INVARIANT CODES FOR SIMILAR TRANSFORMATION AND ITS APPLICATION TO SHAPE MATCHING

Eiji Yoshida and Seiichi Mita
To Toyota Technological Institute 2-12-1 Hisakata, Tempaku, Nagoya 468-8511, Japan

Keywords: Shape matching, curvature, shape classification, shape similarity measure.

Abstract: In this paper, we propose a new method for the measurement of shape similarity. Our proposed method encodes the contour of an object by using the curvature of the object. If one objects are similar (under translation, rotation, and scaling) in shape to the other, these codes themselves or their cyclic shift have the same values. We compare our method with other methods such as CSS (curvature scale space), and shape context. We show that the recognition rate of our method is 100 % and 90.40 % for the rotation and scaling robustness test using MPEG7-CE-Shape1 and 81.82 % and 95.14 % for the similarity-based retrieval test and the occlusion test using Kimia's silhouette. In particular, the value of the occlusion test is approximately 25 % higher than those of CSS, SC. Moreover, we show that the computational cost of our method is not so large by comparison our method with above methods.

1 INTRODUCTION

A measurement of shape similarity for shape-based retrieval in image databases should correspond with our visual perception. This basic property leads to the following requirements:

1. A shape similarity measure should present recognition of perceptually similar objects that are not mathematically identical.
2. It should not depend on scale, orientation, and position of objects.
3. A measure must return high similarity when we compare an object with those obtained by varying its shape by moderate articulation and occlusion. For instance, the similarity measure of all hands in Figure 1 must be large when they are compared with each other and small when they are compared with other objects such as heads, faces, and aeroplanes.
4. It should be free from digitization noise and segmentation errors.

We aim to apply a shape similarity measure to the classification of image databases, where the object classes are generally unknown. Therefore, a shape similarity measure is required to be universal in the sense that it allows us to identify and distinguish between objects of arbitrary shapes without any restriction on a shape assumption. In practice, the computational complexity to measure a shape similarity should be small. In particular, we concentrate our effort on the above requirements 2 and 3.

Our method encodes the contour of an object by using the curvature of the object. We use this code as a shape similarity measure. If two objects are perceptually similar (translation, rotation, and scaling) in shape, these codes themselves or their cyclic shift have the same values, and vice versa.

Our method is compared with previous ones such as the well-known Fourier descriptor, CSS (curvature scale space), and shape context. These previous methods have several drawbacks. For example, in the case of Fourier descriptors (FDs), the mapping from the original object to the representation features (e.g. FD magnitudes or phase) is not one-to-one, i.e. the original object cannot be uniquely reconstructed from the representation features. The computational cost of CSS is large. Our method can overcome this drawback.

This paper is organized as follows: Section 2 presents the outline of our method. Here, we define our code and describe the process to construct it and how it is used to compare the similarity of two objects. In Section 3, we report the experimental results obtained using the shape databases MPEG7_CE-Shape1 (Mokhtarian and Bober, 2003)

Yoshida E. and Mita S. (2008).
INVARIANT CODES FOR SIMILAR TRANSFORMATION AND ITS APPLICATION TO SHAPE MATCHING.
In *Proceedings of the Third International Conference on Computer Vision Theory and Applications*, pages 264-269
DOI: 10.5220/0001076002640269

and Kimia's silhouette (Belongie, Malik, and Puzicha, 2002). The experiment is composed of three parts with the following main objectives:

A: robustness to scaling and rotation by using MPEG7_CE-Shape1,

B: performance of the similarity-based retrieval by using Kimia's silhouette, and

C: robustness to occlusion by using Kimia's silhouette.

In Section 4, we summarize our conclusions.

Original Rotated Articulated Occluded

Figure 1: Variations of a sample shape.

