A NEW SHAPE SIGNATURE FOR FOURIER DESCRIPTORS

A. El-ghazal [(1)], O. Basir [(1)] and S. Belkasim [(2)]

[(1)] Department of Electrical and Computer Engineering, University of Waterloo,
Waterloo, ON., Canada, N2L 3G1
Email: akrem@pami.uwaterloo.ca , obaisr@uwaterloo.ca
[(2)] Department of Computer Science, Georgia State University, Atlanta, GA., USA, 30303
Email: sbelkasim@cs.gsu.edu

ABSTRACT

Shape-based image description is an important approach to Content-based Image Retrieval (CBIR). A variety of techniques are reported in the literature that aim to represent objects based on their shapes; each of these techniques has its advantages and disadvantages. Fourier Descriptor (FD), a simple yet powerful technique, has attractive properties such as rotational, scale, and translational invariance. In this paper we investigate this technique and present a novel shape registration method for extracting Fourier descriptors. When evaluated against Curvature Scale Space (CSS) and Zernike Moments (ZM) in shape-based image retrieval, the proposed technique exhibits superior performance.

***Index Terms*—**Fourier descriptors, Image retrieval, Shape.

1. INTRODUCTION

In recent years, there has been an enormous increase in the size of digital image databases. The ease and convenience of capturing and transmitting digital images between digital cameras and image databases residing on dedicated servers is a contributing factor in the immense growth of such databases. These databases cover a wide range of applications ranging from medical to multimedia. The storage format of such image data is relatively standardized, but the retrieval process of the images tends to be complex. Image retrieval becomes a critical tool for searching and handling images queried by databases clients.

Typically, images in a database are retrieved based on either textual information or content information. Early retrieval techniques were based on textual annotation of images. Images were first annotated with text and then searched based on their textual tags. However, text-based techniques have many limitations, including their reliance on manual annotation, which is a very difficult process for large data sets. Furthermore, the rich content typically found in images and the subjectivity of the human perception make the task of describing images using words a difficult if not impossible task.

To overcome these difficulties, Content Based Image Retrieval (CBIR) was proposed [1]. This approach to image retrieval relies on the visual content of images, rather than textual annotations, to search for images and, hence, has the potential to respond to more specific user queries. CBIR techniques use such visual contents as color, texture and shape to represent and index the image. Visual contents such as color and texture have been explored more thoroughly than shape .The increasing interest in using the shape features of objects for CBIR is not surprising, given the considerable evidence that natural objects are recognized primarily by their shape [2, 3]. A survey of users on the cognition aspects of image retrieval indicates that users are more interested in retrieval based on shape than on color and texture [4]. However, retrieval based on shape content remains a more difficult task than that based on other visual features [2].

During the last decade, significant progress has been made in both the theoretical and practical research aspects of shape-based image retrieval [5, 6]. There are mainly two approaches to shape representation, namely, the region-based approach and the boundary-based approach (also known as contour-based approach). Region-based techniques often use moment descriptors to describe shapes. These descriptors include geometrical moments [7], Zernike moments [8, 9], pseudo-Zernike moments [10], Legendre moments [8], and Tchebichef moments [11]. Although region-based approaches are global in nature and can be applied to generic shapes, they fail to distinguish between objects that are similar [12].

In many applications the internal content of the shape is not as important as its boundary. Boundary-based techniques tend to be more efficient for handling shapes that are describable by their object contours [12]. Many boundary-based techniques have been proposed in the literature, including Fourier descriptors [13], Curvature Scale Space (CSS) [14], wavelet descriptors [15], and contour displacement [16]. The most promising boundary-based approaches are those that use Fourier descriptors (FDs). This is due to the fact that the FDs are based on the well-known Fourier theory, making them easy to compute, and simple to normalize and interpret. In addition, the FDs are well suited to online image retrieval because of their computational efficiency and compactness. To derive the FDs of an image, the 2-D image is converted to 1-D signature. Many signatures have been proposed in the literature. The centroid distance (CD), triangular centroid area (TCA), complex coordinate (CC), and chord-length distance(CLD) are some notable signatures which have been used to drive FDs [17]. Fourier descriptors derived from different signatures can have significantly different effects on the result of retrieval [17]. In this paper, we propose a new signature, namely Farthest Point Distance (FPD), and compare it with other frequently used shape signatures.

