

Correspondence

---

## Efficient Shape Retrieval Using Elliptical Shape Coding

JIA-MING LI AND WEI-YANG LIN\*

Department of Computer Science and Information Engineering, National Chung Cheng University,  
Taiwan

### ABSTRACT

This paper presents a novel shape representation, called *elliptical shape coding*, and uses it to develop an efficient shape retrieval algorithm. The elliptical shape coding transforms an original shape, which is a single, closed contour in our experiments, into a periodic signal. The alignment between two shapes is accomplished efficiently by performing convolution of the corresponding periodic signals. Unlike previous approaches, we taking advantage of a priori knowledge hidden in the shape structure to produce shape matches with a significantly lower computational cost. Experimental results on an MPEG-7 shape database demonstrate the effectiveness of the proposed method.

**Key word:** shape matching, shape retrieval, object presentation.

### 1. INTRODUCTION

The recognition of shapes in images is one of the fundamental issues in the development of content-based image retrieval systems. It is challenging due to the deformations within one shape category. The robustness of a shape representation to such variations can have significant impact on the retrieval accuracy. There has been extensive works in the area of shape representation for object recognition. Shapes have been represented by their outline curves (Gdalyahu & Weinshall, 1999; Basri, Costa, Geiger & Jacobs, 1998), structural graphs (Coughlan & Ferreira, 2002; Luo & Hancock, 2001), shape context (Belongie, Malik & Puzicha, 2002), shock graphs (Sebastian, Klein & Kimia, 2004), and curvature scale space (Mokhtarian, Abbasi & Kittler, 1997), among others. In this paper, we introduce a novel shape representation, called *elliptical shape coding*. A shape retrieval algorithm utilizing the elliptical shape coding is also presented. We emphasize that we are not presenting a complete solution to shape retrieval, but rather a shape descriptor suitable for shape retrieval applications.

The rest of this paper is organized as follows. Section 2 provides the details of the proposed elliptical shape coding and shape retrieval algorithm. Section 3 shows experimental results on an MPEG-7 database. We make concluding remarks and discuss future works in Section 4.

### 2. RELATED WORKS

---

\* Correspondence author. E-mail: wylin@cs.ccu.edu.tw

There have been many approaches to shape matching. How to define a shape and how to measure the shape similarity are important steps before matching two shapes. According to a survey conducted in this study, there are two major approaches to shape matching: feature-based and brightness-based.

Feature-based approaches involve the spatial property such as edge or silhouette features. Zahn and Roskies (1972) used Fourier descriptors to describe silhouettes. Amit, Geman and Wilder (1997) utilized decision trees by learning the spatial property of keypoints. Gdalyahu and Weinshall (1999) applied dynamic programming approaches using silhouette curves for matching. Although silhouettes are simple and efficient to compare, they ignore internal contours and are difficult to extract from real images. There are several approaches to the use of shape descriptors for general objects to provide greater discriminative power. Lowe (1999) used gray-level information around the keypoints, called SIFT. Belongie et al. (2002) proposed shape context to collect information around a point.

Brightness-based approaches make use of intensity values of images. Lades et al. (1993) used elastic graph matching that involves the use of geometry and photometric properties to describe features. Turk and Pentland (1991) used principal components analysis (PCA) in the area of face recognition. These approaches also have been used for handwritten digit recognition in Lecun, Bottou, Bengio and Haffner (1998), face recognition in Moghaddam, Jebara and Pentland (2000), and 3D object recognition in Murase and Nayar (1995).

### 3. SHAPE MATCHING AND RETRIEVAL

#### 3.1 The Elliptical Shape Coding

Let  $\psi$  denote the set of discrete sampling points along a contour:

$$\psi = [p_1 \quad p_2 \quad \dots \quad p_N] = \begin{bmatrix} x_1 & x_2 & \dots & x_N \\ y_1 & y_2 & \dots & y_N \end{bmatrix} \quad (1)$$

where  $N$  denotes the number of sampling points. The mean vector  $m$  and the covariance matrix  $C$  of the set  $\psi$  are given by

