Harmonic and Wavelet Transform Combination for Fast Content Based Image Retrieval

by Dimo T. Dimov

Institute of Information Technologies
Bulgarian Academy of Sciences
dtdim@iinf.bas.bg

a presentation at the Conference on
"PDE Methods in Applied Mathematics and Image Processing"
(Sep. 7-10, 2004, Sunny Beach, Bulgaria)

Abstract— Content Based Image Retrieval (CBIR) is a relatively new area of Informatics and covers methods, techniques, and approaches for automatic (or automated) retrieval of images by content description based on simple features as color, texture, shape, movement, etc., as well as on structures over them. Each CBIR system comprises a Database of Images (IDB), for which it should provide content-based image access methods being enough fast and reliable from a user viewpoint. From the other hand, the CBIR can be considered also a pattern recognition approach using a large dictionary of image examples, where the accents are put on images preprocessing to adapt the appropriate recognition technique.

A new heuristic approach to fast and reliable CBIR will be presented, namely arranged as a combination of Fourier transform and wavelet transform for images considered in a polar coordinate system. Experimental results will be also committed in comparison with other CBIR approaches experimented by the EIRS (Experimental Image Retrieval System of IIT) on a real IDB of hallmark images.

Obviously, the approach proposed can be considered based on a linear 2D transform for global processing of initial images. The approach combination with differential filtering techniques of first and second order for possible image contouring will be also brought to discussion due to the Conference topic.

Key words— Content Based Image Retrieval (CBIR), fast and reliable CBIR, Fourier's and wavelets' transform of images.

Contents

slide No./46

Introduction

- CBIR (Content Based Image Retrieval)
- Image Data Bases (IDB)
- Some examples (GEMINI, QBIC (IBM), CORE, ARTISAN) of the "early" CBIR
- How they work:
  - color, tone or texture histograms, directional histograms of sub- objects, elastic constellations of them, etc..
  - full shape representation (segmentation troubles, problems of noise)
- Search in IDBs (sequential or indexed access)
- EIRS (the Experimental Image Retrieval System of IIT)

The CBIR state of the art in brief

- CBIR – the retrieval of images on the bases of features automatically derived from the images themselves.
- CBIR draws many of its methods from the field of image analysis and processing, computer vision, statistics, and pattern recognition.
- An ordinal CBIR system includes a database (IDB) for keeping image data and for maintaining the typical user requests, a graphic user interface for request wording and result visualization, and suitable indexing techniques for the feature vectors treasuring.
- CBIR systems examples: TRADEMARK (1992), QBIC (1993), Photobook (1994), ARTISAN (1999), Picture-Finder (2002),…

  cf. [Eakins], [Stanchev], [Smeulders et al], [Vasconselos et all], [Zaniolo et all].

The CBIR state of the art in brief (continue 1)

• 3 levels of CBIR [Eakins]:

- by primitive features - color, texture, shape (representative points and their histograms), etc., e.g. “all pictures containing yellow stars in a ring <?>”

- by attributes or logical features, involving some degree of inference among objects in the picture: e.g. ”a train crossing a bridge <?>”

-by abstract attribute; e.g. ”all pictures illustrating pageantry”

• The future CBIR levels 2 & 3 (will) rely on AI areas of Informatics: knowledge representation, syntactic/semantic approaches, machine learning, etc..

• CBIR level 1 <=> “early” CBIR, and it seems everybody are still therein [Eakins], [Smeulders et al].

The CBIR state of the art in brief (continue 2)

- Possible applications of the “primitive” CBIR:

  -for large amount of given class images: trademarks, postmarks, stamps, fingerprints, industrial components, etc..

  -Accent on the CBIR system’s interaction with the user: search by association, by example, by category; Vienna convention in the practice of patent offices, assay offices, etc..

  -Image DB: extension over conventional DBs and/or DB Management Systems (DBMS).

  -The problem of effectiveness (access speed and noise tolerance).

The EIRS (Experimental Image Retrieval System)

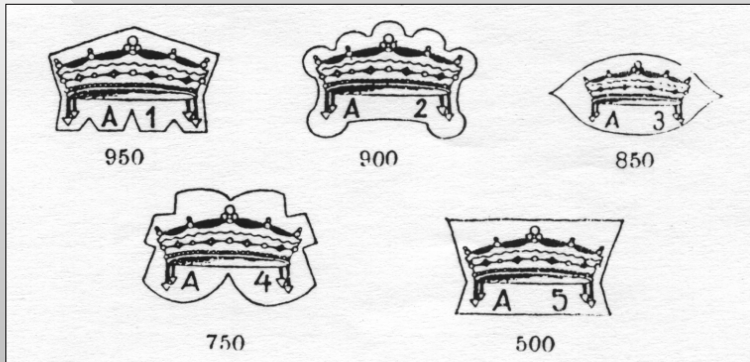
The images of interest

(Examples the EIRS methods are designed for)

