similarity measure, *medical image*,
the Hough transformation,
the Fourier transformation

Krzysztof WRÓBEL*, Piotr PORWIK*

COMPARATIVE INVESTIGATIONS OF SIMILARITY MEASURE EXPLOITED IN MEDICAL IMAGES PRESELECTION

Nowadays, the most significant impact of digital image processing in the area of applications are real–world problems. Many new technological trends in medicine and digital processing have been implemented. Several factors indicate such development. A major one is the perpetually declining cost of the computer equipment required. Both processing unit and capacity of storage devices continue to become less expensive year by year. Another factor is the increasing availability of equipment for digitising and displaying images. In modern image processing, images have to be compared each other because such approach allows us to automate of retrieval process. Computer image retrieving is today especially important in medical diagnostics [1,7] or in preliminary images selection [8,9]. Today, in the digital image processing are used techniques and methods which have well known mathematical backgrounds. It can be observed, that in the area of digital signal processing, the Hough and well known the Fourier transform are exploited very often. These transforms are frequently use in image retrieving and can be implemented as computer applications. In many cases the mentioned methods give promising results in images classification or preselection [1,2,4,5,11]. Special properties of such transforms can be used in statistical or comparative goals, especially when searched information has graphic form. Taking into account the mentioned applications, transforms as methods of preliminary medical images selection have been investigated. From this reason pictures, analysing in the paper, to medical images have been limited.

1. INTRODUCTION

Transmission and storage information is mostly done by computers. Hence, there is a need to perform document image analysis by identification of graphics blocks in images and extraction information. It is obvious, that the health service has to store medical information in any case for any patient. Hence, such information should be organised in the form of database. Because stored data are mostly very large, it is required fast access to stored information. If information has a text form, this one can be searched by means of descriptors use. Unfortunately, frequently it is necessary to find similar objects from thematic data collection in order to statistical or medical investigations. Such data are medical images for example.

It can be observed [1,2,6,7,9] that in many practical cases, image similarity is build on the basis of Euclidean distance: the smaller the distance is, the more similar the image is.

* University of Silesia, Institute of Informatics, 41-200 Sosnowiec, ul.Będzińska 39, Poland

In measuring similarity, there are few mathematical formulas such as Dice similarity coefficients, Jackard, Otsuka, Simpson, Manhattan, Robinson, and the like. In this paper the Hough transformation approach has been used because in this case Euclidean distances are also calculated. The different method based on spectral coefficients analysis is Fourier transform, where similarity between images is calculated immediately from spectrum of images.

2. THE HOUGH TRANSFORMATION

In this paper the Hough transformation [1,2,4,5] which is one of the popular techniques for images analysis is used. The Hough transformation maps each point in the original (x, y) plane to all points in the (θ, ρ) parameter plane that is Hough space of lines through (x, y) with slope θ and the distance ρ from origin. A line in the original image forms a cluster in the parameter plane. Once the locations of the clusters are determined, the skew of each line and the average skew are easily estimated. The Hough transformation is very useful because it detects not only solid lines but also broken lines. To reduce processing time it is enough to map edge points to the parameter plane instead of entire foreground pixels. In practice, to accelerate calculations only simple forms are isolated from images, in our case – straight lines. Such transformation is described more precisely in [1,6].

In this paper image elements, isolated with the aid of the Hough transformation, have been analysed and stored in appropriate set of features. For any set of features, Czekanowski's coefficient has been calculated. In the Czekanowski's method, average differences between elements of set are calculated [3]. Each element of the set has characteristic feature and each feature has a number description. For the first time, mentioned method in anthropological investigations has been used. As it has been shown in this paper, such approach also can be successfully adapted as image similarity measure. From analysed image, with the aid of the Hough transformation, line sections have been analysed. Each section can be described by means of some features: length, slope to X–axis, and values of section centre co–ordinates. Each type of feature constitutes the adequate set of features. Hence, average difference between sets of a given type can be calculated as follows:

$$Sim_0(M_1, M_2) = 1 - \frac{\sum_{i=1}^{n} \left[|M_{1i} - M_{2i}| \cdot S_i \right]}{n} \tag{1}$$

where:

Sim_0 – average difference isolated features between images M_1 and M_2,
M_{1i} – the value of *i-th* feature for the first compared element,
M_{2i} – the value of *i-th* feature for the second compared element,
S_i – the (attribute) weight of *i–th* feature,

n – the global number of features of all types.

