International Conference on Computer Systems and Technologies - *CompSysTech'2003*

Fast, Shape Based Image Retrieval [(*)]

Dimo Dimov

Abstract: *The paper presents a method for fast access by content into database (DB) of images. The method proposed is based on the image tree of contours defined for graphic images, as well as on the one-dimensional complex Fourier transform of the contours. In this way, problems connected with the image normalization concerning translation, rotation, scaling, reflection and intensity are currently solved. The most informative data of the image are ordered by importance in a key of fixed length, on which the fast access is performed using the well-known index access methods of a conventional DB management system (DBMS). The method is tested on a DB of about 4000 images of hallmarks. The experimental system EIRS, intended for the Bulgarian Patent Office practice, is also described here in brief.*

Key words: *content based image retrieval (CBIR), fast image retrieval, representation of graphic images, Fourier's transform of contours, image databases.*

INTRODUCTION

Content Based Image Retrieval (CBIR) is a new area of Informatics [1, 13] covering techniques for automatic (or automated) retrieval of images from database of images (IDB). The essential content of search (and retrieval) is usually represented by simple image features as color, texture, shape, movement, and/or structures of them. Naturally, CBIR draws many of its methods [13] from the field of image analysis, computer vision, statistics, and pattern recognition [14].

A few CBIR systems are known in world practice as GEMINI, QBIC (IBM), CORE, Image Engine (Virage), Image Retrieval Ware (Excalibur), ARTISAN [1, 13, 17, 18]. A larger list of CBIR systems and/or IDB-s can be also found in [15]. And they all should be attached to the so-called "early", i.e. the contemporarily CBIR [13], because of being based predominantly on image "primitive" features and/or structures [1] as color histograms, directional histograms of image simple sub-objects (edges, curvatures, representative points), elastic configurations of image sub-objects, etc. The approaches of more detailed representation of shape are avoided in principle because of classical difficulties with image segmentation even in cases of low noise [17]. The recent CBIR developments are no exception of these rules, e.g. [2, 16].

A general characteristic of CBIR systems [13] and available IDB as well [15] is that they are developed over conventional DB management systems (DBMS), which methods for fast data access by indexes are optimized for text, number and logical data. The known extensions of these methods for multimedia data (R-trees, k-trees, etc. [13,18]) do not respect the CBIR specifics of image recognition. Generally CBIR systems tend to a sequential data access, what is unacceptable, especially in cases of large IDB.

In this way the interest in CBIR naturally arises towards approaches better adapted to the index access techniques of conventional DBMS. A system called EIRS [4, 5, 6], has been designed to support experiments in this aspect on a real IDB of about 4000 images of hallmarks. One of the three approaches of EIRS, namely a new heuristics for fast CBIR based on full contour data of the images is considered in the paper.

THE EXPERIMENTAL SYSTEM IN BRIEF

At present, the EIRS (an Experimental Image Retrieval System) is to be mainly considered as an instrumental software for experimenting three CBIR techniques, namely:

(**T1**): using a two-dimensional wavelet transform,
(**T2**): using a two-dimensional Fourier transform, and
(**T3**): using a heuristic decomposition of the images to contextual contour parts.

[(*)] This research has been elaborated in the frames of the project "Access Methods for Image Databases", partially supported by research funds of BAS, and maintained by Institute of Information Technologies (IIT No. 010043).

Fig.1. The EISR search engine's table of results for given target image: two rows of images by content similarity are retrieved, before and after the target.

Although applying of different recognition principles (and feature spaces) the above three techniques are built on the following common statements [4, 5, 6]:

- The search content is the input image itself or a sketch of it.
- The most essential image data are automatically extracted and arranged in a key string of a fixed length (following the chosen technique for key derivation – T1, T2, or T3).
- The fast access is performed on this key data using conventional index access methods of the given DBMS.

As already mentioned, the T3 technique will be generally considered hereinafter.

What the EIRS generally performs, as illustrated in Fig.1, is to retrieve fast a number of IDB images that are close enough by their content to the input target image. Of course, to provide the necessary experimental environment, the EIRS keeps also a full option of an IDB browser, image-editing subsystem, specific tools for experiment maintenance, etc.

The EIRS is C++ written, on Borland C Builder v.5.0, and respectively, the DBMS used is Paradox. The IDB of EIRS is organized as a table of pointers to the image files (BMP and/or JPG), which are stored outside the Paradox DB part (cf. Fig. 2), in a separate folder "Images".