The paper is organized as follows: Section 2 introduces the proposed signature. Section 3 explains how the Fourier transform along with a normalization scheme is applied to the shape signatures. In Section 4, experimental results are presented to compare the proposed signature with frequently used signatures. In addition, the proposed signature is compared with two notable techniques: the Curvature Scale Space (CSS) and the Zernike Moments (ZM) techniques. Conclusions derived from the study and suggestions for future work are presented in Section 5.

2. FARTHEST POINT DISTANCE SIGNATURE

In this section we present Farthest Point Distance (FPD), a novel technique that exploits some differential properties of shapes, such as corner points and/or transition details.

The Farthest Point Distance (FPD) is developed to overcome some of the shortcomings of existing techniques. The value of the signature at a given point *a* is defined as the distance between *a* and the point farthest from it, say *b*. The signature is calculated by adding the Euclidian distance between point *a* and the centroid *c* to that between the centroid *c* and the farthest point.

Assuming that shape boundary coordinates $(x(u),y(u)), u=0,1,\ldots N-1$, have been extracted, the FPD signature at boundary point $(x(u),y(u))$ is calculated as follows:

$$FPD(u) = \sqrt{([x(u)-x_c]^2 + [y(u)-y_c]^2)} + \sqrt{([x_{fp}(u)-x_c]^2 + [y_{fp}(u)-y_c]^2)} \qquad (1)$$

where $(x_{fp}(u), y_{fp}(u))$ is the farthest point from $(x(u),y(u))$, and (x_c, y_c) is the centroid of the shape.

Figure 1 depicts how the distance from point *a* to its farthest point b is calculated. This signature captures distances between corners. Transition points and corners are elements of focus to the human visual system. In most shape matching techniques, corner points play a major role. Figure 2 depicts examples of FPD signatures for two different classes. It can be seen from the figure that shapes within the same class have similar FPD signatures.

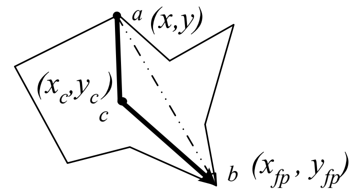

Fig.1 Example of the Farthest Point Distance (FPD) signature.

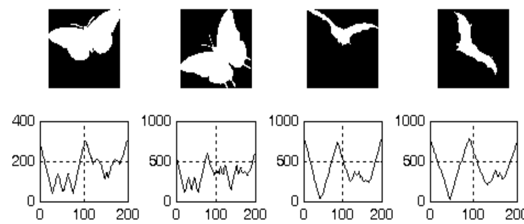

Fig.2 The FPD signatures for shapes from two different classes.

3. GENERATION OF FOURIER DESCRIPTORS

The Discrete Fourier Transform (DFT) of an arbitrary signature *Z(u)* is defined as [18]:

$$a_n = \frac{1}{N}\sum_{u=0}^{N-1} Z(u).e^{-j2\pi nu/N} \qquad (2)$$
$$n = 0,1 \ldots. N-1$$

The coefficients a_n ($n=0,1,..N-1$) are called the Fourier descriptors (FDs) of the shape, and are denoted by FD_n. The rotation invariance of the FDs can be established by taking into consideration the magnitude values of the descriptors and ignoring the phase information.

Since the Farthest Point Distance (FPD) is real valued, only half of the FDs are needed. It is straightforward to prove that the FPD signature is translation invariant. Scale invariance is achieved by dividing the magnitude of the first half descriptors by the DC components [19]:

$$f = [\frac{|FD_1|}{|FD_0|}, \frac{|FD_2|}{|FD_0|}, \ldots., \frac{|FD_{N/2}|}{|FD_0|}] \qquad (3)$$

The similarity measure between two shapes indexed with Fourier descriptors is the Euclidian distance between the two Fourier descriptors vectors.

4. EXPERIMENTAL RESULTS

Given the lack of a standard database, the evaluation of the different shape signatures is not an easy task. Researchers in this field tend to develop their own databases, which are often limited in size or application scope or both. The MPEG-7 developers have set up a database of a reasonable size and generality [12]. This database consists of three main sets: set A, set B, and set C. Set A consists of two subsets A_1 and A_2, and each subset includes 420 images. A_1 is used for testing scale invariance; set A_2 is used to test for rotation invariance. Set B consists of 1400 images that are classified into 70 classes, each class having 20 images. Set B is used to test for similarity-based retrieval performance, and to test the shape descriptors for robustness to various arbitrary shape distortions including rotation, scaling, arbitrary skew, stretching, defection, and indentation. For these reasons, set B is selected to evaluate the performance of the proposed signature. Set C consists of 200 affine transformed Bream fish and 1100 marine fish that are unclassified. The 200 bream fish are designated as queries. This set is used to test the shape descriptors for robustness to non-rigid object distortions. Samples of the images from this database are depicted in Figure 3.

