SELF-CLOSED PARTIAL SHAPE DESCRIPTOR FOR SHAPE RETRIEVAL

Huijie Fan[1,2]*, Yang Cong*[1]*, and Yandong Tang*[1]

[1] State Key Laboratory of Robotics, Shenyang Institute of Automation,
Chinese Academy of Science, Shenyang 110016, China
[2] Graduate University of the Chinese Academy of Science, Beijing 100049, China

ABSTRACT

We propose a discriminative partial-based algorithm for shape recognition and retrieval. A key distinction of our approach is that we use pairwise geometric relations between contour fragments containing important and salient shape information to establish self-closed partial descriptor (SCPD), it can capture similar local parts in matching shape contours and meanwhile overcome part occlusion and distortion. We establish local coordinate system for each fragment to make sure SCPD is invariant to RST (rotation, scaling, and translation) transformation. In the matching stage, a scale approximation scheme is used to get rid of invalid matches. We experiment on MPEG7 shape database, and experimental results illustrate that our algorithm performs well on shape retrieval.

Index Terms— partial shape descriptor, shape retrieval, MPEG7 shape database, RST transformation

1. INTRODUCTION AND RELATED WORK

Image retrieval using shape information is very important in computer vision that has received considerable attention from many researchers. However, shape representation and description is a difficult task. Since the 2D image loses one dimension of object depth information in 3D real world, and the shape extracted from the image are often corrupted with noise, defects, arbitrary distortion and occlusion.

In order to solve this problem , many promising approaches have been proposed. There are mainly two kinds of approaches on shape representation, namely, region-based approach and contour-based approach. Region-based methods consider pixels not only on the boundaries but also within the shape region. Common region-based methods use moment descriptors [8], fourier descriptor [3], shape matrix or region bags [9] to describe shapes and they have good performance on simple image retrieval. Compared with region-based methods, contour-based shape representation have attracted more attention in recent years because in many applications, the shape contour can provide sufficient information for successful object detection or retrieval and the shape interior content is not interested. Contour-based shape techniques only exploit shape boundary information and there are mainly two types of very different approaches: global shape descriptors and partial shape descriptors. The global shape descriptors [2] require the presence of the whole shape contour to be one or more closed curves, and use simple shape descriptors (such as area, circularity and eccentricity), shape signatures, chain code or curvature scale space [4] to describe shape contours. These methods are always very fast in shape matching process but they can not handle distortion, local occlusion and inter-class deformation very well, because they cannot provide local details of shape.

Recently, a number of partial shape descriptors have been proposed. Belongie et al. proposed Shape Context (SC) [1]to handle small nonlinear transformation, and it has been widely used in partial shape matching [11, 16]; Ferrari et al. proposed a scale invariant descriptor called k-Adjacent Segments (kAS) [6] to work together in a team to match the model parts, which performs well on background clutter images [5]. However, there is always a trade-off between accuracy and efficiency in partial shape matching. For example, SC is efficient but it is not rotation invariant, and kAS needs a multi-scale sliding-windows mechanism to localize previous unseen instances in test images, which is time consuming. Ma et al. [12] selected fragments with same length from two arbitrary position to define the distance matrix and the orientation matrix to obtain high accuracy, while the matching is time consuming because it is evaluated on all sample edge points.

In this paper, we proposed a new partial-based shape retrieval algorithm which is invariant to RST (rotation, scaling, and translation) transformation. Different from many machine learning based methods, we use single model image as a query in retrieval. We broke image contour into several fragments at the curvature extreme points (CEP) and connected each fragment into a closed curve using a semicircle to ensure the quantitative partial shape descriptors have the same length. The local coordinate system is established for each fragment to guarantee the partial descriptor is invariant to translation and rotation and the self-closed partial descriptor (SCPD) is constructed by dividing the local coordinate space equally into L pieces according to angle (an example is shown in Fig.1) and computing the average distance as weight of each bin for each fragment.

The remainder of this paper is organized as follows: Section 2 represents the fragment selection technique and the self-closed partial shape descriptor; Section 3 describes the partial shape matching between pairwise shape fragments. The experiment results are presented in Section 4 and final conclusions are given in Sec. 5.

2. SELF-CLOSED PARTIAL SHAPE DESCRIPTOR

The first step in our approach involves extracting image contours using the Berkeley edge detector [13] and breaking the model and testing image contours into several fragments. Contours around the curvature extreme points (CEP) contain the critical messages of object shape for human visual system, and some of them keep similar when object appears different or part occlusion. Therefore, we believe there are three CEPs in each selected fragment: two are endpoints and the other is on the fragment. i.e., for each CEP, we introduce its previous CEP and next CEP as the break points to obtain fragment, which is different from [14, 11] separating contours at two adjacent CEPs.

