Applied Aspects In Static Images Processing

Gergana Markova[1]

Abstract – **The content extraction from static images is an actual investigation task, because in the modern information networks the amount of information, which is stored, shared and processed under this form, is vastly increasing**

***Keywords* – Image metadata, Contextual metadata, Image Preprocessing.**

I. Introduction

The metadata (MD) could be examined as one of the ways for formal description of the information content, presented by one or more digital storage devices. The scope and structure of the metadata is determined by the nature of information processing and expected output result. In this sense we can talk about structure of the metadata, oriented to the nature of information presenting to the place, mechanism and characteristics of introduction presenting as well as to the opportunities for further processing. When setting a task for image recognition it is often skipped the stage at which there may be put limits on the input data.The inclusion of this intermediate stage is directly dependent on the fact that each of the algorithms are defined sets of constraints that determine the effectiveness of the algorithm. The MD from an image can participate in the preprocessing stage and through research and definition of certain algorithms, non-structural sets of images can be dynamically classified or limited on the base of MD. In this sense we can define a specific research task, which aims to analyze the necessity of adding one functional layer in the process of imaging called preprocessing, in which the MD of the image will be included. This is shown in Fig. 1.

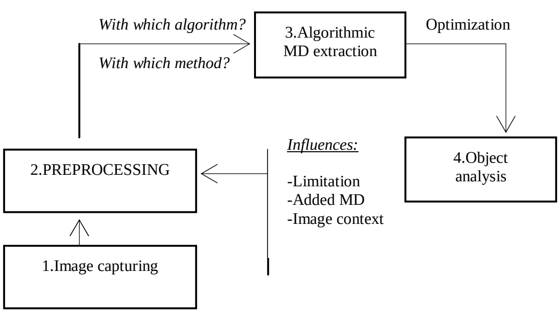

Fig. 1. Stages Image Processing with intermediate layer Preprocessing.

II. Explanation of the stage preprocessing

1. Image processing by the human visual system

In order to be accepted, processed and classified the environment or only object from this environment, it is projected into human's mind since the first weeks of his birth. Nixon and Aguado [1] described the human visual system as sequence of three models: 1. The eye as a **physical model**; 2. The neural system as **experimental model**; 3. The image processing by the human brain as **psychological model**. The Human vision can cope well with the relative distance, but the problem is the absolute distance while in the direction Computer vision is on the contrary. These three steps, as models and as stages determine this configuration as a preprocessing of the image from the human senses. This information is based on the logic used in basic research and development directions of Image Processing and Computer Vision.

2. Classical approach for image preprocessing

One classical approach in the filed of computer vision is described in [2, 3, 4] and includes:

- ✓ Digital presentation, filtration of the image;
- ✓ Separation of borders and edges for the objects in the image by usage of different operators Robert's, Prewitt's, Sobel's, Huckel's, Canny's, Hough's, Walsh's, Adamar's, Karunen-Loev's and transformation;
- ✓ Methods for discovery of lines in binary images. Segmentation.

In terms of processing of visual information, these authors define the discrete representation of the image and the above described stages for low-level stages. These low-level stages are discussed as *preprocessing*.

3. Preprocessing, applied in identification of moving object

In the behavioral biometrics one of the sustainable methods of passive and automated extraction of distinguishing features of gait, is usage of low-resolution video material [5].

The main concept of the preprocessing is shown in Fig. 2

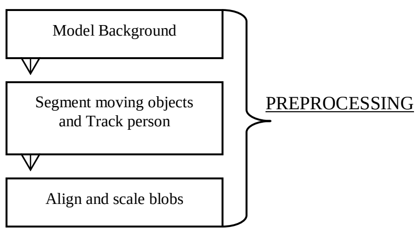

Fig. 2. Image Preprocessing in identification of moving object.

In this case in order the silhouette parameters to be extracted, the noise is decreasing and the signal is increasing. In order the input parameters to be correct for the observed object it is important to be situated far from the camera. In this case the process of segmentation is realized by a technique for separation of the background with great stability in light changing. On the other hand, in the identification it is used the proportions of the human body, based on which is applied

[1]Gergana Markova, Faculty of Mathematics and informatics, 3 arch. G.Kozavov str., bild 3, 5000 Veliko Turnovo, Bulgaria, e-mail: gerganavioletova@abv.bg.

algorithm which would not give result in incorrect image capturing, i.e. using MD from the hardware.

