Joint Source-Channel CELP Coding

Snejana G. Pleshkova-Bekiarska[1]

Abstract – **The method of speech coding CELP is extensive used in many voice communications, multimedia, video conference and other systems. There are a lot of articles related to CELP coding characteristics improvement and as a final result to realize more speech quality after decoding. Also most of the articles are directed to lower the rate of CELP coded speech transmission. One problem related to low rate CELP speech transmission with a good speech decode quality is noise in the channel for transmission. The goal of this article is to combine the advantages and the possibilities of CELP speech coding to reduce the rate of transmission and the methods of channel coding to protect the most important CELP coding parameters in each speech frame such as line prediction coefficients, excitation indexes etc.**

Keywords – **Source-channel Coding, CELP Coding, Speech coding, Speech Processing**

I. INTRODUCTION

The Shannon communications coding theory is based to the assumption of separately and sequentially source coding (compression) and channel coding (error protection) [1]. This is true only in the case of asymptotically long block lengths of data. In many practical applications, such a speech coding and especially in CELP coding [2,4] this condition is not satisfied perfectly. Thus, it is the goal of this article to propose a joint source-channel CELP coding method. Also, to it is interesting to study the proposed method and to show the effectiveness from the combination of the advantages of transmission rate reduction with CELP speech source coding and the error protection with channel coding of the CELP coded parameters.

The reason for the joint source-channel CELP coding is the presence of inter and intra-frame redundancy in the CELP coded and quantified parameters. Also it is known, that the importance of each of these parameters for the quality of the decoded speech is different. This fact can be exploited to achieve an additional effectiveness in the joint source-channel CELP coding. The theory of joint source-channel is based of the proposition to design jointly the source and channel coding operation. This means the channel coding, which is mainly used for error protection work not independently from the source coding for speech in CELP method – the statistic characteristics of CELP coded and quantified parameters representation for each CELP speech coded frame.

[1]Snejana G. Pleshkova is with the Faculty of Telecommunications, bul.Kl. Ohridski 8, Sofia, Bulgaria, E-mail: snegpl@tu-sofia.bg

II. THE JOINT SOURCE-CHANNEL CELP MODEL

The proposed model for joint source-channel CELP coding is shown as a sequence of block in Fig.1. The speech signal source is considered as a sequence of frames, each with 30 ms duration (240 samples for 8 kHz sample frequency).

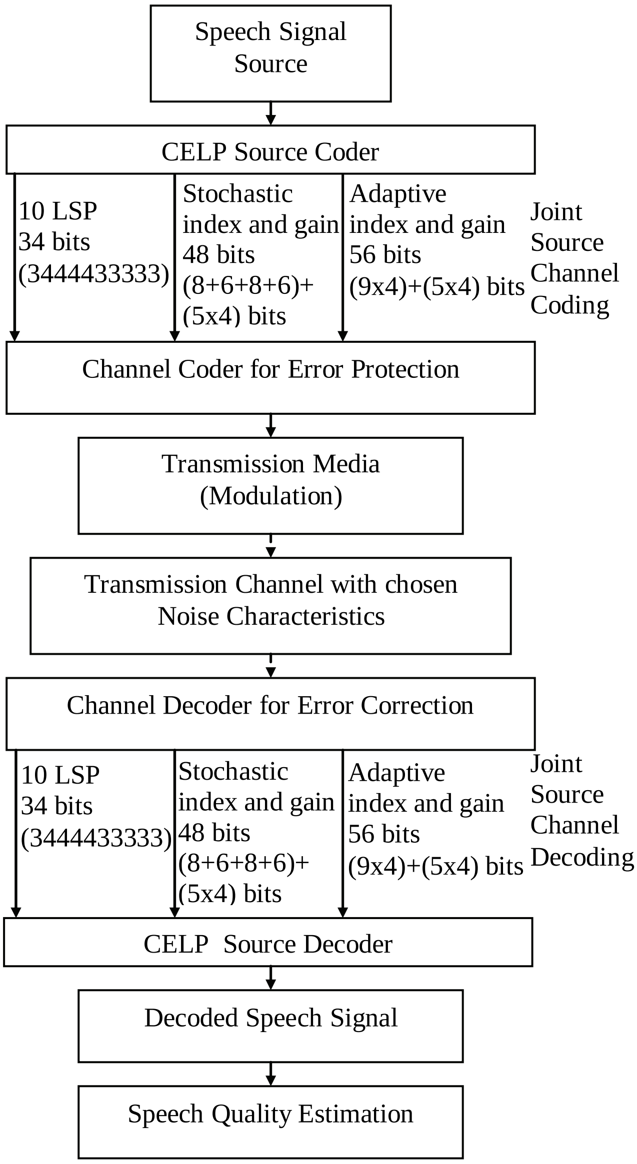

Fig.1.