Let P_1, P_2 and P_3 denote the CEPs in fragment $\mathbf{s}_i$ (shown in Fig.1), where P_1 and P_3 are the two endpoints of $\mathbf{s}_i$. O_i is the midpoint of line $\overline{P_1P_3}$. To achieve the rotation and translation invariant, we first establish a local coordinate system whose x-axis is $\overline{P_1P_3}$ and the origin is O_i. The established local coordinate system for fragment $\mathbf{s}_i$ can be seen in Fig.1(represented by the black arrow). The fragment $\mathbf{s}_i$ in local coordinate system can be represented as $\mathbf{s}'_i$:

$$\mathbf{s}'_i = \begin{pmatrix} cos(\varphi_i) & -sin(\varphi_i) \\ sin(\varphi_i) & cos(\varphi_i) \end{pmatrix} (\mathbf{s}_i - O_i) \qquad (1)$$

where φ_i is the rotation angle between the local coordinate system and the world coordinate system.

Open curve matching always face the problem that shape descriptors of different fragments always have different length after quantization, we propose establishing the efficient descriptors with the same length for different partial shape fragments. Let Fig.1 as an example, $\mathbf{s}'_i = \{u_1, u_2, \cdots, u_n\}$ denotes a fragment in its own local coordinate system, we connect the two endpoints by using a semi-circle whose diameter is $\overline{P_1P_3}$ to transform the open curve into a closed curve. We divide the space equally into L pieces according to angle and compute the average distance as weight of each bin to describe the spatial distribution of fragment $\mathbf{s}'_i$, and it can be encoded into a L-dimensional self-closed partial descriptor (SCPD):

$$d_i(k) = \begin{cases} \frac{\sum_{l=1}^{L_k} |\overline{O_l u_l}|}{L_k}, & L_k > 0 \\ \frac{|\overline{P_1P_3}|}{2}, & L_k = 0 \end{cases} \qquad (2)$$

where $u_l \epsilon bin(k)$ and L_k is the number of points in $bin(k)$. The SCPD of fragment A and B in Fig.1 demonstrate that the proposed partial descriptor is invariant to rotation.

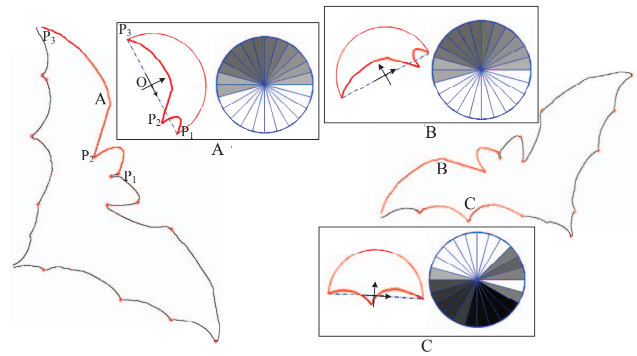

Fig. 1. An example of proposed partial descriptor. The red points are the CEPs and rectangle A shows the closed curve and the SCPD for bold fragment A in left shape contour. Rectangle B and C represent that for fragment B and C in right contour. Local coordinate space are divided equally into 24 bins and value in each bin is the average distance of origin to points in this bin. (dark = small value).

3. MATCHING BETWEEN PAIRWISE FRAGMENTS

Given an image I, a set of fragments $T = \{\mathbf{t}'_1, \mathbf{t}'_2, \cdots, \mathbf{t}'_M\}$ is generated. Our goal is to find the best matching for parts of image fragment $\{\mathbf{t}'_{j1}, \mathbf{t}'_{j2}, \cdots, \mathbf{t}'_{jn}\} \subseteq T$ and parts of model fragment $\{\mathbf{s}'_{i1}, \mathbf{s}'_{i2}, \cdots, \mathbf{s}'_{in}\} \subseteq S = \{\mathbf{s}'_1, \mathbf{s}'_2, \cdots, \mathbf{s}'_N\}$. Pairwise partial shape matching are performed between each $\mathbf{s}'_i \epsilon S$ and each $\mathbf{t}'_j \epsilon T$ to identify the similarity of two matching fragments and to select the hypothesis corresponding parts from them. Let d_i and d_j denote the SCPD of $\mathbf{s}'_i$ and $\mathbf{t}'_j$ respectively, and C_{ij} denotes the matching cost of these two fragments. The matching cost C_{ij} between two vectors is calculated using the χ^2 test statistic:

$$C_{ij} \equiv C(\mathbf{s}'_i, \mathbf{t}'_j) = \frac{1}{2} \sum_{k=1}^{L} \frac{[d_i(k) - d_j(k)]^2}{d_i(k) + d_j(k)} \qquad (3)$$

Once we get the set of costs C_{ij} between all pairs of fragments $\mathbf{s}'_i$ on query image shape and $\mathbf{t}'_j$ on the testing image shape, we use bipartite graph matching [15] algorithm to find corresponding matching fragments by minimizing the cost

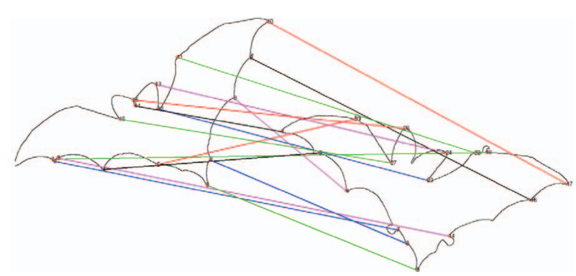

Fig. 2. Bipartite graph matching results before scale approximation scheme.