III. Contextual information and metadata generation

1. Categorization of MD according to the context

Modern smart mobile devices generate a lot of context information. This type of MD can be categorized into several categories:

✓ Computing context describes the available resources of the devices. The terminal profile defines the technical features, including: • Type and size of the screen; • Support functions, such as WAP and Java; • Maximum size of memory for WAP decks and Java MIDlets; Support of different content types.

✓ User context contains the user profile like location, close friends, social status, opinion about various subjects and personal interests.

✓ Physical context describes lighting, noise levels, traffic conditions, temperature. The information is used to identify gestures and movement, positioning of the device.

✓ Timed context contains the current day, hour, week, month, year.

✓ Computing context is used to adapt content and transfer it.

Bluetooth and GPS context is generated by mobile phones or small multisensory digital cameras, which accurately locate and retrieve location and events by automatically capturing, at certain intervals [6, 7]. By using sensors are detected changes in the levels of luminance, motion, room temperature. The cameras of the modern mobile phones allow collection of MD in terms of image capturing. Systems are realized, which collect information about the location of the captured image, as well as the use of contextual MD in terms of capturing. The idea is the use of spatial-temporal-social context, in order the location of the captured object to be identified in sharing platform. It is used simplified faceted metadata hierarchy [8, 9]. The data are the object in the image, that we are interested in, its sharing with other users and comparison with already available objects in a database stored on the server [10]. Over the past decade, research has focused on the description of algorithms for extracting content from low-level metadata i.e. from the hardware Fig. 3. According to the data quality from the Onboard light sensor and Motion sensor, functional analysis and classification for further user processing are structured.

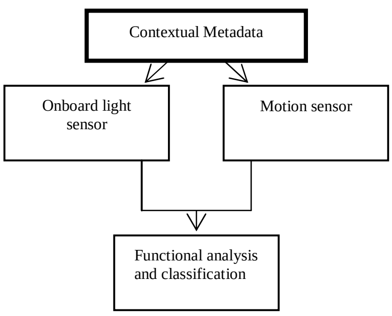

Fig. 3. Process of processing of Contextual Metadata.

The producers allow by manually adding of data by user to be shared and identified MD, which include semantic information. Today the objective is the extraction of content to be completely automated process and to exclude the participation of users.

2. Devices as sources for contextual metadata.

Some hardware devices could generate metadata with the help of software. With them we can understand the conditions under which undergo the process of fixing the image. The most information-rich is the standard RAW. The ROW format has its own variations and each producer supports his own decisions: .raf (Fuji), .crw, .cr2 (Canon), .kdc (Kodak), .mrw (Minolta), .nef (Nikon), .orf (Olympus), .dng (Adobe), .ptx, .pef (Pentax), .arw (Sony), .x3f (Sigma). If the input data are précised and additionally processed the images can “candidate” for inclusion in one or another algorithm, which aims extraction of definite context for further analysis.

3. Preprocessing in Medical diagnostics. Image diagnostic.

In the image diagnostics in visualization of images are generated shapes of objects by discovering of photons or the usage of electromagnetic waves. When forming the image in medical devices, some of the most important features are *noise level* and image resolution [11, 12, 14]. In these static and video images for the purpose of research, forecasting and diagnostics is necessary the search for a definite type of contextual content. The development of techniques for detection and identification of constituent views in each video, extract clinically important frames, and generate time-compressed clinical summaries, solve specific tasks using metadata - generated and added [9].

3.1. **CT** (Computed tomography) have three advantage:

- Compression of 3D volume of the structure of the human body into a 2D image, where the over and underlying tissues and structures overlap, leading to reduced visibility and contrast.
- The entire volume can be reformatted into one of three positions - sagittal, coronal and axial, depending on what projection should be seen the image.
- The high level of sensitivity of CT, depending on the level of scattered photons ensures improving of the effectiveness. In sequential CT is used graphical matrix 512x512 and 16-bit resolution.

3.2. **MRI** (Magnetic resonance imaging) works in the radio-frequency range. As an advantage it is shown the excellent soft tissue contrast, which can be managed as transmitted by the device input signal.

3.3. **SPECT** (Single-photon emission computed tomography) - gamma camera operates in two modes – for a flat image, capturing an image projection of the body or part of the body and SPECT mode by using different angular distances. In this mode, single or multiple detectors revolve around the patient by defined angular steps. At each step, usually is 64x64 or 128x128 in 2-D projection. These projections are reconstructed into 64x64x64 or 128x128x128 topographic volumes in order to be visualized. One serious problem is the process of creating and restoring of lost resolution and contrast enhancement.