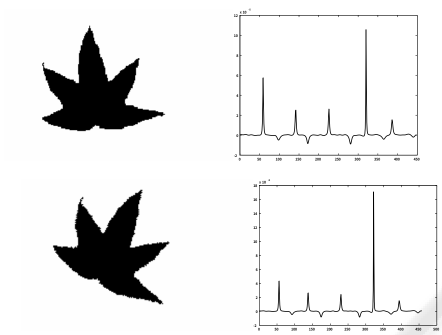


Figure 2: Curvature images of a leaf image and its rotated image.

2 PROPOSED METHOD

First, we discuss the properties of curvature; then, we define our code.

2.1 Curvature of a Planar Curve

The curvature of a planar curve is invariant under rotation and translation. It is also inversely proportional to the scale. Therefore, it is useful to compare objects by using their curvatures. However, direct use of curvature is difficult due to the digitization errors. Figure 2 describes the curvature images of a leaf and its rotation image. The values of their curvature functions are slightly different. On the other hand, the shapes of their curvature images are similar. In particular, the positions of their extreme points are almost equal to each other. This property is an important aspect of our method and extreme points of a curvature are used in our method.

2.2 Definition of our Code

Our method encodes the contour of an object as follows:

1. Compute the curvature of the object. For this, we use the same method as that used in CSS (Mokhtarian and Mackworth, 1992, Costa and Cesar, 2001).
2. Extract extreme points of the curvature function and select points of an object corresponding to them. Hereafter, we call these points as "interest points." Intuitively, these points are like the corner points of an object.
3. Let $\{p_1 = p_{n+1}, p_2, p_3, \ldots, p_n = p_0\}$ denote the set of interest points obtained by the above step. Then, we construct a code of the object defined as follows:

$$\begin{pmatrix} c_{11} & c_{21} & c_{31} & . & . & . & . & c_{n1} \\ c_{12} & c_{22} & c_{32} & . & . & . & . & c_{n2} \\ c_{13} & c_{23} & c_{33} & . & . & . & . & c_{n3} \\ c_{14} & c_{24} & c_{34} & . & . & . & . & c_{n4} \end{pmatrix} \quad (1)$$

where

$$c_{i1} = \frac{p_{i-1}p_i}{p_{i-1}p_i + p_i p_{i+1} + p_{i-1}p_{i+1}} \quad (2)$$

$$c_{i2} = \frac{p_i p_{i+1}}{p_{i-1}p_i + p_i p_{i+1} + p_{i-1}p_{i+1}} \quad (3)$$

($p_i p_{i+1}$ is the length of a segment $p_i p_{i+1}$ with $i = 1,2,3,\ldots,n$),

$$c_{i3} = {}^{l_i}\!/_{l} \quad (4)$$

(l_i is the area of a triangle $p_{i-1} p_i p_{i+1}$, $l = \sum_{i=1}^{n} l_i$),

and

$$c_{i4} = \begin{cases} 1 & (k_i > 0) \\ 0 & (k_i < 0) \end{cases} \quad (5)$$

(k_i is the value of the curvature of an object corresponding to p_i).

Figure 3 represents an outline of the construction of our code.