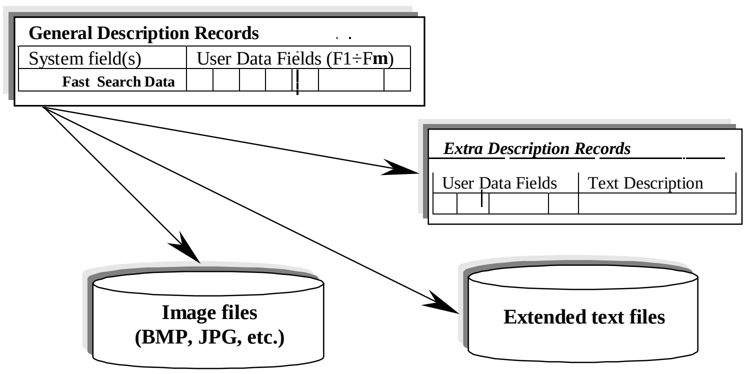

Fig.2. A simplified logical scheme of the IDB of EIRS

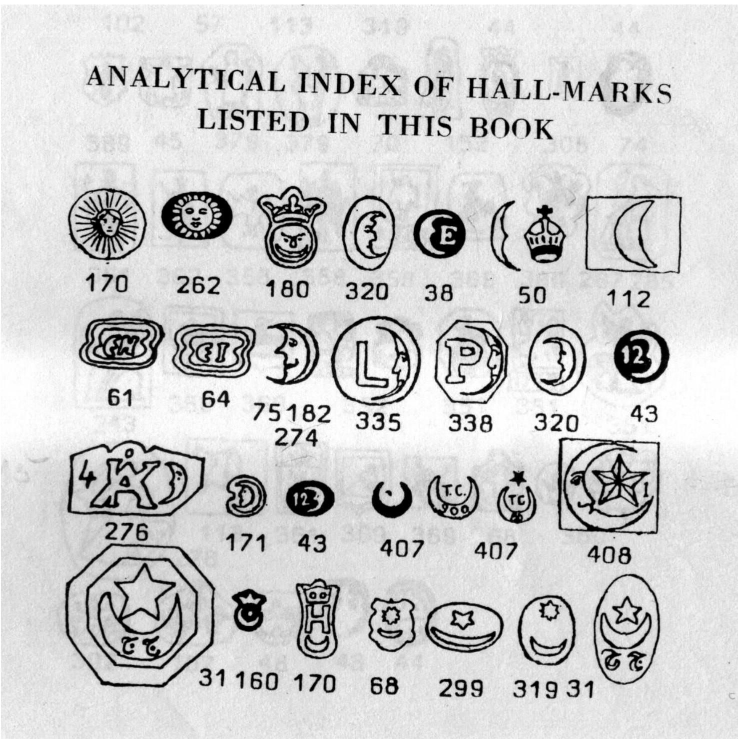

Fig.3. The first 29 images of hallmarks (from the all 3937 ones loaded into the IDB of EIRS) as they are semantically ordered in [8].

For the time being, the images the EIRS operates are graphic images, gray scale or true color and not very noisy. The system is tested on an IDB of about 4000 half-toned images of hallmarks, scanned from [8], cf. Fig.3.

Recently, about 150 color trademark images from the Bulgarian Patent Office practice [11] were additionally loaded into IDB of EIRS.

THE ESSENCE OF THE METHOD PROPOSED

The proposed method for fast IDB access for a shape based similar image retrieval combines two strategies:

(1) to represent given images by an appropriate informative structure (e.g. a multidimensional feature space), and

(2) to use this structure in the "fast retrieval by key" paradigm of conventional DBMS. In brief, the idea is to represent the essential image content as a well-structured DB key of fixed length.

As it is well known [9, 10, 18], all fast DB access methods work using the convention of a "key" and give the result of DB search as follows:

$$\langle \text{the key value of the retrieved object (record)} \rangle \prec \langle \text{the key value of the searched object (record)} \rangle ,$$

where "⟨i⟩ $\prec$ ⟨ii⟩" means that ⟨i⟩ is the greatest value less than or equal to ⟨ii⟩. The key usually corresponds to a feature (or a combination of features) defined for all DB objects of a given type. As a rule, key fields should be simple fields of fixed and not very long length.

