimplification of two pulses.- trapezoidal and triangular.

Figure 1a and Figure 1c are showing idealized (simplified) but still representative models with straight lines and “rounded” angles. They are built by two basic element:

* The first element is a non-ideal “rounded” angle which could be approximated with exponential or sinusoidal function in the simpler cases.
* The second element is a linear function (straight line) with finite SR.

In the connection point these two elements have the same value and same SR.

The models on Figure 1 b and Figure 1d are oversimplified and misrepresentatives. The oversimplified (over generalized) models (OSMM) are containing artificial (non existing) elements as ideal angles and straight lines with infinite SR. They should be abandoned because theirs properties are too far to the properties of the real signals and are leading to the wrong conclusions.

VI. Conclusions

Several important conclusions could be made:

1. The DIC (discontinuity, derivation, limited number of maximum and minimums, etc.) are not related to the application of FS to approximate RS in the SST. Every RS is satisfying the DIC. Consequently DIC should not be mention in the SST in relation with FS and RS.
2. Although the FS is a good approximation of many periodical RS it is incomplete, e.g. it cannot approximate in a simple way the RS constructed by a sum of the simplest band limited signals (SBLS) and containing SS and/or CS with non-harmonic frequencies and an individual DC component. Consequently, a more general sum containing SS and CS with non-related parameters (amplitude, frequency, phase and DC component) should be used to approximate any RS [3].
3. The Dirac comb (DCM) is based on the rule “take and forget” and does not correspond to the real world application. In fact the RS is always replaced with a step function (SF) based on the rule “take and memorize the present sample until the next sample became available”.
4. A difference should be made between the physical (engineering) world and the SMM. In the world of mathematicians a lot of things are possible but in the world of the engineers a few things are realizable and observable. For example an SMM, which have to be used in the engineering world, have to have the same nature as the modeled engineering process. Not respecting the nature of the engineering (physical) process is leading to MRM and to wrong conclusions.
5. The GP is declared, “non-existing in the engineering world” due to the fact that the RRP is in practice “a trapezoidal-like signal with rounded angles”. Consequently the GP is not applicable to it and is not observable. The GP and the natural effect of “ringing” due to improper loading or bandwidth are declared with different nature.
6. The CST is considered “non efficient” due to the oversimplification, lack of universality and no errors evaluation. The DEF, DCM and the SXF are considered “irrelevant to the sampling and reconstruction process”.
7. With the exception of the FS, which is still utilizable in some cases and after “reparation”[3], the rest of the classical concepts stated above are irrelevant to the SST. In order to be more useful the FS should be extended with new parameters (individual phase and DC component of every SS/CS component) and with non-harmonic frequencies of the SS/CS components.
8. Postulates about the composition of RS are formulated. The examples with DTMS and MTMF are given in [3].
9. The oversimplifications and the overgeneralizations are important sources of errors in the classical SST and should be avoided. Also it is important to test the representativeness of every SMM and to evaluate the errors between SMM and replaced RS. The main sources of the oversimplification into the NRM are the replacement of the finite SR with infinite and the rounded angles with ideal or “broken” angle.
10. During the sampling and reconstruction process the approximation of the RS is with non-ideal trapezoidal stair case function (TRSF with finite slew rate).
11. The SMM but still RMM for RS should be constructed from two segments: “lines” and “rounded angles”. This SMM must not be further simplified because an oversimplification will occur and a MRM will be produced with non-existing and not useful artificial effect (e.g. GP and CST)

The paper is making revision of some of the fundamental principals in SST and DSP and replacing them with more realistic models.

The paper is developed for the educational and research purposes. It is giving new ideas and examples for evaluation of the relations between the basic properties of the real signals and the representatives and non representatives simplified mathematical model of these signals. It is helping the students in the engineering and mathematical field to understand the nature of the real signals and the relations between them and the RMM and NRM. New models could be developed with these ideas and the old modes could be replaced.