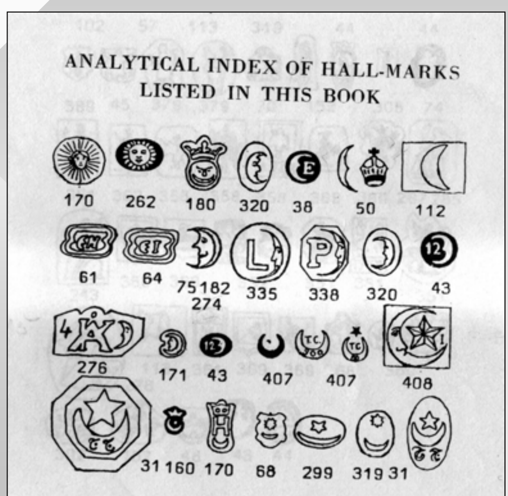
An example of a complete hallmark, parts (from left to right) represent the sponsor’s mark, the mark for sterling silver, the London Assay Office mark and the date mark for 1976, cf. [A booklet about Hallmarks on gold, silver & platinum, the Assay Office Publication Dept., London., 10p.]

The images of interest
(Other examples suitable)

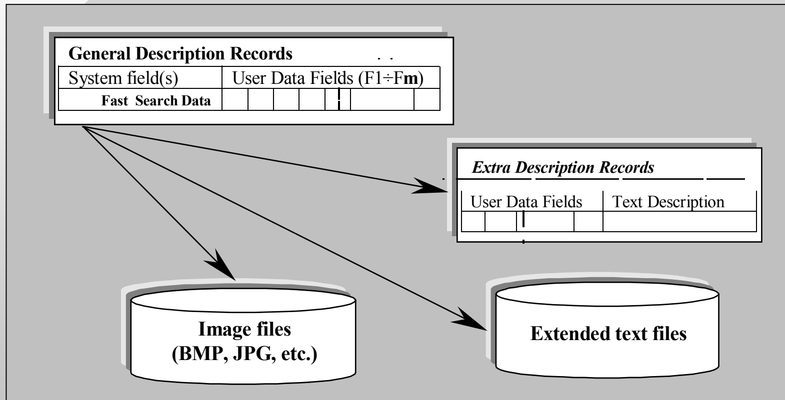
Examples of Bulgarian hallmarks on old silverware, used as maker's marks in 1907-1919. Numbers below the images of Tsarina's crown represent silver standards, cf. [Int. Hallmarks].



IDB of EIRS: the first 29 images of hallmarks – from the 4000 ones already loaded into the test IDB (as they are semantically ordered in [Int. Hallmarks]):



The IDB of EIRS simple logical scheme

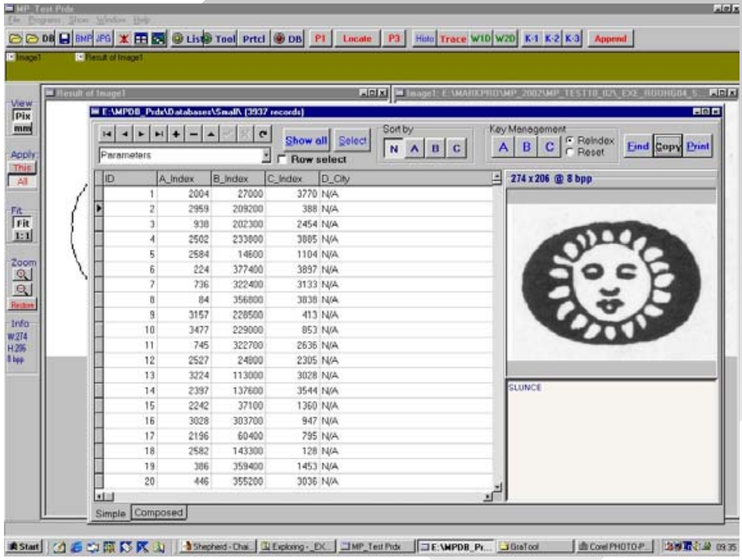


The EIRS system (a DEMO)

- If the Demo is not possible show next 4 slides:

- EIRS: the IDB browser
- EIRS: the Search-engine start menu
- EIRS: the Search engine (result table I)
- EIRS: the Search engine (result table II)