After block of CELP Source Coder are presented the CELP coded parameters for one frame: 10 LSP – Line Spectral Pairs; Stochastic code book index, Stochastic code book index and gain; Adaptive code book index and gain. More detailed explanation of these parameters and their bits allocation is shown on the Table 1. Next is the Channel Coder for Error

Protection, which can be made to work independently or joint with the CELP Source Coder. The coded information is input to Transmission media with a chosen type of modulation. The transmission channel depending from the type of modulation, but also it is necessary to choose a type of Noise Characteristics in the channel of transmission.

TABLE I

	LSP	Stochastic	Adaptive
Update time	30 ms (240 samples)	30/4=7,5 ms (60 samples)	30/4=7,5 ms (60 samples)
Bit per frame	34 bits [3444433333]	Index:8+6+8+6 Gain: 5x4	Index: 9x4 Gain: 5x4

In the Receiving Media the received signal is demodulated and then it is put to the Channel Decoder. At the output of the Channel Decoder are the decoded CELP parameters. They are used in CELP Source decoder to restore finally the speech decodes signal. In the Fig.1 is added also a block for Speech Quality Estimation. This can be use for comparatively evaluation of speech quality with and without joint source-channel CELP coding.

III. JOINT CHANNEL CODING OF CELP PARAMETERS

The assumption, that CELP coded parameters are redundant, gives the reason to propose a joint source-channel coding of these parameters both, for to error protection and for using in this channel coding their redundancy to improve the effectiveness jointly from the source and channel CELP coding.

It is chosen to use a common applied in channel coding family of so called rate compatible punctured convolution codes (RCPC) [3]. With these codes is possible to achieve higher rate $R = k/n$ toward the conventional code with rate $R = 1/n$, where k and n are the number of input and output bits, respectively. These codes can be performed easy from the basic a puncturing matrix. The basic convolution code is periodically performing with this matrix. For example to produce a punctured convolution code with rate $R = P/(P+\delta)$ and period P from a basic code with rate $R = 1/n$ it is used a punctured matrix $A(\delta)$ with dimension $(n\,x\,P)$ and $\delta \in [1,(n-1)P]$. For $n = 2$ basic code have a rate $R = 1/2$ and for a period $P = 4$ are possible four different code rates $R = 4/5, 4/6, 4/7$ or $4/8$ and the punctured matrix is:

$$A(\delta)_{2x2} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad (1)$$

For the basic code with length 3 the generator matrix is:

$$G(D) = \left[1 + D^2, 1 + D + D^2\right]. \quad (2)$$

With (1) and (2) can be performed a resulting code with rate $R = 2/3$ and with a generator matrix:

$$G(D) = \begin{bmatrix} 1+D & 1+D & 1 \\ 0 & D & 1+D \end{bmatrix}. \quad (3)$$

The corresponding trellis for this punctured convolution code is shown in Fig.2 for period $P = 2$, output bits X0, X1 and “X”, witch means punctured bits.

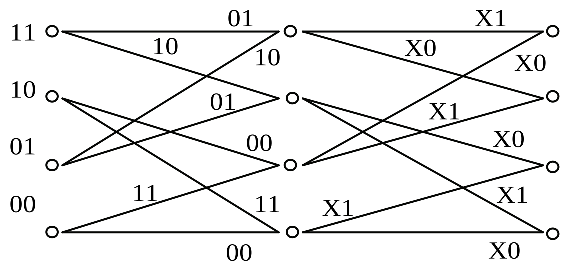

Fig.2.