VII. Abbreviations in the paper

BLS – band limited signal
BRPS – basic property of the real signal
CS – co sinusoidal signal
CST – classical sampling theorem (Kotelnikov-Whittaker-Shannon – etc.)
DSP - digital signal processing

DC - Direct current
DCM – Dirac comb
DEF – delta function
FS - Fourier series
IRP - ideal rectangular pulse
ISTS - ideal saw tooth signal
GP – Gibbs phenomenon
MF – mathematical function
MM – mathematical model
MRM - misrepresentative mathematical model
NC - Noise component
OSMM – over-simplified mathematical model
RRP – real rectangular pulse
RSF - rectangular staircase function
RMM – representative mathematical model
RS – real signal
RSF – rectangular staircase function
TP – technical process
TPIA – trapezoidal pulse (signal) with ideal angles
TRSF – trapezoidal stair case function
SBLS – simplest band wide signal $A= A_m\sin(2\pi F+\varphi)+C$
SC – signal component
SST - signal sampling theory
SS - sinusoidal signal
SSF – signal sampling factor N=Fd/Fs.
SST – signal sampling theory
SR – slew rate
SSM – simplified mathematical model
RST – representative sampling theorem
SXF – sin(x)/x function
USF – unit step function

REFERENCES

[1] Petre Tzv. Petrov, Sampling analog signals (abstract with formulas, tables and figures), Brochure printed in Technical University –Sofia, February 2004, Bulgaria, ISBN 954.438.385.9

[2] Petre Tzv. Petrov , Method of calculating the sampling factor N=Fd/Fs for analog signal conversion when the maximum amplitude error Emax is known and vice versa. Journal " Success of Modern Electronics", vol. 7, 2004/Foreign radio-electronics 2/2004, Moscow, Russia.

[3] Petre Tzv. Petrov, Basic properties of the real signals. ICEST-2006, Sofia, Bulgaria

[4] F. Manneville, J. Esquieu, Théorie du signal et composants, Dunod, Bordas, Paris, 1989. ISBN 2-04-018915-7. France.

[5] C. E. Shannon. A mathematical theory of communications. The Bell Systems Technical Journal, pp 379-423, 623-656, July, October. 1948. USA.

[6] Openheim A. V., A.S. Wilsky, I. T. Young. Signals and systems, Prentice Hall Inc. 1983. USA

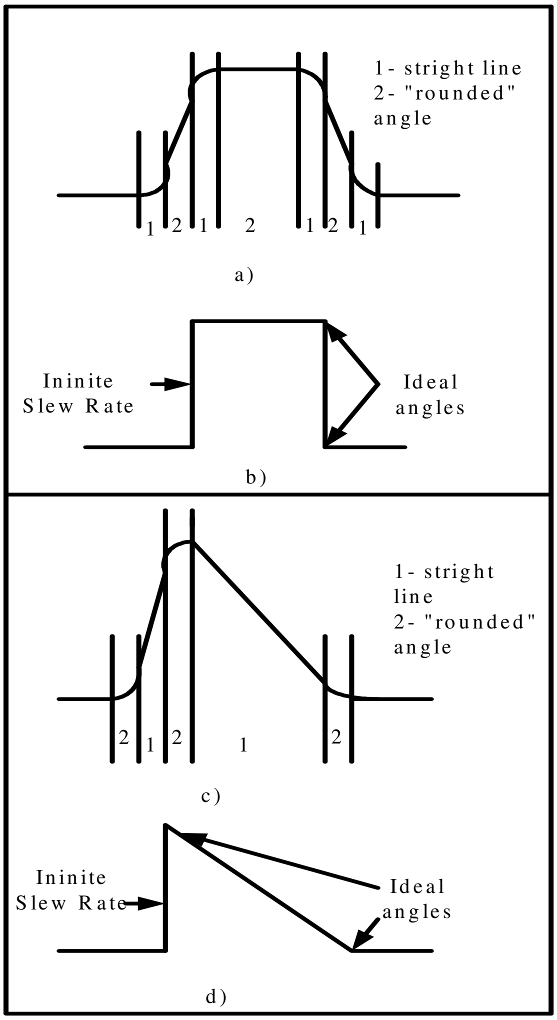


Figure 1

AUTHOR INDEX