EIRS: the IDB browser



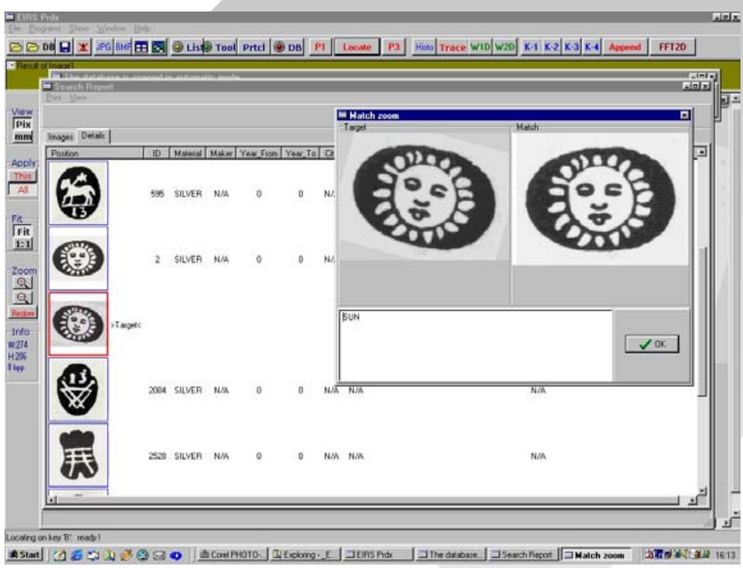
13

EIRS: the Search-engine start menu

14

EIRS: the Search engine (result table I)

EIRS: the Search engine (result table II)



The EIRS system in brief

- A view on a retrieval experiment
- 3 fast CBIR techniques proposed:
  - T1: ToCFT using a 1/D Fourier's transform to the image tree of contours.
  - T2: 2DWT using a 2/D wavelet transform.
  - T3: 2DFT using a 2/D Fourier's transform (to be extended hereinafter).
- Common principles for the 3 techniques:
  - 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
  - The fast access is performed on this key data using ***conventional index access methods*** of the DBMS of current use.

All fast access methods work using the convention of a "key" and give the search result as follows:

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

where "(i) $\prec$ (ii)" means "(i) is the greatest value less than or equal to (ii)".

Intuitive Requirements to the image key performance

- An intuitive *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.

Requirements to the image key definition:

- 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 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 length there can be more than one image to correspond.
- All DB images of one and the same key should be considered as equivalent images.
- The key length should be chosen in such a way that the image key to be enough informative one (from user viewpoint).

This talk aim:

- Brief comparison of the 3 basic techniques of EIRS for fast and reliable CBIR.
- Accent on the rotation invariance – against possible accidental rotation of input images to search.
- The other type of invariance against translation, scaling, mirroring and/or illumination irregularities can be considered resolved – by preliminary evaluation and/or compensation. This approach is considered bad practice to realize rotation invariance, cf. also [Reiss]. Internal integration of rotation invariance within EIRS key derivation techniques is necessary.
- Integration of rotation invariance within EIRS techniques:

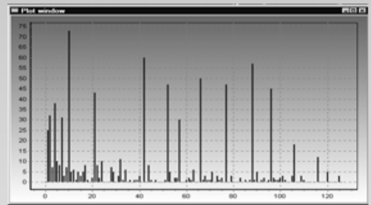
  - in T1 (ToCFT) – resolved by the definition of 1DFT applied to ToC (the tree of contours). But it has noisy troubles with very close couples of contours.

  - in T2 (2DWT) – principal difficulties for integration because of the 2DWT definition (nevertheless the good properties for noise tolerance organization here).

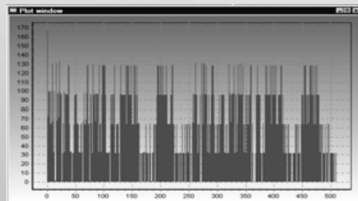
  - in T3 (2DFT) – the most promising for both the rotation invariance integration and the good properties for noise tolerance organization.

EIRS: Image key generation
(a visual comparison)

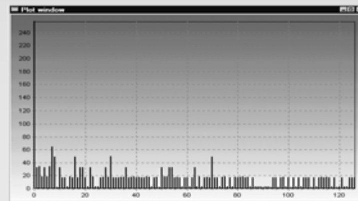
An original mark image.



Its ToCFT key
(i.e. by a 1D Fourier's transform on the image's Tree of Contours).



Its 2DWT key
(i.e. by a 2D Wavelets' transform on the whole image).

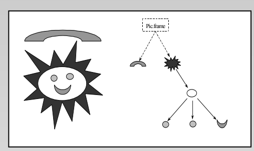


Its 2DFT key
(i.e. by a 2D Fourier's transform on whole image).

