IMAGE RETRIEVAL EXPERIMENTS BASED ON SHAPE

Andreea Soimu*
Anca Loredana Ion**
Dorian Cojocaru*

University of Craiova,
Faculty of Automation, Computers and Electronics
**Mechatronics Department, **Software Department*
andr33a_soimu@yahoo.com, anca.ion@software.ucv.ro, cojocaru @robotics.ucv.ro

Abstract: The effective administration of huge multimedia document collections represents a challenge for a multitude of disciplines from signal processing to databases and artificial intelligence. This paper presents a content-based visual retreival application on the shape characteristic inside an image database. More shape signatures have been used to derive Fourier descriptors, all these being described in the paper. A new function was also suggested in order to improve the existent signature methods, this being called the area signature.

Keywords: image processing, shape, Fourier transform.

1. INTRODUCTION

The shape is one of the most important visual characteristics and one of the primitive characteristics used to describe the contents of multimedia data: images and videos. Describing the shape contents is a difficult task, because it is really difficult to define the perception characteristics and to measure the shapes similarity. The difficulty of this process is also increased by the fact that shapes are often influenced by noise, defects, distortions and occlusions.

Several different methods of classifying the shapes representation techniques are available [Aslandogan, et. al., 1999]. The most common and general classification is based on using the points located on the outline of the object (shape), instead of using the points located inside it. The resulted classes are known as: boundary and global [Pavlidis, 1980]. The shapes representation techniques can be distinguished between the spatial domain and the characteristic domain. The spatial domain methods compare the shapes using the points or the points' characteristics, while the characteristic domain techniques compare the shapes using the characteristics vectors. The variety of shape representation techniques are first classified using methods that are based only on contours, for which information is extracted only from the contour and region or surface based methods, for which the information is extracted from the entire shape region.

The contour based shape descriptors based on contours usually imply little calculation and small dimension in comparison with the region - based descriptors [Pass et. al, 1996; Berretti et. al., 2000]. These descriptors can be applied in applications in which the shape contours can be distinguished. In literature, contour descriptors are more popular than the region based ones, due to their clear properties and simple implementation.

There are 3 types of Fourier descriptors (FD): conventional, related and short time. In this paper the conventional Fourier descriptors will be treated. The conventional FD descriptors derive by applying the Fourier transforming on a single dimension signature function. More shape signatures have been used in order to derive the Fourier descriptors, all these being described bellow. A new signature function has been also suggested in order to improve the existent signature methods, this being called the area signature.

2. IMAGE PRE-PROCESSING

The shapes in our database are acquired from the shapes of the real-life objects. The shapes are images in grey levels. In the pre-processing phase information is extracted from the contours, or the points co-ordinates on the shape contour.

2.1 The signatures of shapes

Usually, a shape's signature represents any single dimension function that describes a bi-dimensional surface. In this study 3 shape signatures have been treated: central distance, cumulative angular function and a new suggested method, the maximum area function.

The reason for choosing the first 2 signatures is due to their frequent usage in the newest Fourier descriptors implementation and to the fact that they have proved to be the most practical methods of the general representation of the shape. Next, we are going to assume that the contour points co-ordinates (x (f),y(f), f= 0, 1, ..., L-1, have been extracted in the pre-processing phase.

2.1.1 The central distance function

The central distance function is represented in the distance between the points on the contour and the weight centre (xc,yc) of the shape:

$$r(t) = [(x(t) - x_c)^2 + (y(t) - y_c)^2]^{1/2} \qquad (1)$$

Due to the diminution of the weight centre, that represents the shape's position, from the contour points co-ordinates, this distance does not variate at translation.

2.1.2 The cumulative angular function

The shape can be represented by the angle between the countour points, but the angle tangent θ(t) has values in the [-π, π] or [0, 2 π] values. The θ(t) function generally contains 2π dimensional discontinuities. That is why, the cumulative angular function is introduced in order to eliminate the discontinuity problem. Intuitively, the angle tangent indicates the change of the shape's contour angular directions. This is very important for the human perception. The shape can be then represented by the tangent of the angle formed by 2 points situated on the shape's contour:

$$\theta(t) = \arctan \frac{y(t) - y(t-w)}{x(t) - x(t-w)} \qquad (2)$$

where w can be any whole, and in practice it is the sample step of the points on the contour.