Intuitive Requirements to the Image Key Performance

- The key should consist of the (most) essential information of the image.
- This key information should be of significantly smaller volume than the image itself.
- The key information volume should be written in a 1/D array of fixed length.
- The key information should be structured in such a way that essential parts to appear in frontier positions.
- The image key is not obligatory to be unique, i.e. for a given key (of given length) there can be more than one image to correspond.
- All DB images of one and the same key should be considered as equivalent ones.
- The key length should be chosen in such a way that the image key to be informative enough.

Thus, the following intuitive definition about key can be written [5, 6]:

A definition**:** The *image key* or simply the *key* should consist of the most essential information of given image, structured in descending order into a one-dimensional (1/D) array of fixed length.

The Tree of Image Contours

The images of EIRS interest (T3 technique) are considered as compositions of one-connected regions (i.e. without any holes) placed one over another, and each of them filled by its respective color (or half-tone). This representation can be arranged in such a way that each two regions to be either entirely included one in another or non-crossing ones.

Thus, a given image can be represented by all its "colored" contours structured in a tree of inclusion, and this representation is usually called "*tree of image contours*" (TIC), cf. Fig.4. Besides, for the EIRS general case of two colors' images (a black object over white background), it is obvious that this representation is isomorphism. Furthermore, we can consider as *noise* all the differences between the original image and the reversed one (obtained after respective filling of contours). Because of large diversity in similar structures [12, 14], the relation between TICs and the so-called homotopic trees [14] is also illustrated in Fig.4.

Finally, from the viewpoint of image key definition here, the following can be intuitively said about the essential image data and their ordering by importance:

- The closer a given contour is situated to the root of the TIC, the more important it is.
- The most important contour among the ones at a given level should be the greatest area contour, or perhaps the curliest one, or the one of longest perimeter per an area, etc.
- The root can be only one in case the image frame is considered also a contour. Of course, it does not mean the root is an important contour, but its primary successors should be.

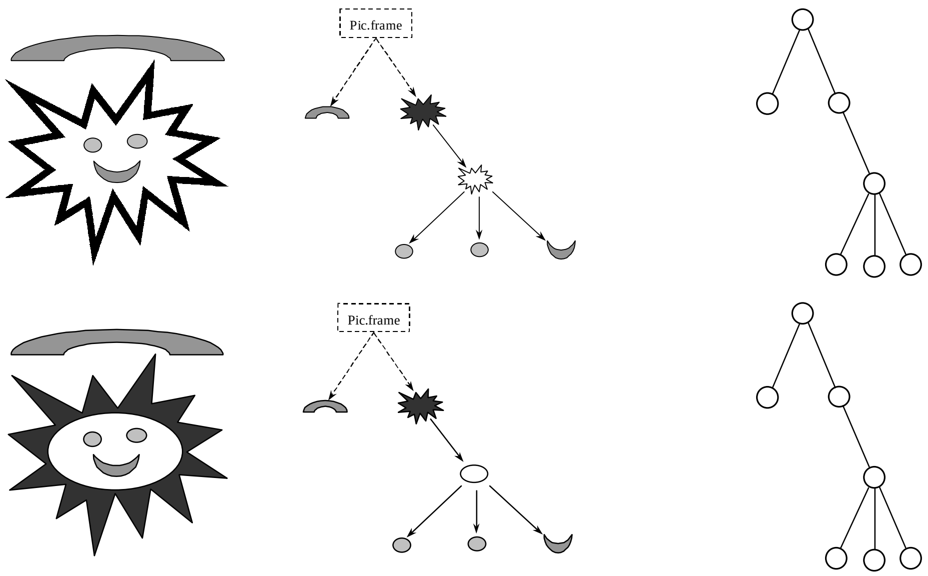

Fig.4. Two similar graphics images:
their trees of contours differ, but their homotopic trees are equal.

The Full Spectra Matrix and the Image Key Derivation

Without loss of generality we can consider the TIC of given image defined by the 1/D complex Fourier's spectra of respective contours [3, 7, 14]. In this way, the images of interest are represented invariantly, as follows:

- regarding intensity (because of a preliminary binarization, before the contouring),
- regarding translation (the image origin is moved into the image center of gravity),
- regarding rotation (by considering only the modules of the Fourier's spectra),
- and with respect to reflection (by the spectra folding towards the middle).

Besides, the invariance concerning scale can be obtained by all TIC spectra' normalization to a predefined value (in EIRS, it is the first frequency item of the greatest area contour).