About the 3 techniques of EIRS:

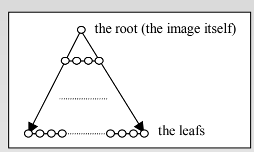
(T1): ToCFC:

Image :-> ToC (the Tree of image Contours) :->
:-> 1DFT to ToC :->
:-> a heuristic arrangement of the frequency modula :->
:-> IDB key



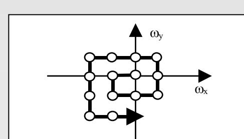
(T2): 2DWT:

Image :-> 2DWT of the image (WT package, a quad-tree of decomposition) :->
:-> consider only the leafs :->
:-> order them by "importance" :->
:-> IDB key



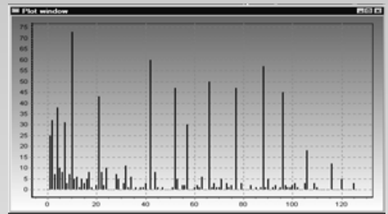
(T3): 2DFT:

Image :-> 2DFT of the image :->
:-> scan appropriately the 2d frequency space :->
:-> IDB key



EIRS/T1: the ToCFT method to image key generation:

(a) An original mark image.

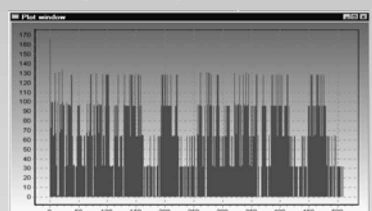


(c) The resulting ToCFT key

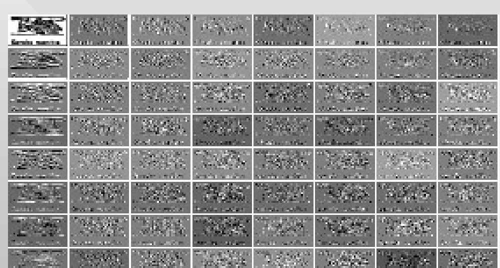
(b) All contours of the image -> the image Tree of Contours (ToC)

EIRS/T2: the 2DWT method to image key generation:

(a) An original mark image.



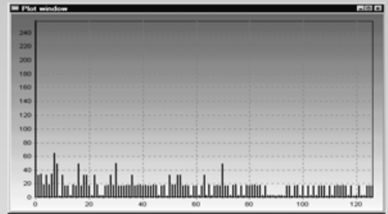
(c) The resulting 2DWT key



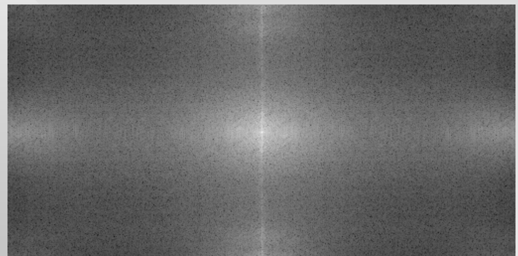
(b) 2D Wavelets’ transform on the binarized image.

EIRS/T3: the 2DFT method to image key generation:

(a) An original mark image.



(c) The resulting 2DFT key



(b) 2D Fourier's transform on the binarized image (a logarithmic scale for the modula-spectrum).

Image geometric transforms' influence on IERS/T3 approach (2DFT and its extensions)

2DFT: $\boxed{G(\boldsymbol{\omega}^T) = \boldsymbol{F}(g(\mathbf{x}^T)) = \int g(\mathbf{x}^T)\exp(-j\mathbf{x}^T\boldsymbol{\omega})d(\mathbf{x}^T)\,; \quad \mathbf{x},\boldsymbol{\omega} \in \mathbf{E}^n, n = 2}$

or $$G(\omega_x,\omega_y) = \iint g(x,y)\exp(-j(x\omega_x + y\omega_y))dxdy$$

or $$G(\omega_x,\omega_y) = \iint g(x,y)e^{-j(x\omega_x + y\omega_y)}dxdy$$

Translation:

Modula-spectrum is invariant.

$$\tilde{g}(\tilde{\mathbf{x}}^T) \Leftrightarrow g(\mathbf{x}^T) \quad , \quad \mathbf{x} = \tilde{\mathbf{x}} + \boldsymbol{\Delta}$$
$$\tilde{G}(\tilde{\boldsymbol{\omega}}^T) \Leftrightarrow G(\boldsymbol{\omega}^T) = \exp(-j\boldsymbol{\Delta}^T\tilde{\boldsymbol{\omega}})\tilde{G}(\tilde{\boldsymbol{\omega}}^T)\,, \quad \boldsymbol{\omega} = \tilde{\boldsymbol{\omega}}\,.$$
i.e. $\left|\mathbf{F}(\tilde{g}(\tilde{\mathbf{x}}^T))\right| = \left|\mathbf{F}(g(\mathbf{x}^T))\right|$

Scale:

Modula-spectrum vastly depends on scale matrix.