The cumulative angular function p(f), defined by Zahn and Roskies [Zahn et. al, 1979] is the angular difference between the z(t) position and the initial position z(0) on the shape's contour:

$$\varphi(t) = [\theta(t) - \theta(0)] \bmod (2\pi) \qquad (3)$$

The nominated angular function ψ(t) is defined as:

$$\Psi(t) = \varphi(\frac{L}{2\pi} t) - t \qquad (4)$$

where L represents the perimeter of the shape.

The diminution of t from the angular cumulative function implies: ψ(t) =0 when the shape is a circle and ψ(t) <>0 when the shape is not a circle. This function is invariant at rotation, scalar and of translation.

2.1.3 The maximum area function

We suggested a new method in order to represent the shape. The weight centre C is calculated:

$$x_c = \frac{\sum_{i=o}^{N-1} x_i}{N} \quad y_c = \frac{\sum_{i=o}^{N-1} y_i}{N} \qquad (5)$$

We have defined the maximum axe of the shape as being the segment that unites the weight centre with the point that is situated on the contour at the maximum distance from the weight centre: [P_0C] so that: dist(P_0, C) = max. The maximum area function $a_m(t)$ is calculated as being the area of the triangle formed by P_0, C and P_i, where P_i are points on the shape's contour.

$$a_m(t) = \frac{l_{CP_0} \cdot l_{CP_0}(t-w) \cdot \sin\theta}{2} \qquad (6)$$

where w can be any whole, and in practice it is the sample step of the points situated on the contour. l_{CP0} is the length of the CP_0 segment, and $l_{CP(t-w)}$ is the length of the segment $CP_{(t-w).}$

The area of the formed triangle CP_0P_1 is:

$$a_m = \frac{l_{CP_0} \cdot l_{CP_0} \cdot sin\left(arcsin\frac{|x_c - x_{P_0}|}{l_{CP_0}} + arcsin\frac{|x_c - x_{P1}|}{l_{CP1}} \right)}{2} \qquad (7)$$

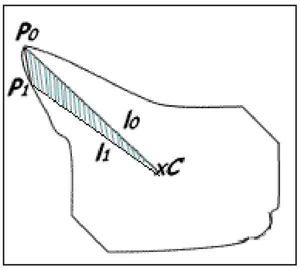

Fig.1. The maximum area function

The maximum area function is invariant at rotation, translation and scalation.

3. THE INDEXATION OF THE SHAPES USING FOURIER DESCRIPTORS

The Fourier transforming of the shapes signature functions is extremely used in the shape analysis, although there are other attempts to explore the retrieval based on the [Sajjanhar, 1997] shape.

The coefficients of the Fourier transformate form the shape Fourier Descriptors. These descriptors represent the shape of an object in the frequency domain. The low-frequency descriptors contain information about the general characteristics of the shape. Although the number of coefficients generated by the transformation is large, a certain subset of coefficients is enough to capture all the characteristics of the shape. The high-frequency information describes the little details of the shape and this is not as important in the description of the shape, so it can be ignored. As a result, the dimensions of the Fourier descriptors used to index the shapes, are significantly reduced.

For a certain shape signature described in the above section s(t), t = 0, 1, ..., L, assuming that it has been nominised at N points, the discreet Fourier transformate of the s(t) function is:

$$u_n = \frac{1}{N}\sum_{t=0}^{N-1} s(t)\exp\left(\frac{-j2\pi nt}{N}\right) \qquad (8)$$

The u_n, n = 0, 1, , N-1 coefficients are usually named shape Fourier descriptors (FD), written FD_n, n = 0, 1, …., N-1.

In the image retrieval process based on shape, the user is only interested in the contour characteristics of similar shapes, the position, the dimension, the rotation, are not important. In order to modelate shapes and to make them comparable, their representations must be invariated at translation, rotantion and scalation.