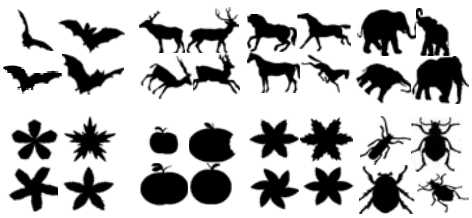

Fig. 3 Samples of images from set B of the MPEG-7 database.

To evaluate the performance of the different techniques in image retrieval, an efficient performance measure is required. The precision and recall measures are the most commonly used measures and are deemed appropriate for measuring retrieval performance on classified datasets.

Precision measures the retrieval accuracy, whereas recall measures the capability to retrieve relevant items from the database [17].

To evaluate the performance of the proposed signature, experiments are conducted using set B of the MPEG-7 database. All the contours of the images are resampled to so as each will consist of 128 points.

The proposed farthest distance signature is compared with the best performing signatures reported in the literature [17]. These signatures include the centroid distance (CD), the triangular centroid area (TCA), the complex coordinate (CC), and the chord-length distance (CLD) signatures [17]. The comparison is evaluated in terms of the recall and precision measures. The average precision rates for low ($\leq 50\%$) and high (>50%) recalls are shown in Table 1.

From Table 1, it can be seen that the proposed FPD signature's performance is the highest, whereas the CLD and CC signatures' performance are the lowest among the signatures. The TAC signature is much better than the CLD and CC signatures. The CD and FPD signatures have comparable results; however, in the case of high recall the FPD does better than the CD.

Table 1. The average precision for low and high recalls for the FPD and other signatures.

Signature	Low Recall The average precision for recall rates $\leq$ 50%	High Recall The average precision for recall rates $>$50%
FPD	75.82 %	42.13 %
CD	75.69 %	41.77 %
TCA	73.40 %	38.50 %
CLD	57.80 %	24.00 %
CC	64.76 %	22.59 %

In another set of experiments, the proposed farthest point distance signature is combined with four simple global descriptors (SGD): solidity, circularity, eccentricity, and aspect ratio [12]. The combined signatures are then compared with the Zernike moments [9] and the curvature scale space (in this case the CSS is combined with the four simple global descriptors to maintain consistency) [20]. The recall and precision curves using set B of the MPEG-7 database are plotted in Figure 4. The proposed signature outperforms the CSS technique, as can be deduced from Figure 4. The ZM technique and the proposed signature yield comparable results in the low recall case; however, the proposed signature does better in the case of high recall. The average precision rates for low and high recalls are shown in Table 2.

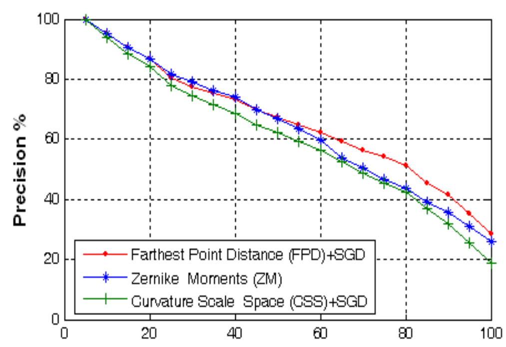

Fig. 4 Precision-recall curves of the proposed signature combined with SGD, curvature scale space combined with SGD, and Zernike moments.

Table 2. The average precision rates of low and high recalls for the FPD+SGP, ZM and CSS+SGP methods.

Method	Low Recall	High Recall
	The average precision for recall rates $\leq$ 50%	The average precision for recall rates $>$50%
FPD+SGP	81.56 %	56.61 %
ZM	81.97 %	51.59 %
CSS+SGP	78.61 %	48.03 %

5. CONCLUSIONS

This paper presents a new Fourier descriptor shape signature. The proposed signature is evaluated against several commonly used shape signatures. The performance of the proposed signature is examined based on several experiments using set B of the widely known MPEG-7 database.