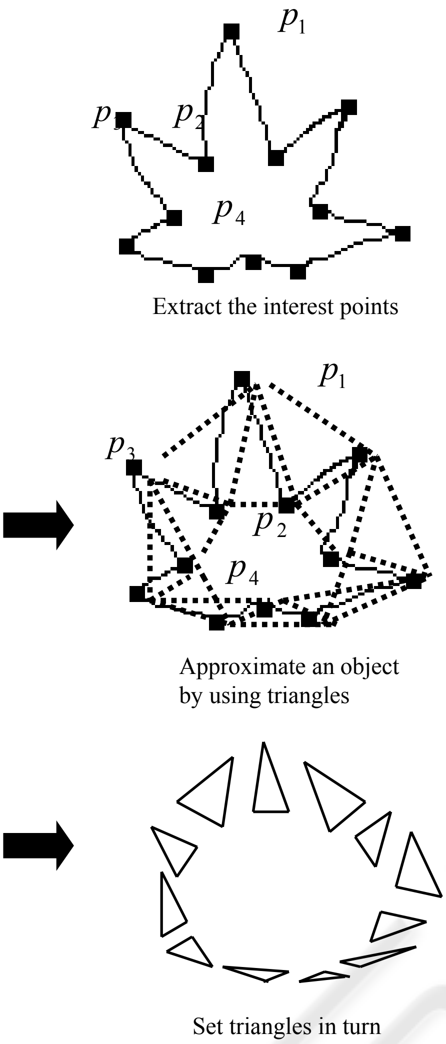


Figure 3: Outline of the construction of our code.

Each column of our code in (1) represents the information of triangles $p_{i-1}p_i p_{i+1}$ with $i = 1,2,3,...,n$. The meaning of the components c_{i1}, c_{i2}, c_{i3}, and c_{i4} of the i-th column of the above code is the following:

c_{i1}, c_{i2} : information of a side of $p_{i-1}p_i p_{i+1}$,

c_{i3} : information of the area of $p_{i-1}p_i p_{i+1}$,

c_{i4} : information of convexity.

By using this code, we compare the similarity between two objects. It is invariant under translation and scaling. Rotation causes only a cyclic shift of a column in it. Moreover, the mapping from an object to this code is injective.

2.3 Application to Shape Matching

The process of comparing two objects by using our code is as follows.

Let $\{p_1, p_2,..., p_m\}$, $\{q_1, q_2,..., q_n\}$ denote the sets of interest points of two objects A and B , respectively.

1. Compute the value of the curvature corresponding to p_i and q_j with $i = 1,2,3,..., m$ and $j = 1,2,3,..., n$ ($m \geq n$).
2. If $m > n$, remove the interest points of A such that the values of the curvature are the top $m - n$ ranked points in the ascending order of the absolute value.
3. Construct both codes.
4. The codes of A and B are denoted by (6) and (7), respectively

$$CA = \begin{pmatrix} a_{11} & a_{21} & a_{31} & . & . & . & . & a_{n1} \\ a_{12} & a_{22} & a_{32} & . & . & . & . & a_{n2} \\ a_{13} & a_{23} & a_{33} & . & . & . & . & a_{n3} \\ a_{14} & a_{24} & a_{34} & . & . & . & . & a_{n4} \end{pmatrix} \tag{6}$$

$$CB = \begin{pmatrix} b_{11} & b_{21} & b_{31} & . & . & . & . & b_{n1} \\ b_{12} & b_{22} & b_{32} & . & . & . & . & b_{n2} \\ b_{13} & b_{23} & b_{33} & . & . & . & . & b_{n3} \\ b_{14} & b_{24} & b_{34} & . & . & . & . & b_{n4} \end{pmatrix} \tag{7}$$

(The meaning of each component is the same as that in Section 2.2.). Then, we determine that the i-th column of CA and the j-th column of CB are common if the following conditions are held:

a. $| a_{i1} - b_{j1} | < t_1$,

b. $| a_{i2} - b_{j2} | < t_1$,

($t_1 \approx 0.1$)

c. $\begin{cases} 1/t_2 < a_{i3}/b_{j3} < t_2 & (b_{j3} \neq 0) \\ a_{i3} - b_{j3} = 0 & (otherwise) \end{cases}$

($t_2 \approx 2.5$)

d. $a_{i4} - b_{j4} = 0$.

We construct the $n \times n$ -matrix $COM = (com_{ij})$ such that if the i-th column of CA and the j-th column of CB are common, $com_{ij} = 1$; otherwise, $com_{ij} = 0$.