The invariation of shapes is hard to achieve in the spatial domain, especially rotation, this assumes complex calculation. The invariation of shapes is easy to achieve for the Fourier descriptors (FD).

All 3 signatures described above are invariated at translation, although the corresponding FD are also invariated at translation. The invariation at rotation of the FD is realised by ignoring the phase and by taking in account only the magnitude of the FD descriptors. The invariant at scalation is realised by dividing the magnitude values of all descriptors. For the central distance, because the signature function is real there are only N/2 different frequencies of the Fourier transformate, so there are enough only N/2 FD to index the shape.

The invariant at scalation is obtained by dividing the magnitudes of the first N/2 FD with the DC component.

$$f = \frac{|DF_1|}{|DF_0|}, \frac{|DF_2|}{|DF_0|}, \ldots, \frac{|DF_{N/2}|}{|DF_0|} \qquad (10)$$

The angular cumulative function is in itself invariant at translation, rotation and scalation, so the derived FD can be used directly to index the shape. But due to the fact that this function is real, there are only N/2 different frequencies of the Fourier transformate, so there are enough only N/2 FD to index the shape.

The characteristic vector in order to index the shape described by the angular cumulative function is:

$$f = [|DF_0|, |DF_1|, \ldots, |DF_{N/2}| \qquad (11)$$

And the maximum area function is in itself invariant at translation, rotation and scalation, although the derived FD can be used directly to index the shape. But due to the fact that this function is real, there are only N/2 different frequencies of the Fourier transformate, so there are enough only N/2 FD to index the shape.

The characteristic vector in order to index the shape described by the maximum area function:

$$f = [|DF_0|, |DF_1|, \ldots, |DF_{N/2}| \qquad (12)$$

For 2 shapes idexed by the FD characteristic vectors FD f_m= [f_m^1, f_m^2,..,f_m^N], and f_d= [f_d^1, f_d^2,..,f_d^N], because both of the characteristics are invariated at translation, rotation and scalation, the Euclidiana distance can be applied between the two characteristic vectors as the metrics of similitude:

$$d = \left(\sum_{i=0}^{N} \left|f_m^{\,i} - f_d^{\,i}\right|^2\right)^{\frac{1}{2}} \qquad (13)$$

where N is the number of Fourier coefficients used to index the shape.

4. THE INDEXATION OF THE SHAPE USING GEOMETRIC MOMENTS

The contour-based descriptors are used only if the shape edges can be represented. The utilization of these descriptors is difficult in the case of disjunct shapes with content inside them. Because of that, the contour – based shape descriptors have limited applicability. In image retrieval, it is prefferable that a descriptor could be application independent.

The region-based methods use the information from shape contour and from inside, and the edge detection is not necessary. In this paper, the geometric moment descriptor is studied and implemented. The technique based on geometric moment invariants for shape representation and the similitude metric are extremly used in object recognition [Prokop, 1992].

The geometric moment invariants are derived from shape moments and are invariant to 2D shape geometric transformation. The central moments of p+q order of a bidimensional shape represented by $f(x, y)$ function are given by:

$$\mu_{pq} = \sum_{p}\sum_{q}(x-\bar{x})^{p}(y-\bar{y})^{q} f(x,y) \quad (14)$$

where $\bar{x} = \mu_{10}/m$ and $\bar{y} = \mu_{01}/m$, m is the weight of the region shape. μ_{pq} are invariant to translation.

7 geometric moments normalized, invariant to translation, rotation and scaling, are well known in literature. A vector consisting of the seven moment invariants are used to index each shape in the database:

$$f = (\Phi_1, \Phi_2, \Phi_3, \Phi_4, \Phi_5, \Phi_6, \Phi_7) \quad (15)$$

The distance between two feature vectors is determined by city block distance. The values of computed moment invariants are usually small, and the values of moment invariants of higher order are closer to zero in some cases. The advantage of GMD is because it offers a compact representation of shape and the computation is low.

5. EXPERIMENTS

In order to test the performance of the image retrieval based on the shape characteristic using the Fourier descriptors, we develop a software system, based on Java technologies. In order to test the retrieval performances the database with MPEG-7 contours has been used, from the real world and medical surgical instruments. The sets of shapes are used to test the invariant at scaling, rotation and translation.
The experiments have been done on each class of images in order to analyze which shape descriptor based on contour will give the best results after the content-based visual retrieval after the shape characteristic.