The experimental results demonstrate that the proposed signature and the centroid distance (CD) signature yield comparable results; however, the proposed signature (FPD) performs better in the case of high recall. This improvement is due to the fact that the FPD signature tends to capture the information about an object's corners, which are shape points at which we focus our visual attention.

Furthermore, to enhance the ability of the proposed signature in capturing global shape information, it is combined with four global descriptors. The performance of the signature is then compared with two commonly used techniques: the curvature scale space and the Zernike moments. The results show that the proposed signature does better than the curvature scale space in both high and low recall. The results also show that the proposed signature does better than the Zernike moments for the high recall while maintaining comparable results for the low recall.

The overall results demonstrate the effectiveness of the farthest point distance signature for image retrieval applications. For future work, the computational complexity of the proposed signature will be evaluated, and compared with other techniques.

6. REFERENCES

[1] K. Hirata and T. Kato, "Query by visual example - content-based image retrieval," in *EDBT'92, Third International Conference on Extending Database Technology*, pp. 56-71, 1992.

[2] D. Mumford, "Mathematical theories of shape: Do they model perception? "in *Proceedings of SPIE Conference on Geometric Methods in Computer Vision*, vol. 1570. San Diego, California, pp. 2-10, 1991.

[3] I. Biederman, "Recognition-by-components: a theory of human image understanding," *Psychological Review* vol. 94, pp. 115-147, 1987.

[4] L. Schomaker, E. d. Leau, and L. Vuurpijl, "Using Pen-Based Outlines for Object-Based Annotation and Image-Based Queries," *Visual Information and Information Systems*, pp. 585-592, 1999.

[5] Z. Wang, Z. Chi, and D. Feng, "Shape based leaf image retrieval," presented at IEE Proceedings, Vision Image and Signal Processing, 2003.

[6] M. E. Celebi and Y. A. Aslandogan, "A Comparative Study of Three Moment-Based Shape Descriptors," in *International Conference on Information Technology: Coding and Computing (ITCC'05)*, vol. 1, pp. 788-793, 2005.

[7] M. Hu, "Visual Pattern Recognition by Moment Invariants," *IRE Transactions on Information Theory*, vol. IT-8, pp. 179-187, 1962.

[8] M. Teague, "Image analysis via the general theory of moments," *Journal of the Optical Society of America*, vol. 70, pp. 920-930 1980.

[9] A. Khotanzad, "Invariant Image Recognition by Zernike Moments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 489-497 1990.

[10] S. O. Belkasim, M. Shridhar, and M. Ahmadi, "Pattern Recognition with Moment Invariants: A Comparative Study and New Results," *Pattern Recognition*, vol. 24, pp. 1117-1138, 1991.

[11] R. Mukundan, S. H. Ong, and P. A. Lee, "Image Analysis by Tchebichef Moments," *IEEE Transactions on Image Processing*, vol. 10, pp. 1357-1364, 2001.

[12] F. Mokhtarian and M. Bober, *Curvature scale space representation: Theory application and MPEG-7 standardization*: Kluwer Academic 2003.

[13] T. P. Wallace and P A Wintz, "An efficient three dimensional aircraft recognition algorithm using normalized Fourier descriptors," *Computer Graphics and Image Processing*, vol. 13, pp. 99-126, 1980.

[14] F. Mokhtarian and A. Mackworth, "Scale-based description and recognition of planar curves and two-dimensional shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, pp. 34-43, 1986.

[15] G. Chauang and C. Kuo, " Wavelet descriptor of planar curves: Theory and applications," *IEEE Transaction on Image Processing*, vol. 5, pp. 56-70, 1996.

[16] T. Adamek and N. E. O'Connor, "A Multiscale Representation Method for Nonrigid Shapes with a Single Closed Contour," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 742-753, 2004.

[17] D. S. Zhang and G. Lu, "Study and Evaluation of Different Fourier Methods for Image Retrieval," *International Journal of Computer Vision*, vol. 23, pp. 33-49, 2005.

[18] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*: Addison-Wesley, 2002.

[19] H. Kauppinen, T. Seppänen, and M. Pietikäinen, "An Experimental Comparison of Autoregressive and Fourier-Based Descriptors in 2D Shape Classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, pp. 201-207 1995.

[20] S. Abbasi, F. Mokhtarian, and j. Kittler, "Curvature Scale Space Image in Shape Similarity Retrieval," *Multimedia Systems*, vol. 7, pp. 467-476, 1999.