$$\tilde{g}(\tilde{\mathbf{x}}^T) \Leftrightarrow g(\mathbf{x}^T) \quad , \quad \mathbf{x}^T = \mathbf{S}\tilde{\mathbf{x}}\,, \quad \mathbf{S} = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix}$$
$$\tilde{G}(\tilde{\boldsymbol{\omega}}^T) \Leftrightarrow G(\boldsymbol{\omega}^T) = |\mathbf{S}|\tilde{G}(\tilde{\boldsymbol{\omega}}^T)\,, \quad \mathbf{S}\boldsymbol{\omega} = \tilde{\boldsymbol{\omega}}\,.$$
i.e. $\left|\mathbf{F}(\tilde{g}(\tilde{\mathbf{x}}^T))\right| = |\mathbf{S}|^{-1}\left|\mathbf{F}(g(\mathbf{x}^T))\right|$

Rotation:

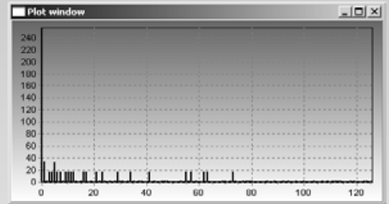
Modula-spectrum is rotated as the original image is.

$$\tilde{g}(\tilde{\mathbf{x}}^T) \Leftrightarrow g(\mathbf{x}^T) \quad , \quad \mathbf{x}^T = \mathbf{R}\tilde{\mathbf{x}}\,, \quad \mathbf{R} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$
$$\tilde{G}(\tilde{\boldsymbol{\omega}}^T) \Leftrightarrow G(\boldsymbol{\omega}^T) = \tilde{G}(\tilde{\boldsymbol{\omega}}^T)\,, \quad \boldsymbol{\omega}^T = \mathbf{R}\tilde{\boldsymbol{\omega}}$$
i.e. $\left|\mathbf{F}(\tilde{g}(\tilde{\mathbf{x}}^T)\right| = \left|\mathbf{F}(g(\mathbf{x}^T)\right|$

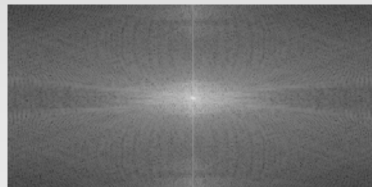
EIRS/T3: an extension:

a simple Polar Map (SPM) of the image before it's 2DFT:

(a) An original mark image.



(c) The resulting SPM-2DFT key



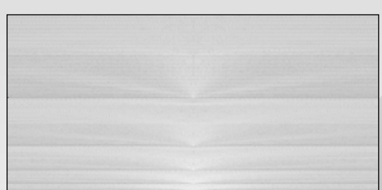
(b1) Preliminary SPM of the binarized image, and
(b2) it's 2D Fourier's transform (a logarithmic scale for modula-spectrum).

EIRS/T3: next extension (still under development):

= SPM + FTWT, FTWT = FT on angles (horizontally), get modula and then a WT on distances (vertically).

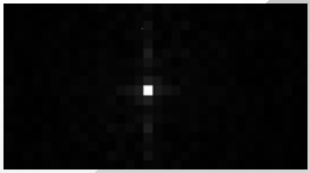
(a) An original mark image.

(!) Zoom of the Spectra essential part
(linear scale for visualization)

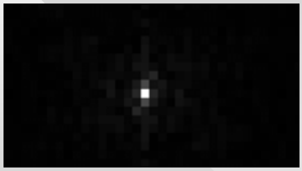


(b1) Simple Polar Map of the binarized image, and
(b2) it's FTWT (a logarithmic scale for visualization).

EIRS/T3: A comparison of the essential spectrum areas for the above 3 approaches:



(a) 2DFT of the original image.



(b) Simple Polar Map and 2DFT (!)
promising experiments,
cf. also [Reiss], and [Zhang & Lu].

(c) Simple Polar Map and FTWT (?)
(a heuristics to be experimented
and proved)

Experiments (notes to results'tables)

(1) *Pure access time*: it includes the time for image key generation as well as the time for fault-tolerant search of the generated key among all IDB-image keys preliminary loaded into the main memory. It can be theoretically evaluated to about ~ $\log_2|IDB|$.

The *full access time* (for CPU Intel Celeron 333MHz, MM SDRAM 192MB, HDD IDE-47 20GB) is about 3÷5s for the main test IDB (3937 images) and less that a 1s for the extra test IDB (146 images). Note that 2 extra seconds are necessary for the EIRS to visualize the results.

(2) *IDB load time*: it can represent the time for a *conventional sequential access* to IDB, by the averaged formulae ~ |IDB|.

(3) *Number of bad experiments*: For each image of the given IDB a retrieval experiment is conducted, i.e the number of all experiments is |IDB|. An experiment is denoted "bad" if no exact match occurs between the input and the most similar image retrieved.