In order to avoid sophisticated processing of graph matching [13, 14, 17] and to link the contradicting requirements for TIC presentation as a IDB key, we construct a frequency items' matrix called herein "*supreme spectra matrix"* (SSM), and containing the main frequency items of the contours ordered by their area, cf. Fig. 5.

The idea of the IDB key derivation for a given image is illustrated in Fig.5, by line *L* (sloped at α, α a given constant) that moves to infinity starting from the SSM origin. So, the IDB key items appear consequently with the movement of *L*. In other words, the SSM can be considered a multidimensional feature space currently reduced to its most informative subspace by a criterion of the image contour frequency items of importance.

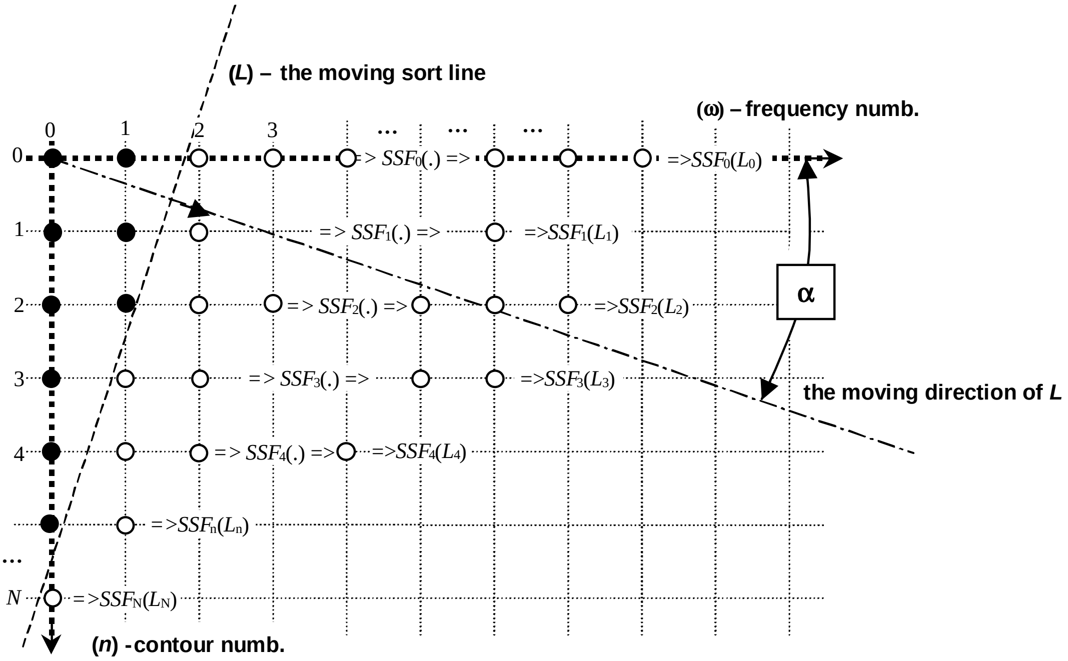

Fig.5. The Super Spectra Matrix and its sorting scheme producing the IDB key

The Algorithm for Key Derivation

The method proposed for fast image retrieval can be outlined algorithmically by the following sequence of steps (for a given image):

A1. Image binarization playing here as tone/color intensity normalization

A2. Creation of the respective TIC

A3. 1/D complex Fourier's transform for each image contour in the TIC

A4. Translation, scale and reflection normalization of the contours

A5. Composition of the SSM ordering the contours by importance

A6. Derivation of the image key items appropriately scanning the SSM (at given α).

A sequential illustration of these steps is given by Figs. 6÷15 on a real example. For more details about the method, see also [6].

EXPERIMENTS

An IBM compatible PC has been used for the EIRS development and test, namely CPU Intel Celleron 330MHz, MM 128MB, and HDD 15GB.

The testing IDB consists of 3937 images that consume 63MB (as JPGs) or 313MB (as BMPs) of HDD memory, while the Paradox part of IDB occupies only 13MB.

At present, the above-described technique (T3) is the most efficient among the three ones experimented by the EIRS. The T3 retrieval time, averaged on the whole IDB, is about 3÷5 sec., cf. also Fig.1. The "real" experiments, with a rough preliminary rotation of IDB images (at an angle of 7°), show that only 71 of all the 3937 images can not be retrieved exactly, i.e. the rate of T3 exact recognition is about 98,2%. The SSM scanning angle is chosen here $\alpha = 72°$. More details about possible improvement of this rate can be found in [6]. By the way, figures 6÷15 are produced also using the tools of EIRS.