5. Compute the following number:

$$s = s_1 + s_2 + s_3 + s_4 \tag{8}$$

where

$$s_1 = 1 - \frac{1}{n}\#\{i \mid \exists j = 1,2,\ldots,n \ \ s.t.\ \ com_{ij} = 1\} \tag{9}$$

$$s_2 = 1 - \frac{1}{n}\max\{\sum_{i=1}^{n} com_{iu_i} \mid (u_1,u_2,\ldots,u_n) \in U\} \tag{10}$$

$(U = \{(i,i+1,i+2,\ldots,n,\ldots,i-1) \mid i = 2,\ldots,n\} \cup \{(1,2,3,\ldots n)\})$

$$s_3 = \sum_{i=1}^{n} s_3^{(i)} \tag{11}$$

with

$$s_3^{(i)} = \begin{cases} \min\{|a_{i3} - b_{j3}| \mid com_{ij} = 1\} & (if \ \sum_{i=1}^{n} com_{ij} \neq 0) \\ |a_{i3} - b_{u'_i 3}| & (else) \end{cases}$$

$$(\sum_{i=1}^{n} com_{iu'_i} = \max\{\sum_{i=1}^{n} com_{iu_i} \mid (u_1,u_2,\ldots,u_n) \in U\})$$

$$s_4 = \frac{m-n}{m+n} \tag{12}$$

We use above s to measure the similarity between two objects. We call s the similarity number.

6. For all i with $1 \le i \le t_3 - 1$ ($0.5n \le t_3 \le 0.8n$), remove the interest points of A and B such that their absolute values of curvature sorted in the ascending order are less than or equal to that of $(m-n+i)$-th and less than or equal to that of i-th, respectively. Thereafter, repeat steps 3 to 5 by replacing n with $n-i$.
7. We denote the similarity number of the i-th trial by s^i ($s^0 = s$). Compute the mean value of each similarity number, i.e.

$$S = \frac{1}{t_3}\sum_{i=0}^{t_3-1} s^i \tag{13}$$

S is the definition of the similarity between A and B.

Here, we provide an additional explanation about S and the parameters s_1, s_2, s_3, s_4. If S is small, the similarity between A and B is high, and vice versa. The role of s_1 is to measure how many common parts of a shape exist between A and B. It is used to measure the rough similarity. s_2 is used to calculate how many common connected parts of the shape exist between A and B. It is used to measure the close similarity. In fact, s_2 is not small unless the shapes of A and B are considerably close (e.g. A and B are similar in shape). s_3 is mainly used to compute the local difference between A and B. s_4 is large if $m-n$ is large. This parameter plays a role to distinguish dissimilar shapes.

Due to step 6, S remains small if the curvatures of A and B are similar; otherwise S becomes large. Therefore, S is small when the shapes of A and B are similar or almost similar (e.g. moderate articulation and occlusion).

3 EXPERIMENTS

The first experiment evaluates the robustness to scaling and rotation by using MPEG7_CE-Shape1. The second and third experiments evaluate the performance of the similarity-based retrieval and robustness to occlusion by using Kimia's silhouette.

3.1 Robustness to Rotation and Scaling

3.1.1 Robustness to Rotation

The database MPEG7_CE-Shape1 includes 420 shapes: 70 basic shapes and 5 derived shapes from each basic shape by rotation through angles: 9°, 36°, 45°, 90° and 150°. Each of these 420 images was used as a query image. The number of correct matches was computed in the top 6 retrieved images. Thus, the best result is 2520 matches. Figure 4 shows some shape instances in MPEG7_CE-Shape1.

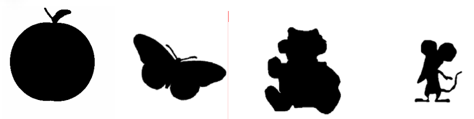


Figure 4: Shape instances in MPEG7_CE-Shape1.

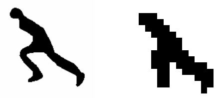
3.1.2 Robustness to Scaling

The database includes 420 shapes: 70 basic shapes are the same as in 3.1.1 and 5 shapes are derived from each basic shape by scaling digital images with factors 2, 0.3, 0.25, 0.2, and 0.1. Each of these 420 images was used as a query image. The number of correct matches was computed in the top 6 retrieved images. Thus, the best possible result is 2520 matches. In Table 1, the results of rotation and scaling tests are presented. The presented results except for our method are based on (Mokhtarian and Bober, 2003, Latecki, Lakamper and Eckhardt, 2000).