(4) *Experimental recognition rate*: an "ambitious" evaluation by the bad experiments' percentage relatively to the all experiments' number (|IDB|).

Experiments (results for two test IDB-s)

EIRS method used	IDB access [(1)] time [s /img] (pure | full)	IDB load [(2)] time [s]	Number [(3)] of bad experiments	Experimental [(4)] recognition rate [%]
ToCFT	**0.44 | 3÷5**	**1076 ~ 18 min**	**71**	**1.8**
2D-FT	**0.95 | 3÷4**	**2113 ~ 35 min**	**127**	**3,2**
2D-WT	**2.79 | 4÷7**	**6365 ~ 106 min**	**975**	**24.7**

Table 1. Experimental results on the main test IDB,
|IDB| = 3937 half-tone images of hallmarks

EIRS method used	IDB access [(1)] time [s /img] (pure | full)	IDB load [(2)] time [s]	Number [(3)] of bad experiments	Experimental [(4)] recognition rate [%]
ToCFT	**0.19 | <1.0**	**14**	**57**	**38**
2D-FT	**0.65 | <1.0**	**74**	**22**	**15**
2D-WT	**0.68 | <1.0**	**78**	**64**	**44**

Table 2. Experimental results on the extra test IDB,
|IDB| = 146 multicolor/tone images of POBG trademarks

Near future work

- Possible application in the current practice of Bulgarian Patent Office (PORB).

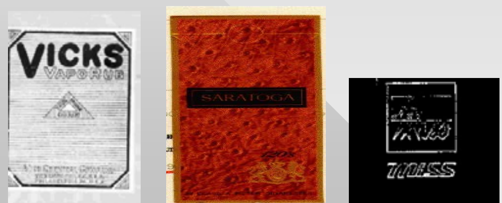
Trademark examples from the current practice of POBG:
15 trademark images (from the extra test IDB) are shown, derived in 5 typical categories:


(1) Enough clean images


(2) Relatively clean images

(3) Multicolor and/or halftone images

(4) Images with a great level of artifacts



(5) Very dirty images

Other possible applications:

Extensions/Modifications (of EIRS & IDB) for:

- numismatics
- banking, bond hobbies
- philately.
- forensic expertise
- machinery design

In all mentioned cases (excepting the forensic expertise) the images can be considered "static" ones, because of the original are usually produced by precise machines (of printing and publishing industry)

Additionally "dynamic" image recognition by examples can be performed for short video clips: e.g. face speaking in a frame, palm gesticulating a sign language, object flying in the sky, etc..

2 general versions of IDB performance:

- a local IDB: cf. test IDB of EIRS.
- a distributed IDB, Internet applications, etc..

The EIRS techniques' restrictions:

- The techniques are restricted to pictures containing of well- localizable essential graphics (i.e. gray scale or colored objects vs. a picture background).

- The pictures should not contain a great level of noise, and especially an "artificial" noise (artifacts).

An idea to avoid the restrictions: by a level of manual preprocessing of images: the user important fragment(s) could be preliminary and interactively (!) extracted, and that means a loss of automation (what is unacceptable).

Current Problems to Solve:

- The 3 basic techniques of the EIRS has been developed in a concurrence to met different image types.

- EIRS/T1 is the best one for now. It is invariant to image accidental rotation, but is sensible to a specific noise – small and/or regular noise may cause tremendous changes in the ToC representation.

- EIRS/T2 & T3, and more precisely their combination described here as SPM-FTWT, is very promising to overcome EIRS/T1 in CBIR effectiveness (i.e. processing speed, noise tolerance and invariance to accidental images transform outside).

- A common problem to solve (concerning all the EIRS access techniques) – to isolate or to suppress the relatively great level of "artificial" noise in images of PORB practice.

Conclusions

- A fast access method for image retrieval by graphic content has been proposed. The method can be associated to the new informatics area of CBIR.
- The method basic idea is to represent the essential image content as a well-structured DB key of fixed length, namely - the more important image data to take more front positions in the key.
- Three (3) basic techniques has been developed by the method - ToCFT, 2DWT, and 2DFT. The techniques ascend on 3 different types of image shape features.
- An new 2D transform (SPM-FT-WT) is proposed based on a preliminary polar mapping of the image, Fourier transform modula on the angle parameter, and wavelet transform on the distance parameter of polar mapping. The SPM-FT-WT is rotation invariant by definition and preserves the noise tolerance properties characteristic of the FT and WT.
- The experimental image retrieval system (EIRS) has been developed for proving the proposed methods on an IDB of about 4000 images of hallmarks.