SPECT is a rapidly changing field, and the past several years have produced new developments in both hardware technology and image-processing algorithms [12]. New clinical devices include high-count sensitivity cardiac SPECT systems that do not use conventional collimation and the introduction of diagnostic-quality hybrid SPECT/CT systems. While there has been steady progress with reconstruction algorithms, exciting new processing algorithms have become commercially available that promise to provide substantial reductions in SPECT acquisition time without sacrificing diagnostic quality.

The γ-ray emissions collected by the SPECT systems are not linearly related to the ray sums of activity in the patient because of tissue attenuation. As a consequence, if these attenuated projections are reconstructed, the resulting tomographic slices contain artifacts and will not accurately reflect the true internal distribution. Factors other than attenuation limit the quality of reconstructed images. These include spatial resolution, scattered radiation, and statistical fluctuations (image noise).

3.4. **NMR** (Nuclear magnetic resonance) - in result of the used technology, it is obtained an image with dimensions 256x256, fixed pixel size 1 mm^2. One problem is the high relation signal/noise, as the noise has a different character. The temperature noise caused by heating of tissues, which deforms the incoming data. Such a problem causes the movement of the patient during the session, heartbeat, breathing, peristalsis, fatty deposits can cause pressure or displacement result is the same.

By using each technology image reproduction for medical purposes, the direction in which they are improved is the removal of fixed problems arising from the used devices and the subject of study. One way to improve the output image is additional image processing before the visualization. The assessment of the image as final result is still being analyzed by a specialist, not automatically. There are already thousands of libraries with available images used for training of specialists and for the needs of telemedicine. Problem remains their classification.

3.5. **Some approaches.**

The basis of most tomographies is the idea that the internal structure of the object can be represented by a series of parallel cross sections. The method of obtaining the 2D tomographic image has two stages: first stage of forming the projection data, the second stage - the projection data to recover (reconstruction and restoration) image cross sections (for example using Fourier theorem for cuts). For the Image Preprocessing in the first phase, incoming data depend on the technical characteristics of the device and setup options. They define the structure in the second stage. If they are not correct, the result of the reconstruction phase would not be correct, this applies to CT and NMR. Therefore, Medical Image processing is considered as one of the most complex areas. In PET and SPECT the actual image is the output of an inverse reconstruction algorithm. It is important to keep in mind, that these kind of images have a very high noise level (10% and more). This makes the interpretation of these images very difficult, especially if the image is used as input of a pattern classification algorithm. In medical imaging such an algorithm will normally try to find the type of tissue at each pixel or voxel.There are used approaches for Medical Image Preprocessing that include filters which are applied to remove noise while preserving semantically important structures such as edges - methods based on nonlinear Partial Differential Equations (PDEs) [13]. It is used for medical images such as mammograms, CT and MR images.

4. Preprocessing in Character Recognition.

Text recognition relates to many fields of implementation - scanning of documents and card indexing, reading a text from photos, automatic reading of meta data from documents, recognition of car registration numbers, numbers of wagons, parcels, road signs, land marks. The quality of meta data of the input data in document scanning is related to brightness, contrast, legibility of symbols, size of scanned image and file format, in which the file is saved. If this step of Preprocessing is neglected, further processing for recognition and classification will be with low reliability. Applied algorithms either will not work out, or will register objects less in number. In one of the methods every character set is a set of templates [14]. On Fig 4. are illustrated several defined stages of sub-divided recognition process:

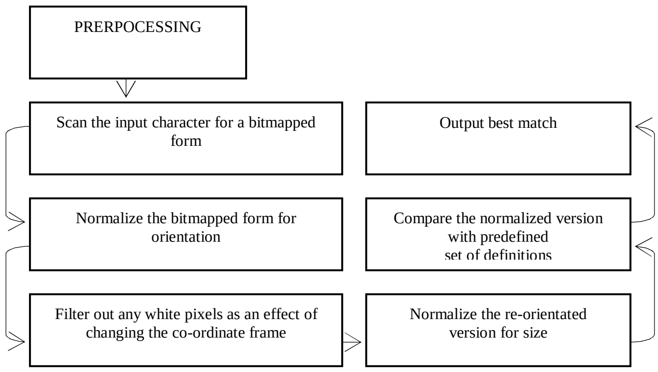

Fig.4. Process of Character recognition in one of the algorithms for recognition.