The difference (1) is located in the range [0,1], where „1” stands for identical features of images M_1, M_2 and „0” indicates that images M_1 and M_2 do not have common features.

Because differences between sets of features can be significant, all features have appropriate attribute (weights). Influence of used features has been investigated for various images. For such assumption, for any feature its influence on quality of image similarity can be checked.

Before investigation, images especially with medical content, have been prepared. The three–stage process of image preparation has been shown in Fig.1.

a) b) c)

Fig. 1 Example of the medical image preparation.
a) X–ray image b) binarization and contouring c) contour after Hough transformation.

3. THE FOURIER TRANSFORMATION

The Fourier transformation is a powerful tool in linear system analysis. When it is necessary to compute the Fourier transform of a sampled signal or image, we use the Discrete Fourier Transform (DFT). If $\{f_i\}$ is a sequence of length L, which is obtained by taking samples of a continuous function at equal intervals, then its the 1D discrete Fourier transform (DFT) can be performed. In 2D dimensions case, if $g(i,k)$ is an $N \times N$ array (such array can be treated as substitution of real image), such as obtained by sampling a continuous function of two dimensions, then its two–dimensional DFT is the array given by [1,6]:

$$G(m,n) = \frac{1}{N} \sum_{i=0}^{N-1} \sum_{k=0}^{N-1} g(i,k) \cdot e^{-j2\pi\left(m\frac{i}{N} + n\frac{k}{N}\right)} \tag{2}$$

and the inverse DFT is:

$$g(i,k)=\frac{1}{N}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1}G(m,n)\cdot e^{j2\pi(i\frac{m}{N}+k\frac{n}{N})} \tag{3}$$

where:
$N\times N$ – vertical and horizontal dimension of image,
m,n – current coordinates of spectral coefficient,
i,k – current coordinates of the table g.

In spectral method, two cases of coefficients analysis have been performed.

In the first method, for pattern and given image, Fourier coefficients have been calculated. Hence, for images M_1 and M_2 two tables of coefficients F^{M_1} and F^{M_2} has been established. Additionally, values of spectra were normalised in this way that each coefficient value was divided by value of the first coefficient. Normalisation was carried out separately for each table. In next stage similarity coefficients between M_1 and M_2 was calculated:

$$Sim_1(M_1,M_2)=1-\sqrt{\sum_{y=0}^{N-1}\sum_{x=0}^{N-1}(F_{x,y}^{M_1}-F_{x,y}^{M_2})^2} \tag{4}$$

where:
x,y – the current column and row of the appropriate table F^{M_1} or F^{M_2}, where appropriate spectral coefficients are stored,
$F_{x,y}^{M_1},F_{x,y}^{M_2}$ – value of spectral coefficient located in position x,y in the tables F^{M_1} and F^{M_2}, for pattern and given image, respectively.

The values obtained by (4) can be negative, therefore values less than zero are exchanged for zero values:

$$Sim_1(M_1,M_2)=\begin{cases}Sim_1(M_1,M_2) & for\quad Sim_1(M_1,M_2)\ge 0\\ 0 & for\quad Sim_1(M_1,M_2)<0\end{cases} \tag{5}$$

Such a procedure allows standardizing obtained results, with described above the Hough transformation. Hence, the range of Sim_1 values is [0,1], and similar like previously, „1" stands for identical images M_1, M_2, and „0" indicates that M_1 and M_2 do not have common features.

In the second method an amplitude spectra A of images have been processed. An amplitude spectra, like before, was stored in the tables F^{M_1} and F^{M_2}, respectively:

$$A^i_{x,y}=\sqrt{\mathrm{Re}\left(F_{x,y}^{M_i}\right)^2+\mathrm{Im}\left(F_{x,y}^{M_i}\right)^2},\quad x,y=0,1,...,N-1,\ i=1,2 \tag{6}$$

The values of spectra were normalised like before. Then, similarity coefficient has been computed:

$$Sim_2\left(M_1, M_2\right) = 1 - \sqrt{\frac{\sum_{y=0}^{N-1}\sum_{x=0}^{N-1}(F_{x,y}^{M_1} - F_{x,y}^{M_2})^2}{K}} * 1000 \qquad (7)$$

where:

$F_{x,y}^{M_1}, F_{x,y}^{M_2}$ – the value of an amplitude spectra, located in position x, y in the tables F^{M_1} and F^{M_2}, for pattern and given image, respectively,

K – number of elements in the tables F^{M_1} and F^{M_2}. Obviously $K = N^2$.