In figure 2, 3 and 4 experiments are presented.

6. CONCLUSION

In this paper there is presented an application of a content-based visual search on shape characteristic in an image database.

There have been studied algorithms that use information from shape edge and from shape region. Conventional FD descriptors are derived by applying Fourier transformation on a one-dimensional signature function. More shape signatures has been used for deriving Fourier descriptors, all these being described below.

A new signature function has been proposed for improving existing signature methods, this is called *surface signature*. In this paper, the geometric moment descriptor is studied and implemented as region based shape descriptor.

As it can be observed from experiments and from search results comparison, the method with best results is central distant signature, followed by maximum surface signature ant this last one with very good results on all classes of images. Cumulative angular distance has recorded the worst results for all classes of images. The geometric moment descriptor records good results (four relevant retrieved images from 6 retrieved images) on the image collection of medical surgical instruments.
In the future, the study will be extended by improving contour detection algorithm (edge enclosing, edge thinning), taking into consideration also other visual characteristics: color and texture.

REFERENCES

Aslandogan, Y. A., and C. T. Yu (1999). Techniques and Systems for Image and Video Retrieval. In *IEEE Trans. On Knowledge and Data Engneering, 11(1):56-63.*

Berretti, S., A. D. Bimbo and P. Pala (2000). Retrieval by Shape Similarity with Perceptual Distance and Effective Indexing. In *IEEE Transactions on Multimedia, 2(4):225-239.*

Cojocaru, D. (2003). *Achizitia, prelucrarea si recunoasterea imaginilor*, Editura Universitaria, ISBN 973-8043-146-6

Fan, S (2001). Shape Representation and Retrieval using Distance Histograms. *Technical Report – University of Alberta.*

Pass, G. and R. Zabih (1996). Histogram refinement for content based image retrieval. In *IEEE Workshop on Appl. of Computer Vision 96-102*.

Pavlidis, T. (1980). Algorithms for Shape Analysis of Contours and Waveforms. In *IEEE Trans. PAMI- 2(4):301-312*.

Prokop R. J. and A. P Reeves (1992). A Survey of Moment-based Techniques for Unoccluded Object Representation and Recognition. In *Graphical Models and Image Processing, Vol.54, pp.438-460.*

Sajjanhar, A. (1997). *A Technique for Similarity Retrieval of Shapes*. Master thesis, Monash University, Australia, 1997.

Sajjanhar, A. and G. Lu (2003). Effect of Spatial Coherence on Shape Retrieval. In *CISST'03 Proceedings, Las Vegas, 23-26 June.*

Zahn, C.T. and R.Z. Roskies (1972). Fourier Descriptors for Plane closed Curves. In *IEEE Trans. On Computer, C-21(3):269-281*.