Conclusions (continuation)

- The EIRS method can be also considered a recognition method by closeness to examples (centers of classes for recognition).
- The method accents on the most “important” image information based on shape instead of particular features as intensity, color, texture, etc..
- The method is well adapted to the conventional index access methods of conventional DB management systems.
- The method can be applied as image search engine in a conventional DB of images, for instance in information and image retrieval systems for marks, hallmarks, trademarks, postmarks, cultural heritage archiving, etc..
- Besides of “static” images considered herein, also “dynamic” recognition of image sequences, e.g. short video clips, can be performed by similar approaches. Besides 2D-signals (as the images herein), also, 1D- and/or 3D-signal recognition can be developed by analogy to the techniques proposed.

References:

Eakins, J.P.: Towards intelligent image retrieval, *Pattern Rec. J.*, 35 (2002), pp. 3-14.
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')
Duda, R. and P. Hart: *Pattern Classification and Scene Analysis*, A Wiley-Inter science Publication, 1973.
International Hallmarks on Silver, collected by TARDY, Paris, France, 1981, 429-523.
Pavlidis, T.: *Algorithms of Computer Graphics and Image Processing*, Comp. Sci. Press, Inc., 1982.
Reiss, T. H.: Recognizing Planar Objects Using Invariant Image Features. Lecture Notes in Computer Science, Vol. 676. Springer-Verlag, Berlin Heidelberg New York (1993).
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.
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.
Stanchev, P.L.: Content Based Image Retrieval Systems, in: *Proceedings of CompSysTech'2001*, June 21-22, 2001, Sofia, P.1.1-6.
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.
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.
Zhang D., and G. Lu: Enhanced Generic Fourier Descriptors for Object-Based Image Retrieval, Int. Conf. On Acoustic, Speech, and Signal Processing (ICASSP'2002), May 13-17, 2002 Orlando Florida, USA.

Author publications in the topic:

Dimov, D.: Fast Retrieval of Images by Graphics Content, *Cybernetics and Information Technologies, IIT-BAS*, Vol.4, 2, 2004, (to appear).
Dimov, D.: Fast access by content to multimedia databases, in Abstracts of Int. Congress MASSEE'2003, Sept. 15-21, Borovets, Bulgaria, 2003, pp.94-94.
Dimov, D.: Fast, Shape Based Image Retrieval, Proc. of CompSysTech'2003, June 19-20, Sofia, 2003, pp.3.8.1-7, (http://ecet.ecs.ru.acad.bg/cst/Docs/proceedings/S3/III-8.pdf)
Dimov, D.: Experimental Image Retrieval System, *Cybernetics and Information Technologies, IIT-BAS*, Vol.2, 2, 2002, pp.120-122.
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.

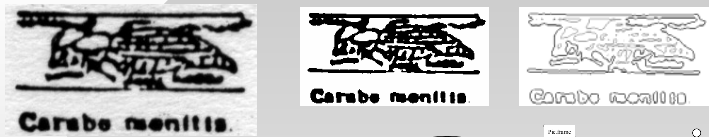
Acknowledgements:

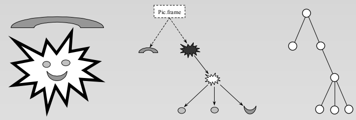
This research is jointly supported by the following grants/funds:
(i) the Grant #RC6/2004 (IIT #0210143) of the ICT Development Agency at Bulgarian Ministry of Transport & Communication,
(ii) the Grant #I-1306/2003 (IIT #020057) of the National Science Fund at Bulgarian Ministry of Education & Science, and
(iii) the BAS research funds of IIT disposal (IIT #010056).

Special thanks to Ognyan Kounchev and Svetozar Margenov (from BAS) for the perfect organization of the Conference as well as to all participants for the fruitful discussion of this work.

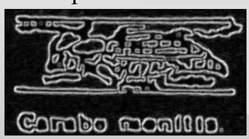
Extra page 1:

The way the EIRS (T1) obtains the image contours:
(a way for possible syntactic/semantic extensions)





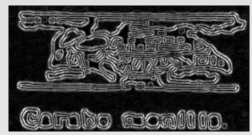
The following approaches are also possible:



$$|\text{grad } g(x,y)| = \sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2}$$

$$\psi = \arg\left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right)$$

via 1st gradient filter (Sobel vers. here)

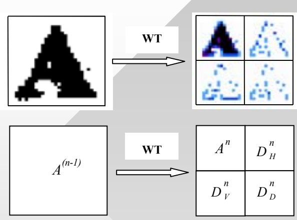


$$\nabla^2 g(x,y) = \frac{\partial^2 g(x,y)}{\partial x^2} + \frac{\partial^2 g(x,y)}{\partial y^2}$$

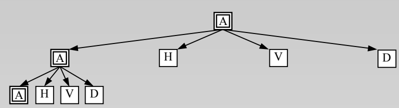
via 2d gradient filter (Laplace oper. here)