Because the values obtained on the basis of equation (7) can be relative small, each result is multiplied by experimental matched constant 1000. Like previously, the range of Sim_2 values is [0,1].

4. RESULTS OF INVESTIGATIONS

Researches into similarity images have been carried out on the basis of the test images (benchmarks). Set of benchmarks presents Fig. 2. Appropriate images present: skulls (C1-C4), pelvises (M1-M4) and chests (K1-K4). Complete results could not be included as over 100 benchmarks were tested. Taking into account such restrictions, only Table 1 with the best results of similarity measure has been stated.

Table 2 includes some results of investigations. For images, analysed by means of the Hough transformation and the Czekanowski's coefficient, the best classification results have been shown.

Quality of similarity measure for described methods has been performed. For any mentioned above method of similarity images investigation, effectiveness of classification on the base of equation (8) has been established:

$$effectiveness = \frac{card(K)}{card(S) + card(K)} \qquad (8)$$

where:

K – number of images classified to correct group,

S – number of images classified to incorrect group.

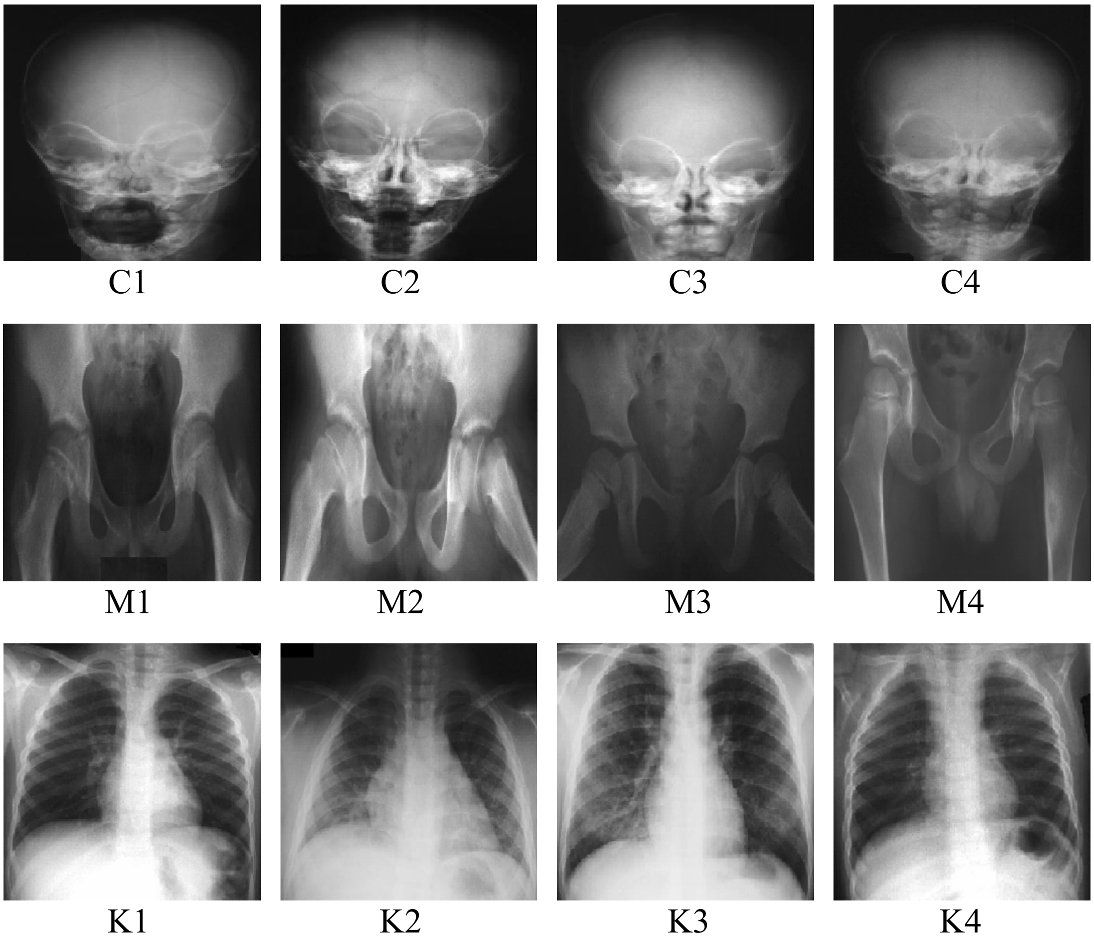

Fig. 2. The set of benchmarks – the selected medical images.