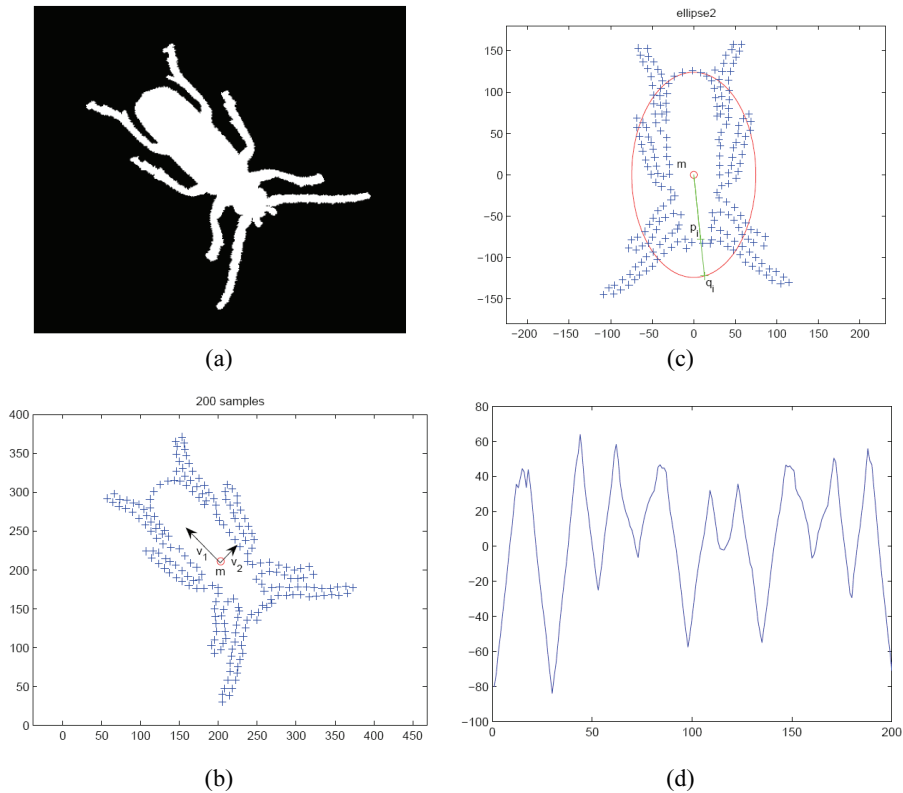
$$m = \left[ \frac{1}{N} \sum_{k=1}^N x_k \quad \frac{1}{N} \sum_{k=1}^N y_k \right]^T; \quad (2)$$

$$C = \frac{1}{N} \sum_{k=1}^N (p_k - m)(p_k - m)^T. \quad (3)$$

Since  $C$  is a symmetric matrix, there exists an orthogonal matrix  $V$  such that

$$C = VDV^T \tag{4}$$

where  $D$  is the diagonal matrix whose diagonal entries  $\lambda_1$  and  $\lambda_2$  are the eigenvalues of  $C$ . The eigenvectors, corresponding to the eigenvalues  $\lambda_1$  and  $\lambda_2$ , are denoted by  $v_1$  and  $v_2$ , respectively. Note that  $v_1$  and  $v_2$  are orthonormal vectors.



*Figure 1.* (a) Original shape in the form of binary image. (b) The set of points sampled along the outer contour of the shape.  $v_1$  and  $v_2$  are the eigenvectors of the covariance metric  $C$ . (c) An ellipse centered at  $m$  and having its major axis parallel to  $v_1$  is drawn. The feature value at the point  $p_i$  is defined as Equation (9). (d) We obtain a periodic signal  $s[n]$  by computing feature value at each point.

In analyzing the properties of a contour, it is convenient to work with a coordinate system associated with the contour. Hence, a new coordinate system is defined by using the origin  $m$  and coordinate frames  $v_1$  and  $v_2$ . An example is shown in Figure 1(b). Given a set of sampling points  $\psi$ , the coordinate transformation from the original coordinate system to the new coordinate system is given by

$$\begin{bmatrix} x_i' \\ y_i' \end{bmatrix} = \begin{bmatrix} v_1^T \\ v_2^T \end{bmatrix} (p_i - m) \tag{5}$$