In the recent past, one software could read, for example, British *license plates* only, other could read plates from Hong Kong only, etc. It was not accidental: the geometry of the plate, as well, as its syntax, were essential parts of the plate reader software. Without the presumption of a fixed plate geometry (character ratios, character distribution, font type, plate’s colour, etc.) and a well defined syntax, the algorithm may no even find the plate on the picture. Nowadays, a good algorithm should read all plates from Europe with the same level of quality. There are indeed a wide variety of plate types in Europe:

- black (dark) characters on white (light) plate;
- white characters on black plates;
- one-row plates;
- two-row plates;
- plates with different character-size;
- latin and cyrillic fonts;
- plates with or without region's shield or special mark, etc.

The authors of one such software [17] offer three additional sub-algorithms before the OCR:
-a plate localizing sub-algorithm, responsible for finding the plate on the picture.
-a contrast/brightness normalization sub-algorithm, responsible for equalizing the plate picture.
-a character segmentation sub-algorithm, responsible for finding and cutting out the individual characters on the plates and pass them to the OCR.

No doubt, the better the quality of the input images are, the better conditions the license plate recognition algorithm has, and thus the higher license plate recognition accuracy can be expected to be achieved. In order to expect reasonable results from a plate recognition algorithm, the processed images should contain a plate: with strictly defined resolution; with good sharpness; with high contrast; under good lighting conditions; in a good position and angle of view [15, 18]. Things described above, and attached examples prove the claim, that Image metadata are of great importance as input data and are of a crucial significance for the performance (or not) of the core License Plate Recognition.

IV. CONCLUSION

In [17], we suggested one approach for increasing the efficiency of algorithms for metadata extraction from static images with aim to expand the volume and content of generated MD and enabling search of content in large libraries of static images. Extraction of contextual content from MD in an image is an important step for its adaptation and transfer for classification.

The preprocessing is an important stage of images processing that direct the correct data input, meeting certain conditions, to correct preprocessing algorithm.

REFERENCES

[1] M. Nixon, A. Aguado, Feature Extraction and Image Processing, First edition 2002
[2] М. Стоева, Ю. Петкова, Обработка на визуална информация, Варна, 2009
[3] Г. Гочев, Компютърно зрение и невронни мрежи, ТУ, София, 1999
[4] D. Marr, E. Hildreth, Theory Of Edge Detection, Psychology Department And Artificial Intelligence Laboratory, Cambridge, 1980
[5] D. Sharmila MCA, Mphil, Dr. E.Kirubakaran, Member, IACSIT, Image and Formula Based Gait Recognition Methods, International Journal of Computer and Electrical Engineering, Vol. 2, No. 2, April, 2010
[6] D. Byrne, B. Lavelle, A. Doherty, Gareth J.F. Jones, A.F. Smeaton, Using Bluetooth & Gps Metadata To Measure Event Similarity In Sensecam Images, CDVP & AIC, Dublin City University, Ireland, 2007
[7] J. Huang, C. Tsai, Omniguider-The Next Generation of GPS Navigation System, International AR Standards Meeting-June 15-16 2011
[8] R. Sarvas, E. Herrarte, A. Wilhelm, M. Davis, Metadata Creation System for Mobile Images, MobiSYS’04, June 6-9, 2004, Boston, Massachusetts, USA, ACM 2004
[9] D. Marc, Editing Out Video Editing, IEEE MultiMedia, 10 (2). 2-12, 2003
[10] D. Marc, S. King, N. Good, R. Sarvas, From Context to Content: Leveraging Context to Infer Media Metadata, University of California at Berkeley School of Information Management and Systems, USA
[11] O. Demirkaya, Image processing with MATLAB: applications in medicine and biologyet, CRC Press Taylor, 2009
[12] M. Madsen, Recent advances in SPECT imaging. / Nucl Med 48: 661-673, 2007
[13] J. Weickert, Ch. Schnцrr, PDE-Based Preprocessing of Medical Images, Department of Mathematics and Computer Science University of Mannheim, Germany, 2000
[14] F. Hussain, J. Cowell, Character Recognition of Arabic and Latin Scripts, Department of Computer & Information Sciences, De Montfort University, 2005
[15] Hsien-Chu Wu, Chwei-Shyong Tsai, Ching-Hao Lai, A License Plate Recognition System In E-Government, National Taichung Institute of Technology, Taiwan, 2006
[16] Е. Витальевна, И. Сафонов, Способ повышения качества цифрового изображения на основе метаданных, описание изобретения к патенту 09.06.2010
[17] G. Markova, O. Asenov, M. Todorova, One approach for increasing the efficiency of algorithms for metadata extraction from static images, ICEST 2011, Nish, Serbia
[18] http://www.platerecognition.info/