Fig.6. An original mark image

Fig.7. The binarized image

Fig.8. All contours of the image

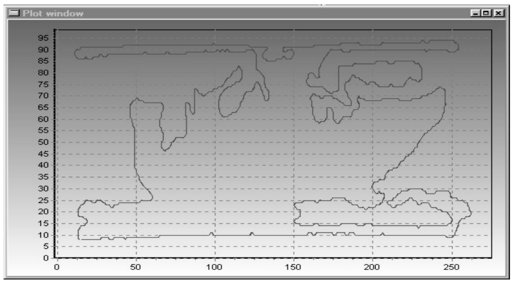

Fig.9. The greatest contour in this case (rotated at 180^o)

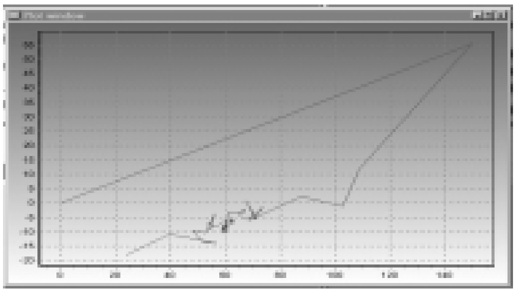

Fig.10. The full spectrum vector chain (the case of the greatest contour).

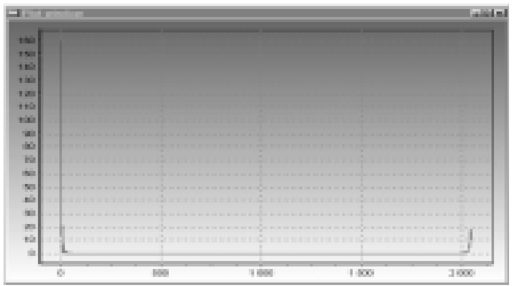

Fig.11. The modulus spectrum only (the case of the greatest contour).

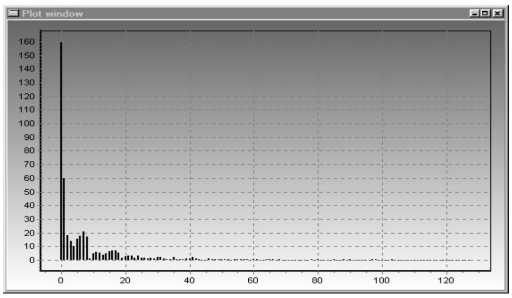

Fig.12. The essential part of the reflected modulus spectrum (see SSM row No.0)

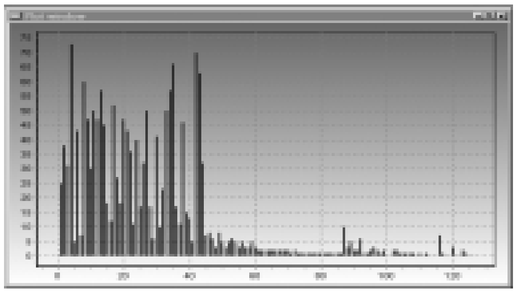

Fig.13. The IDB key obtained by α=0^o.

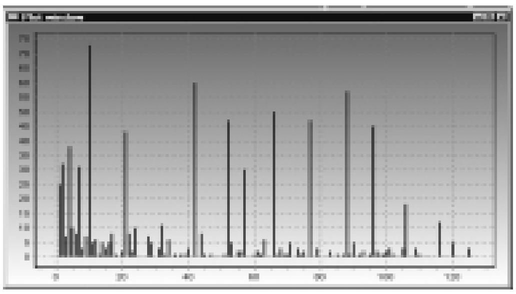

Fig.14. The IDB key obtained by α = 72^o.

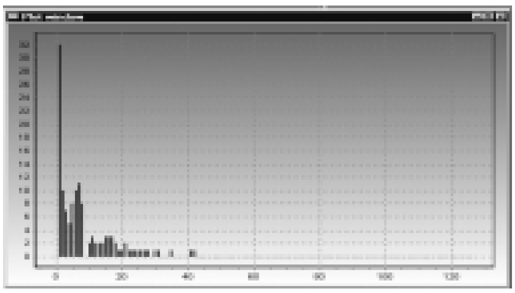

Fig.15. The IDB key obtained by α = 90^o.