More generally for basic code with rate $R = 1/n$ and period P the punctured matrix is:

$$A(\delta) = \begin{bmatrix} a_{11}(\delta) & \dots & a_{1p}(\delta) \\ . & & \\ . & & \\ . & & \\ a_{n1}(\delta) & \dots & a_{1p}(\delta) \end{bmatrix}, \quad (4)$$

where

$$\{a_{ij}\}(\delta) \in \{0,1\}, 1 \le i \le n, 1 \le j \le P^2, \text{and} 1 \le \delta \le (n-1)P. \quad (5)$$

The advantages of punctured convolution codes are less complexity in relation to non-punctured codes and this is reason for choosing them in the proposition to investigate the joint source-channel CELP coding in this article.

The start point for channel coding in Fig.1 is after CELP source coding, when are calculated each frame the CELP coding parameters: 10 LPS, 4 pitch gains, 4 pitch delays, 4 codebook gains and 4 codebook indices.

There are some different possibilities to perform channel coding:

- without using the channel coding;
- independent channel coding from the source CELP coding;
- joint source – channel coding equal for all CELP coding parameters;
- joint source – channel coding unequal or partial for some of CELP coding parameters, towards their importance for the quality of the CELP decoded speech.

Also it is necessary to decide what type or model of channel noise to use CELP coded parameters are affected from the chosen channel noise.

All of these mentioned conditions and possibilities can be modeled and tested and some of them are presented shortly in this article and especially the decoding process from which is depending mainly the quality of the decoded speech.

IV. CELP Parameters Decoding

CELP parameters are sending over a memory less channel. In the receiver the channel decoding of CELP parameters is based the Viterbi algorithm, which chooses the received code sequence:

$$\hat{x}^k = \left(\hat{x}_1, \hat{x}_2, ..., \hat{x}_k \right), \tag{6}$$

that minimizing

$$P_r\left(y^k \middle| \hat{x}^k\right) P_r\left(\hat{x}^k\right), \tag{7}$$

where

$y^k = \left(y_1, y_2, ..., y_k \right)$ is the received sequence of length K, equal to the number of transmitted CELP parameters.

For applying of these equations (5) and (6) it is necessary to define or choose the type of modulation used in transmitter and also the model of the channel noise. Here it is proposed to apply BPSK modulation with AWGN and fully interleaved Raleigh Fading channels and with noise variance $N_0/2$. For this case the minimization of:

$$\sum_{k=1}^{K} \left\| y_k - a_k \hat{x}_k \right\|^2 - N_0 \ln P_r\left(\hat{x}_k\right) =$$

$$= \sum_{k=1}^{K} \left[\left\| y_k - a_k \hat{x}_k \right\|^2 - N_0 \ln P_r\left(\hat{x}_k \middle| \hat{x}_{k-1}, \hat{x}_{k-2}, ... \right) \right] =$$

$$= \sum_{k=1}^{K} \left[\left\| y_k - a_k \hat{x}_k \right\|^2 - N_0 \ln P_r\left(\hat{u}_k \middle| \hat{u}_{k-1}, \hat{u}_{k-2}, ... \right) \right], \tag{8}$$

where

a_k is the sequence of Raleigh Fading coefficients, which are assumed to be available at the decoder. If the channel is chosen as AWGN, then a_k is the all-one vector for all k;

$P_r\left(\hat{u}^k\right)$ - Markov model of CELP parameters;

$\hat{u}^k$ - presentation of CELP parameters as random process.

There are some different possibilities to represent CELP parameters as Markov model of chosen order: first, second etc. Usually it is chosen the first-order or second-order Markov model and each ***i***-th CELP parameter are described as a $\{U_{i,j}\}$ random process, where j is the current frame number from the sequence of frames, defined in speech signal. If the number of CELP parameters in a frame are "l", then

$$U_j = \{U_{1,j}, U_{2,j}, ..., U_{k,j}\} \tag{9}$$

is the random representation of CELP parameters in ***j***-th frame.

The first-order or second-order Markov models of the CELP parameters as random process $\{U_{i,j}\}$ can be used for estimation of their entropy rate and redundancy define and minimize the probability from the equation (6). This equation is related to the first-order Markov model chosen for the transmitting CELP parameters:

$$P_r\left(U_j = u_j \middle| U_{j-1}, = u_{j-1}, ..., U_1 = u_1\right) = P_r\left(U_j = u_j\right) \tag{10}$$

and

$$P_r\left(U_{i,j} = u_{i,j} \middle| U_{i-1,j} = u_{i-1,j}, ..., U_{1,j} = u_{1,j}\right) =$$

$$= P_r\left(U_{i,j} = u_{i,j} \middle| U_{i-1,j}, = u_{j-1}\right) = P_F^{(i)}\left(u_{i,j} \middle| u_{i-1,j}\right), \tag{11}$$

for

$i = 1, 2, ..., l$ and $j = 1, 2, ...,$

where

$P_F^{(i)}$ is the probability, defined for first Markov model.