Extra page 2:

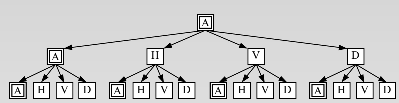
The way the EIRS (T2) processes via 2DWT:



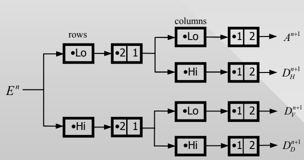
2/D wavelet transform
a visual interpretation



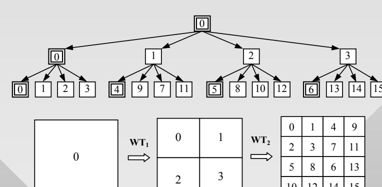
Wavelet chain-tree of 3 levels



Wavelet image package of 3 levels



2/D WT decomposition scheme
at given level *n*

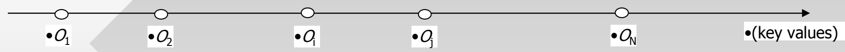


Possible sorting scheme of the leafs

Extra page 3.1:

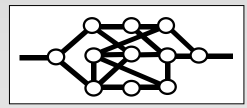
The convention for the DB object's keys:

K(*O*) - a value associated to *O* (an object)
A simple key example: a feature value
DB index <-> ordering the objects by their keys, i.e. a full ordering:



The usual CBIR practice:

A specific similarity function (*S*)
or distance function (*D*) (usually not a metrics)
to compare given input image (input object)
with all the DB images (objects).

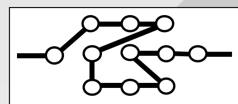


The above techniques usually lead to

(1) a partial ordering, i.e. to a sequential access into DB,
of an access time ~ N = |IDB|

while **the EIRS approach leads to**

(2) a full ordering, i.e. to a DB primary key index,
of an access time ~ $\log_2 N$, N =|IDB|,



what is the main advantage of the EIRS approach (!)

Extra page 3.2:

How the EIRS approach defines full order (the IDB keys):

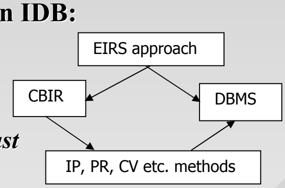
$$K(O) = \sum_{i=1}^{n} x_i M^{n-i} \quad , O \equiv (x_1, x_2, ..., x_n)$$

where *xi* , $i = 1 \div n$ are the *O*-object features ordered by "importance"
(for the current application of EIRS)

Conventional CBIR approaches for accessing an IDB:

- z-filling curves (a way of scanning the feature space)
- k-trees, R-trees, etc. heritage from GIS (Geographic Information Systems)

All they do not match the necessities of an effective (fast and noise tolerant) CBIR (!)



Similarities among objects:

Similarity function *S*. Distance *D*.

S(*A*,*B*) ~ (*D*(*A*,*B*))-1, where *A* and *B* are given objects.

Mehalanobis distance (Gaussian model of errors over given object), it's not a metrics.
Euclidean distance (also often used, because it defines a metrics)

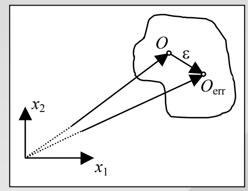
Extra page 3.3:

Squeezing the feature space dimension:

Allows an easier performance (of a PR system)
Simple approaches:
Reordering of the feature space coordinates (a permutation from $n!$ possible ones): Put the most important coord's at front position and cut the rest of "not so important" coord's.
More sophisticated approaches: e.g. Karunen-Loeve (KL principal components analysis (?)

The place of EIRS approach in the CBIR area :

- Searching an object into a DB => Method (or system) for PR (Pattern Recognition), or a DB access method, i.e.
- An object of given DB => A standard object (example) or a center of a class (of given type of objects)



$$O_i \in C_i \subseteq \bigcup_{j=1}^{n} C_j \quad , i = 1 \div n$$

$$C_i \cap Cj \quad , i \neq j$$

Extra page 3.4:

- (Approximated) representation of objects by their features.

- Vector representation (extracted features, e.g. colors, geometric measures, histograms, frequencies, etc.),

$$O \equiv (x_1, x_2, ..., x_n) \in \mathbf{E}^n \quad , \quad n \text{ – number of coordinates of } \mathbf{E}^n$$

- Noised objects:

$$O_{err} = O + \varepsilon \quad , \varepsilon \equiv (e_1, e_2, ..., e_n) \in \mathbf{E}^n$$

- A sufficient condition for a recoverable noise

$$D(O,\ O+\varepsilon) \ \leq \ d_{\min} = \frac{\min\{D(A,B) \mid A,B \in \mathbf{E}^n\}}{2}$$