CONCLUSIONS AND FUTURE WORK

A fast access method for image retrieval by graphic content has been proposed. The method is based on a TIC defined for graphic images, as well as on its restriction to a SSM obtained using an 1/D complex Fourier transform for each of the contours. A heuristic approach for derivation of image key from the respective SSM is proposed. The method is invariant with respect to linear transforms in image color and shape, and is well fitted with performance requirements of contemporary IDB-s. The method efficiency is experimentally proved, in the frames of an EIRS system, on an IDB of about 4000 images of hallmarks.

The method proposed can be applied as image search engine into conventional IDB-s, in information and image retrieval systems for marks, hallmarks, trademarks, postmarks, etc. At present, the method is restricted to pictures containing well-localizable essential graphics (i.e. gray scale or colored objects vs. a picture background). Besides, the picture should not contain a great level of noise, and especially an "artificial" noise (artifacts). The improvement of the method efficiency regarding such kind of noise will be the matter of near future work, due to possible application for the necessities of Bulgarian Patent Office.

REFERENCES

[1] Eakins, J.P.: Towards intelligent image retrieval, *Pattern Rec. J.*, 35 (2002), pp. 3-14.
[2] Eakins, J.P., K. Shields, and J. Boardman: ARTISAN – a shape retrieval system based on boundary family indexing, 12p. (via http://citeseer.nj.nec.com/cs, keyword 'CBIR')
[3] Dimauro, G.: Digital Transforms in Handwriting Recognition, In: Impedovo, S. (ed.) *Fundamentals in HWR*, NATO ASI Series "F", Vol. 124, Springer-Verlag, Berlin Heidelberg, 1994, pp.113-146.
[4] Dimov, D.: Experimental Image Retrieval System, *Cybernetics and Information Technologies, IIT-BAS*, Vol.2, 2, 2002, pp.120-122.
[5] Dimov, D.: Wavelet Transform Application to Fast Search by Content in Database of Images, in: *Proceedings of Intelligent Systems (IS'2002)*, Sept. 10-12, 2002, Varna, Vol.I, pp. 238-243.
[6] Dimov, D.: Fast Retrieval of Images by Graphics Content, in: *Working papers' series of IIT-BAS*, ISSN-1310-652X, WP/151, Dec. 2002, 13p.
[7] Duda, R. and P. Hart: *Pattern Classification and Scene Analysis*, A Wiley-Inter science Publication, 1973.
[8] International Hallmarks on Silver, collected by TARDY, Paris, France, 1981, 429-523.
[9] Khoshafian, S.: *Object Oriented Databases*, John Wiley & Sons, Inc., 1993.
[10] Martin, J.: *Computer Database Organization*s, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1975.
[11] Official bulletin of Bulgarian Patent Office. ISSN 1310-179X, 20311, Vol.10 (2001).
[12] Pavlidis, T.: *Algorithms of Computer Graphics and Image Processing*, Comp. Sci. Press, Inc., 1982.
[13] Smeulders, A.W.M., M. Worring, S. Santini, A. Gupta, and R. Jain: Content-Based Image Retrieval at the End of the Early Years, *IEEE Trans. on PAMI*, 22, 12, (2000), pp. 1349-1380.
[14] Sonka, M., V. Hlavac, and R. Boyle: *Image Processing, Analysis, and Machine Vision*, 2d edition, PWS Publ. at Brooks-Cole Publ. Co, ITP, Pacific Grove, CA, 1998.
[15] Stanchev, P.L.: Content Based Image Retrieval Systems, in: *Proceedings of CompSysTech'2001*, June 21-22, 2001, Sofia, P.1.1-6.
[16] Valova I., B. Rachev: Image organization, querying and retrieval by color distribution features, in: *Proceedings of CompSysTech'2002*, June 20-21, 2002, Sofia, II.3.1.
[17] Vasconselos N., and M. Kunt: Content-based retrieval from databases: current solutions and future directions, in *Proceedings of ICIP'2001*, Oct. 7-10, 2001, Thessaloniki, Greece: IEEE, Vol. 3, pp. 6-9.
[18] Zaniolo C., S. Ceri, C. Faloutsos, R.T. Snodgrass, V.S. Subrahmanian, and R. Zicari: *Advanced Database Systems*, Morgan Kaufmann, Inc., San Francisco, CA, 1997.

ABOUT THE AUTHOR

Assoc. Prof. Dimo Dimov, Ph.D., Institute of Information Technologies, Bulgarian Academy of Sciences, Tel: (+359 2) 8706493, E-mail: dtdim@iinf.bas.bg