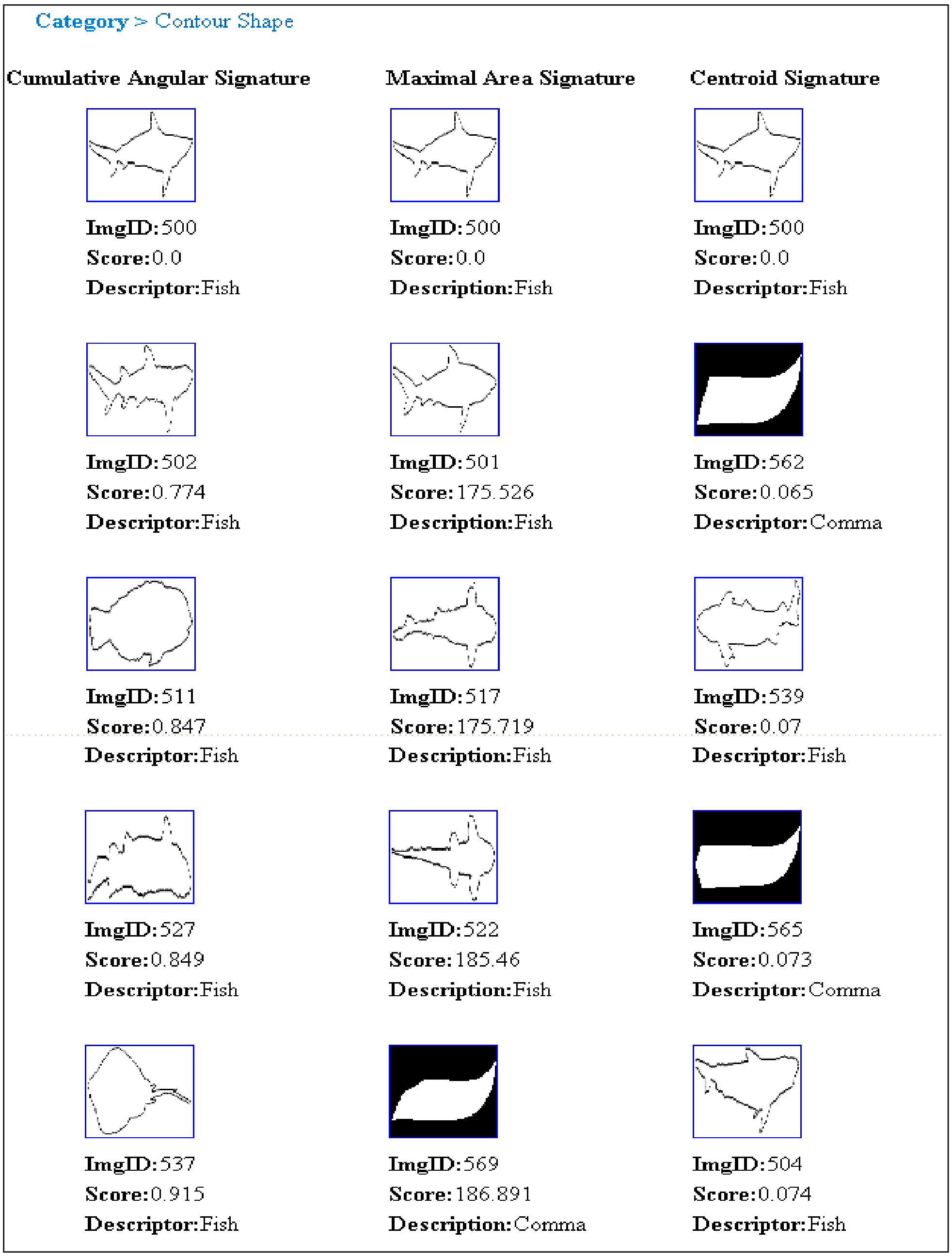

Fig. 2. Query 1 - Image class. Contour - Fishes

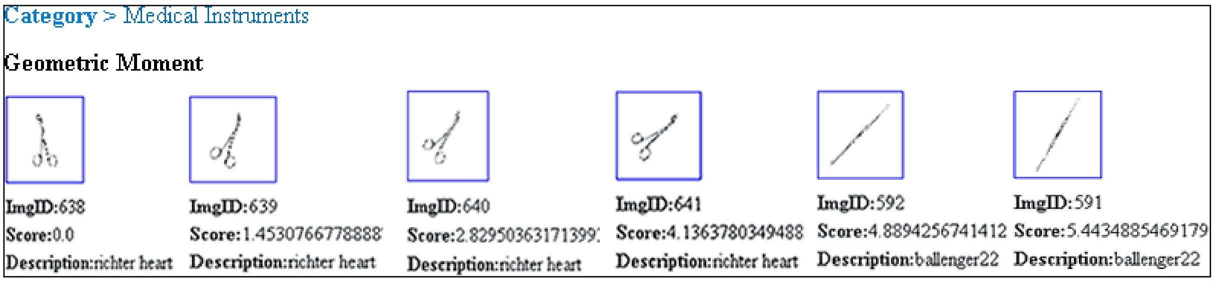

Fig 3. Query 2 - Image class. Medical Surgical Instruments

Fig. 4. Query 3- Image Class: Contour-Comma