The second-order model also can be described and the advantage of this model is the possibility to use not only the intra-frame, but also the inter-frame redundancy:

$$P_r\left(U_{i,j} = u_{i,j} \middle| U_{j-1}, = u_{j-1,}, ..., U_1 = u_1\right) =$$

$$= u_{i,j}, ..., U_{i-1,j}, j = u_{i-1,j} =$$

$$= P_r\left(U_{i,j} = u_{i,j} \middle| U_{i-1} = u_{j-1,j}, U_{i,j-1} = u_{i,j-1}\right) =$$

$$= P_S^{(i)}\left(u_{i,j} \middle| u_{i-1,j}, u_{i,j-1}\right), \tag{12}$$

for

$i = 1, 2, ..., l$ and $j = 1, 2, ...;$

where

$P_S^{(i)}$ is the probability, defined for second-order model.

V. The Simulation of Joint Source-Channel CELP Coding

The proposed method of joint source-channel CELP parameters coding is simulation as a Matlab Simulink model, presented in Fig.3.

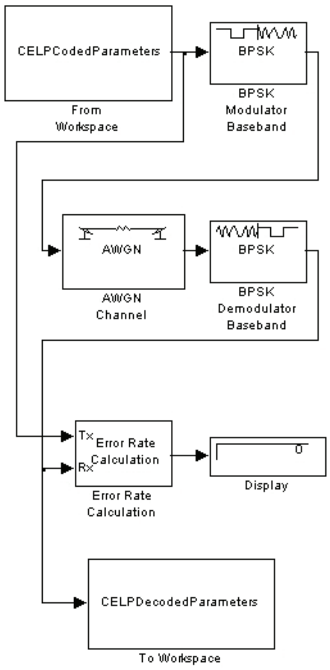

Fig.3.

The CELP parameters are coded using a Matlab program, which work in two modes: with and without joint source-channel coding. From Matlab workspace these calculation and coded CELP parameters are input in the Simulink model shown in Fig.3. It is chosen for transmission, as it is mentioned above, ordinary digital communication system with BPSR modulation and demodulation and with a channel AWGN model. There is shown also in Fig.3 the block for Error rate calculation and estimation. The received and decoded CELP parameters are returned back to Matlab workspace, where they are used a program for speech quality estimation comparing CELP decoded speech with and without joint source-channel coding and the speech signal before CELP coding. Some of the results for male and female speech testing signals and for different SNR E_b / N_0 are shown in the Table II in percentage for perfect quality (without listening quality degradation) and for some quality losses.

TABLE II

SNR (E_b / N_0)	Coding method	Relative quality
-2 dB	Without joint coding	96%
1 dB	Without joint coding	84 %
-2 dB	With joint coding	88 %
1 dB	With joint coding	97 %

VI. Conclusion

The results in the Table II shown and improvement of decoded speech quality when it is used the proposed joint source-channel CELP coding in the cases with a larger SNR values. The briefly presented results show only a little part of the experimental studies, which are made to investigate all properties of the proposed joint source-channel CELP coding for speech signal. Of course it is necessary to continue in some future works these testing procedures in the direction of subjective evaluation of the joint source-channel CELP speech coding quality in comparison with traditional CELP coding

References

[1] C. Shamon, “A mathematical theory of communication”, Bell Sys. Tech. Journal ,vol. 27, pp. 379-423, 1948.

[2] J. Campbell, T.E. Treman, "The Federal Standard 1016 4800 bps CELP voice Coder", Digital Signal Processing, Academic Press, vol.1, no.3, pp.145-155, 1991.

[3] J. Hagenauer, “Rate-compatible Punctured Convolutional Codes (RCPC Codes) and their Applications, IEEE Trans. On Commun., vol.36, no.4, pp.389-400, April 1988.

[4] Sn. Pleshkova, M.Momsthedjikov, “Fast CELP code book search method”, Proceedings of the International Scientific Conference on Energy and Information Systems and Technologic, pp.178-182, Sofia, 2003.