Table 1: Results of rotation and scaling tests.

	Fourier	CSS	Proposed method
rotation	100%	100%	100%
scaling	86.35%	89.76%	90.40%

The parameters t_1 , t_2 and t_3 in these experiments are: $t_1 = 0.1, t_2 = 2.4, t_3 = 0.7n$. The scaling robustness test is difficult because several objects are severely distorted under reduced factors of 0.2 and 0.1. Figure 5 shows a severely distorted sample reduced by a factor of 0.1.



Original 0.1

Figure 5: Shape of a running person and its scaled-down and re-sampled version.

3.2 Performance of the Similarity-based Retrieval

The database Kimia's silhouette includes 99 shapes and is divided into 9 classes of various shapes. Each image was used as a query, and the number of images belonging to the same class was counted in the top 11 matches. Since the maximum number of correct matches for a single query image is 11, the total number of correct matches is 1089. Some of its samples are shown in Figure 6, where the shapes positioned in the same row belong to the same class.

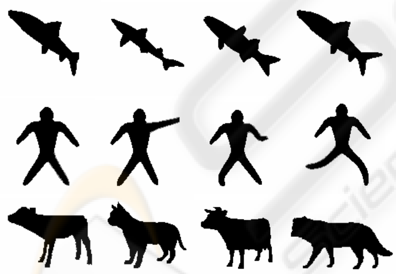


Figure 6: Example shapes in Kimia's silhouette.

Most images of the database consist of several basic shapes and their occluded and articulated shapes. A few images include similar but different animals as shown in third row of Figure 6. In Table 2, the results of this experiment are presented. The parameters t_1 , t_2 and t_3 in this experiments are: $t_1 = 0.1, t_2 = 2.9, t_3 = 0.8n$. Proposed method + SC in Table 2 means a combined method of proposed method and shape context. The recognition rate of proposed method + SC is better than that of proposed method. This is because the interior information of SC is added to the exterior information of our method.

Table 2: Results of the similarity-based retrieval.

	Recognition rate
CSS	73.19%
Shape context(SC)	76.86%
Proposed method	81.82□
Proposed method + SC	87.51□

3.3 Robustness to Occlusion

In this experiment, we took three image classes (fish, aeroplane and art object images) which consist of eight images respectively. We changed their images, and each image was impaired by 10-25% from four different directions (front, rear, right, and left). For each class, 96 images were tested, and the number of the correct matches was counted. Some of the original images and samples of occluded images are shown in Figure 7. In Table 3, the results are presented.

Figure 7: Example shapes of original and occluded images.

Table 3: Results of the occlusion tests.

	10%	20%	25%	total
CSS	91.67%	72.92%	47.92%	70.88%
SC	90.76%	66.67%	54.17%	70.49%
Proposed method	100%	96.88%	88.55%	95.14%

The parameters t_1, t_2 and t_3 in this experiments are: $t_1 = 0.1, t_2 = 2.9, t_3 = 0.7n$. Since our method uses the shape of the contour to measure the similarity, it is more suitable for recognition of partially occluded objects than CSS and SC. There exists a previous result (Krolupper and J. Flusser, 2007) deals with the recognition of the partial occlusion of objects. It also takes into account the invariance to the affine transformation. However, it could not be used to convex objects such as triangle, rectangular, since it is employed the zero-crossing points of curvature. Our method can be applied for those objects.