Table 1. The similarity Czekanowski's coeff., $S_i = 0.25$ for any *i*.

pattern	C1	C2	C3	C4	M1	M2	M3	M4	K1	K2	K3	K4
C1	1	0.887	0.903	0.926	0.820	0.806	0.824	0.813	0.820	0.825	0.815	0.822
C2	0.915	1	0.923	0.915	0.884	0.886	0.886	0.873	0.882	0.886	0.886	0.877
C3	0.934	0.934	1	0.935	0.887	0.887	0.885	0.868	0.894	0.894	0.885	0.881
C4	0.891	0.886	0.877	1	0.842	0.822	0.827	0.827	0.846	0.828	0.843	0.853
M1	0.822	0.855	0.842	0.866	1	0.957	0.948	0.945	0.907	0.907	0.922	0.909
M2	0.792	0.832	0.824	0.811	0.944	1	0.944	0.937	0.866	0.899	0.917	0.903
M3	0.792	0.821	0.805	0.821	0.931	0.943	1	0.934	0.862	0.891	0.901	0.878
M4	0.741	0.803	0.769	0.793	0.928	0.921	0.931	1	0.865	0.876	0.891	0.877
K1	0.843	0.875	0.867	0.861	0.907	0.902	0.901	0.916	1	0.914	0.933	0.940
K2	0.895	0.932	0.918	0.881	0.938	0.938	0.935	0.932	0.944	1	0.951	0.953
K3	0.855	0.909	0.901	0.870	0.945	0.949	0.938	0.947	0.965	0.946	1	0.954
K4	0.867	0.891	0.886	0.892	0.909	0.914	0.910	0.918	0.929	0.930	0.932	1

Table 2. Comparison of images effectiveness classification for the Hough and Fourier transformations.

	Weight of the angle of slope	Weight of the horizontal location	Weight of the vertical location	Weight of the section length	Effectiveness of classification
The Hough transformation and the Czekanowski's coeff.	0,25	0,25	0,25	0,25	94%
	0,1	0,4	0,4	0,1	92%
	0,4	0,1	0,1	0,4	92%
The Fourier transformation	The method 1				71%
	The method 2				78%

5. CONCLUSION

In this paper an important problem of finding of the similarity measure between images has been presented. Proposed solutions can be used as method of medical images ordering or sorting. Our approach allows to search databases where graphic objects are included. Graphic query allows to find all images enclosed in thematic database which are similar to a given query–image. From reported investigations follow, that method based on the Hough transformation gives significantly better results in comparison with spectral method.

BIBLIOGRAPHY

[1] CASTLEMAN K., Digital Image Processing. Prentice Hall. New Jersey, 1996.

[2] CHMIELEWSKI L., Image registration with the Hough transform. In Proc. 13th Polish Conference on Biocybernetics and Biomedical Engineering KBIB 2003, pp. 830-835, Gdańsk, Poland, Sep 10-13, 2003.

[3] CZEKANOWSKI J., Zarys metod statystycznych w zastosowaniu do antropologii, Prace Towarzystwa Naukowego Warszawskiego, 1913.

[4] DUDA R.O., HART P.E., Use of Hough transformation to detect lines and curve in pictures. Comm. ACM 15, 11-15, 1972.

[5] ILLINGWORTH J., KITTLER J., The adaptive Hough transform. IEEE Trans. Pattern Anal. Mach. Intell. 10, 690-698, 1987.

[6] JAIN A. K., Fundamentals of digital image processing, Prentice Hall. International Editions, 1989.

[7] MALINA W., Podstawy automatycznej klasyfikacji obrazów. Wydawnictwo Politechniki Gdańskiej, Gdańsk, 2002.

[8] PORWIK P., WRÓBEL K., WIDUCH S., DOROZ R., The Method of Preliminary Medical Images Selection. Journal of Medical Informatics and Technologies vol.2/2001, ISSN 1642-6037, 2001.

[9] L. G. SHAPIRO, G. C. STOCKMAN, Computer Vision, Prentice-Hall Inc., 2001.

[10] TADEUSIEWICZ R., FLASIŃSKI M., Rozpoznawanie obrazów. Państwowe Wydawnictwo Naukowe, Warszawa, 1991.

[11] WRÓBEL K., The Method of Parameterisation of Binary Image's Contour with Usage of the Hough Transformation. The 6th International Scientific - Technical Conference: Process Control 2004. Kouty nad Desnou, Czech Republic, 2004.