where  $(x_i', y_i')$  represent the  $i$ -th point in terms of the new coordinate system. Figure 1(c) shows the point sequence after performing coordinate transformation. Then, an ellipse centered at the new origin and having its major axis parallel to the new  $y$ -axis is drawn. The ellipse could be specified by the following equation:

$$\frac{x'^2}{\lambda_1} + \frac{y'^2}{\lambda_2} = 1. \tag{6}$$

Or, it could be expressed as functions of one parameter:

$$x' = \sqrt{\lambda_1} \cos \theta; \tag{7}$$

$$y' = \sqrt{\lambda_2} \sin \theta, \tag{8}$$

where  $\theta \in [0, 2\pi]$ . An ellipse obtained in this way contains important information about a contour. For example, if an ellipse is long and thin, the shape behind it must be slender. Now, we could encode shape information into a periodical signal by using the ellipse. For the  $i$ -th point, we draw a straight line from that point to  $m$ . The intersection of the line and the ellipse is denoted by  $q_i$ . The feature value at the point  $p_i$  is defined as

$$s_i = \|p_i - m\| - \|q_i - m\|. \tag{9}$$

A periodic signal  $s[n]$  is obtained by repeating the same procedure on each point. For a contour with  $N$  sampling points, the elliptical shape coding generates a periodic signal  $s[n]$  with fundamental period  $N$ . An example is shown in Figure 1(d).

### 3.2 Shape Alignment

Because shapes are sampled with arbitrary starting points, alignment has to be performed before computing a similarity metric. Consider that a contour is sampled twice, each with different starting point. The periodic signals generated by the elliptical shape coding will be related by an unknown offset  $n_0$ . The discrete-time convolution of these two periodic signals is given by

$$s[n - n_0] \otimes s[n] = \sum_{k=\langle N \rangle} s[k - n_0] s[n - k] \tag{10}$$

where we use the notation  $n = \langle N \rangle$  to indicate summation over any  $N$  consecutive values and the symbol  $\otimes$  to denote periodic convolution. In order to estimate the value of  $n_0$ , we flip one of the signal and then compute the convolution of the flipped signal and the other one. Without losing generality, we assume that the flipped signal is  $s[n_0 - n]$ . The discrete-time periodic convolution of  $s[n_0 - n]$  and  $s[n]$  is given by

$$\xi[n] = s[n_0 - n] \otimes s[n] = \sum_{k=\langle N \rangle} s[n_0 - k]s[n - k] \tag{11}$$

where  $\xi[n]$  is also a periodic signal with fundamental period  $N$ . It is easy to show that the maximum value of the periodic signal  $\xi[n]$  is located at  $n = n_0$ . Hence, the alignment between two shapes is defined to be the solution of

$$\arg \max_n \xi[n]. \tag{12}$$

One could simply find the maximum value of  $\xi[n]$ . Then, the index of the maximum value provides the offset between the two shapes.

### 3.3 The Similarity Measure

After the shape alignment, the similarity  $\sigma_{i,j}$  between the  $i$ -th and  $j$ -th shapes is measured using the normalized cross-correlation (Lin, Boston & Hu, 2005) as follows:

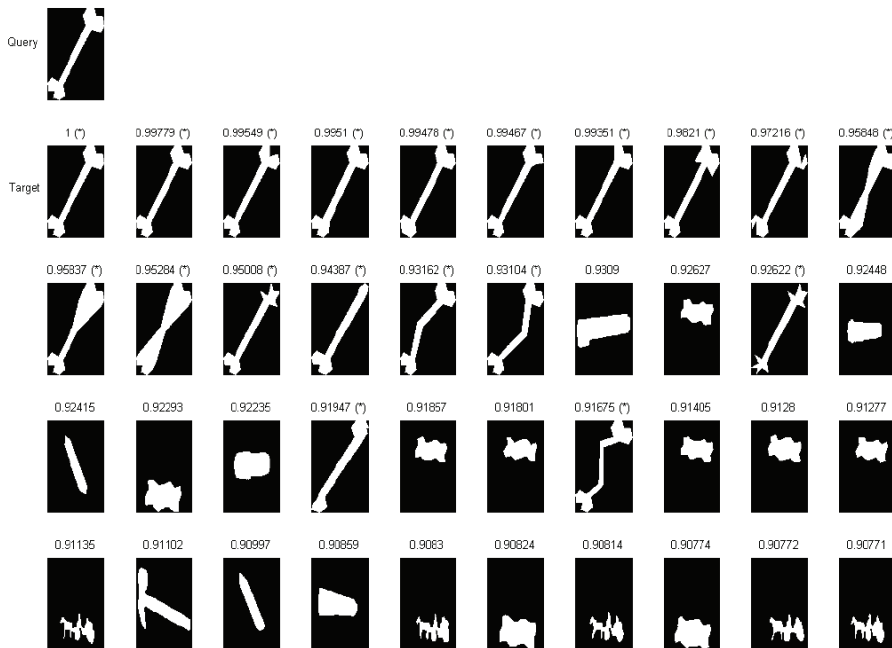
$$\sigma_{i,j} = \frac{\sum_{n=0}^{N-1} s_i[n]s_j[n]}{\sqrt{\sum_{n=0}^{N-1} s_i^2[n] \sum_{n=0}^{N-1} s_j^2[n]}}. \tag{13}$$