3.4 Remark on Computational Cost

CSS involves large computational costs due to the iterations of a Gaussian filter. The cost is at least 100 times larger than that of our method, where the Gaussian filter is used only once. For example, the calculation time by Matlab programming with Pentium(R) D 3.2GHz processor to construct the CSS image of a leaf in Figure 3 is about 150 s, while the calculation time to construct our code is about 1.4 s. The computational cost to compare the similarity of two objects by using our code (except for the complexity of computing the curvature) is low. It requires about 0.25 s. Figure 8 shows the relationship between the length of the contour of a leaf image given in Figure 2 and the calculation time of three methods (CSS, shape context, our method). It follows that the calculation time of our method is about one-hundredth lower than that of CSS, but about five times greater than that of shape context by Figure 8. The vertical line of Figure 8 is the logarithm of the calculation time and the horizontal axis is the length of contour.

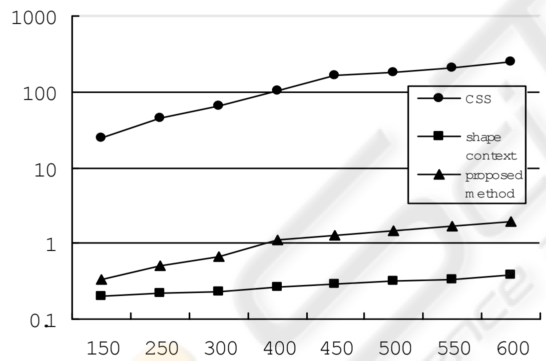


Figure 8: Graphs between the length of the contour of a leaf image and the calculation times of CSS, shape context and proposed method.

4 CONCLUSIONS

We have proposed a new method of shape matching. It is shown that the computational cost of our method is lower than that of CSS. In our method, the recognition rates of the rotation and scaling experiments are 100% and 90.40%, respectively. These results are slightly better than CSS's results. In the similarity-based retrieval and occlusion experiments, the recognition rates of our method are 81.82% and 95.14%, respectively. These results are greater than that of the CSS and SC. Fourier descriptors and shape context have smaller computational complexities than our method due to a Gaussian filter. The recognition performance of our method is better than those of Fourier descriptor and shape context in the above experiments.

ACKNOWLEDGEMENTS

This work was supported by research center for integration of advanced intelligent systems and devices of Toyota Technological Institute.

REFERENCES

F. Mokhtarian and A. K. Mackworth, 1992. A theory of multi-scale, curvature based shape representation for planar curves, IEEE Trans. *Pattern Analysis and Machine Intelligence, vol. 14, no. 8, pp. 789-805.*

F. Mokhtarian and M. Bober, 2003. *Curvature scale space Representation: Theory, Applycations, and MPEG-7 Standardization*, Computa-tional Imaging and Vision, Vol. 25, Kluwer Academic Publishers.

L. J. Latecki and R. Lakamper, 2000. Shape similarity measure based on correspondence of visual parts, *IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 10, pp. 1185-1190.*

L.J.Latecki, R.Lakamper and U.Eckhardt, 2000. Shape Descriptors for Non-rigid Shapes with a Single Closed Contour, *CVPR.*

L.J.Latecki, R.Lakamper and D. Wolter, 2005. Optimal partial shape similarity, *Image and Vision Computing 23, pp. 227-236.*

S. Belongie, J. Malik, and J. Puzicha, 2002. Shape matching and object recognition using shape context, *IEEE Trans. Pattern Analysis and Machine Intelli gence, vol. 24, pp. 509-522.*

A. Khotanzad and Y. H. Hong, 1990. Invariant image recognition by Zernike moments, *IEEE Trans. Pattern Analysis and Machine Inte-lligence, vol. 12, no. 5, pp. 489-497.*

C. T. Zahn and R. Z. Roskies, 1972. Fourier descriptors for plane closed curves, *IEEE Trans. on Computers, C-21, pp. 269-281.*

L. da F. Costa and R.M. Cesar, 2001. *Shape Analysis and Classification, Theory and Practice*, CRC press. F. Krolupper and J. Flusser, 2007. Polygonal shape description for recognition of partially occluded objects, *Pattern Recognition Letters, 28, pp.1002-1011.*