The value of the normalized cross-correlation ranges between  $-1$  and  $1$ . A higher value means a better match. The complexity of the proposed method consists of three parts. First, the computation of the mean vector and covariance matrix can be accomplished in  $O(N)$ , where  $N$  is the number of sampling points. Second, the construction of the periodic signal takes  $O(N)$ . Third, the normalized correlation costs  $O(N)$ . For a personal computer equipped with a Pentium IV 2.8GHz processor and 2GB memory, a single comparison of two shapes with  $N = 100$  takes about 0.69 milliseconds.

## 4. EXPERIMENTS

In this section, the proposed method is tested on the MPEG-7 CE shape-1 database (Latecki, Lak<sup>ˆ</sup>amper & Eckhardt, 2000), which is a widely accepted evaluation platform for shape matching. Note that our method has only one parameter  $N$ , which is the number of sample points on the outer contour of a shape. On the other hand, the state-of-the-art approaches have several parameters and their values could have significant impact on the matching performance. In general, parameter tuning is a nontrivial task for users who do not have a thorough understanding about the algorithm.

The MPEG-7 CE-Shape-1 database consists of 1400 shapes and 70 classes, i.e. each class has 20 shapes. The retrieval accuracy is reported by the so-called Bullseye test: Every shape is matched with all the other shapes in a dataset. We keep 40 shapes with the highest similarity scores and discard the others. Among the 40 retrieved shapes, we count the correct hits, i.e. the number of shapes belonging to the same class as the query shape. There are at most 20 hits for each query. The accuracy of shape retrieval is the ratio of the number of correct hits to the highest possible correct hits (which is  $20 \times 1400$  in this experiment). Figure 2 shows a query image and the 40 images retrieved by the proposed method.



*Figure 2.* A retrieval example on the MPEG-7 database. In this case, the number of correct hits is 19 and hence the retrieval rate is  $19/20 = 0.95$ .

We conduct sampling uniformly along contours and take  $N$  sampling points for each shape. Note that some shapes in the MPEG-7 database have mirror symmetry. To handle mirrored shapes, we match a query to an original shape and also its mirrored shapes, and keep matches with higher similarity scores. Table 1 shows the retrieval rates and average matching times for various  $N$ . We observed that the elliptical shape coding achieves high speed while maintaining good accuracy.

Table 1. *Retrieval accuracy (in %) and speed (in millisecond) for bullseye test on the MPEG-7 database*

$N$	50	75	100	125	150	175
retrieval rates (%)	68.23	68.44	68.49	68.48	68.52	68.43
average matching times(ms)	0.64	0.66	0.69	0.77	0.79	0.83

## 5. CONCLUSION AND FUTURE WORKS

Many approaches have been proposed for shape retrieval. But the accuracy of shape retrieval meets a bottleneck due to its subjective nature, i.e. different persons have different ideas about whether two shapes are similar. We tried to tackle this challenging issue by incorporating online learning from user feedback. However, most of the existing algorithms for shape retrieval are too slow for online applications. In this paper, we introduce a novel shape representation, called elliptical shape coding, and demonstrate its effectiveness in shape retrieval. The proposed algorithm achieves good accuracy with a significantly reduced retrieval time on the MPEG-7 database.

For the future works, we would like to exploit the relevance feedback that involves users in the shape retrieval process. In other words, users are allowed to label retrieval results as relevant or irrelevant. Using online learning from user-labeled samples, the system could iteratively adapt the retrieval results to meet the user's query concept.

## REFERENCES

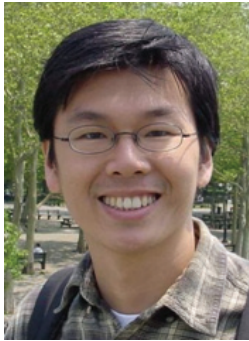
- Amit, Y., Geman, D., & Wilder, K. (1997). Joint induction of shape features and tree classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19, 1300-1305.
- Basri, R., Costa, L., Geiger, D., & Jacobs, D. (1998). Determining the similarity of deformable shapes. *Vision Research*, 38(15), 2365-2385.
- Belongie, S., Malik, J., & Puzicha, J. (2002). Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(4), 509-522.

- Coughlan, J. M., & Ferreira, S. J. (2002). Finding deformable shapes using loopy belief propagation. *Proceedings of European Conference on Computer Vision*, 453-468, London, UK.
- Gdalyahu, Y., & Weinshall, D. (1999). Flexible syntactic matching of curves and its application to automatic hierarchical classification of silhouettes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(12), 1312-1328.
- Lades, M., Vorbruggen, J. C., Buhmann, J., Lange, J., von der Malsburg, C., Wurtz, R. P., & Konen, W. (1993). Distortion invariant object recognition in the dynamic link architecture. *IEEE Transactions on Computers*, 42, 300-311.
- Latecki, L. J., Lakšamper, R., & Eckhardt, U. (2000). Shape descriptors for non-rigid shapes with a single closed contour. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 424-429, Hilton Head Island, SC, USA.
- Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86, 2278-2324.
- Lin, W.-Y., Boston, N., & Hu, Y. H. (2005). Summation invariant and its applications to shape recognition. *Proceedings of IEEE Conference on Acoustics, Speech, and Signal Processing*, 5, 205-208, Philadelphia, PA, USA.
- Lowe, D. G. (1999). Object recognition from local scale-invariant features. *Proceedings of IEEE International Conference on Computer Vision*, 2, 1150-1157, Corfu, Greece.
- Luo, B., & Hancock, E. R. (2001). Structural graph matching using the EM algorithm and singular value decomposition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(10), 1120-1136.
- Moghaddam, B., Jebara, T. & Pentland, A. (2000). Bayesian face recognition. *Pattern Recognition*, 33, 1771-1782.
- Mokhtarian, F., Abbasi, S., & Kittler, J. (1997). Efficient and robust retrieval by shape content through curvature scale space. In A. Smeulders & R. Jain (Eds.), *Image Databases and Multi-Media Search* (pp. 51-58). New Jersey, USA: World Scientific.
- Murase, H., & Nayar, S. (1995). Visual learning and recognition of 3-D objects from appearance. *Journal of Computer Vision*, 14, 5-24.
- Sebastian, T. B., Klein, P. N. & Kimia, B. B. (2004). Recognition of shapes by editing their shock graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(5), 550-571.
- Turk, M., & Pentland, A. (1991). Eigenfaces for Recognition. *Journal of Cognitive Neuroscience*, 3, 71-86.
- Zahn, C. T., & Roskies, R. Z. (1972). Fourier descriptors for plane closed curves. *IEEE Transactions on Computers*, 21, 269-281.





**Jia-Ming Li** received the B. S. degree in computer & information science from Soochow University, Taipei, Taiwan in 2006. He is currently a graduate student in Computer Science and Information Engineering, National Chung Cheng University. His research interests include pattern recognition and object classification.



**Wei-Yang Lin** received the BSEE degree from National Sun Yat-sen University, Taiwan, in 1994, and the MSEE and Ph. D. degrees from University of Wisconsin - Madison in 2004 and 2006, respectively. Since 2006, he has been with the Department of Computer Science and Information Engineering, National Chung Cheng University, Taiwan, where he is currently an assistant professor. His research interests include computer vision, biometric authentication, and multimedia signal processing.