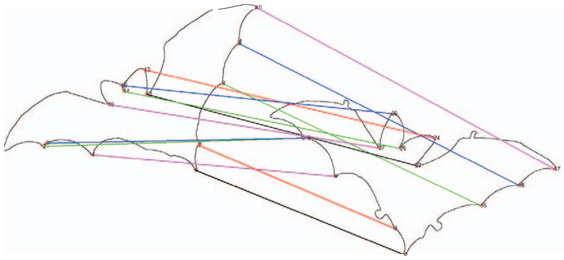

Fig. 3. Fragment matching results after scale approximation scheme.

function:

$$H_{ij} = \sum C(\mathbf{s}'_i, \mathbf{t}'_{\pi(i)}) \tag{4}$$

subject to the constraint that π is a permutation, i.e., the matching is one-to-one. Fig.2 shows the results of the bipartite graph matching. We can see that although most matching results are correct, several obvious invalid match still exist. The reason is that we don't use the global information of two matching shapes and the bipartite graph matching will assign a matching fragment for a query fragment so long as there existing more image fragments not been matched. To overcome this problem, we introduce a scale approximation scheme to remove the matchings with large scale difference and select the best five matchings for each model fragment. The local scale τ_{ij} for matching fragments $\mathbf{s}'_i$ and $\mathbf{t}'_j$ can be calculated from Formula. (2):

$$\tau_{ij} = \sum_{k=1}^{L} \left(\frac{d_i(k)}{d_j(k)}\right) \tag{5}$$

We suppose local scales wouldn't change drastically, then we calculate the scale factor after bipartite graph matching using the mean operator $\bar{\tau}$ and remove the fragment matchings whose local scales τ are outside the interval range $[\bar{\tau}/2, 2\bar{\tau}]$. Fig.3 shows the fragment matching results after getting rid of large scale variation matchings and Fig.4 represents the local scale VS matching cost distribution of matched fragments before and after scale approximation scheme. We can see the matching results are more precise, wrong matchings are eliminated or corrected.

By applying the quantization scheme to each fragment, a correspondence $T(\mathcal{E})$ between a subset of points on the model shape $\mathcal{E}_a \subseteq \mathcal{E}$ and a subset of image points $\mathcal{I}_a \subseteq \mathcal{I}$ is obtained. The matching confidence S_{ij} for model and testing images is evaluated from two aspects.

$$S_{ij} = H'_{ij} \times \left(1 - \frac{\sharp\{p|\, p\epsilon\mathcal{E}_a\}}{\sharp\{p|\, p\epsilon\mathcal{E}\}}\right) \tag{6}$$

where the first score H'_{ij} is the cost of bipartite graph matching after scale approximation scheme, which indicates how well $\mathcal{E}_a$ is corresponding to $\mathcal{I}_a$ considering the geometric arrangement. We also need to measure the proportion of the matched fragments in model shape. This is indicated by the second score. Finally we rank all obtained hypothesis according to the confidence S_{ij}.

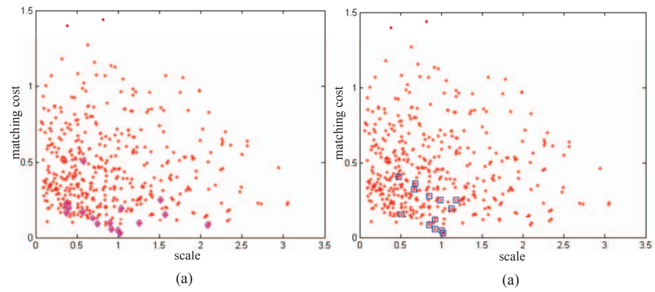

Fig. 4. The local scale and matching cost distribution of matched fragments. The red points are the candidate positions for all matching fragments. (a) Pink diamond show the positions of matched fragments before scale approximation scheme. (b) Matched fragments after scale approximation scheme are marked in blue square.

4. EXPERIMENTAL RESULT

We evaluate the proposed algorithm on the MPEG7 Part-B shape database [7] which consists of 1400 images that are classified into 70 classes, each class having 20 images. For all categories, there are scale changes, rotation and inner-class variations. This makes it very challenging for object retrieval. Fig.5 shows a representative sample for each class. The proposed method is compared with ID(inner-distance) [10], IDSC (inner-distance shape context) [10], Generative Model (GM) [17], SC(shape context) [1] and D-shape [7]. ID and IDSC are skeleton-based approached starting with two chosen landmark points calculates the shortest path between those points. GM use Thin-plate-spline (TPS) to measure the similarity of two matching shapes and D-shape use several points on circles to encode spatial arrangement of shape.

The recognition rate is measured by the so-called Bullseye test: For every image in the database, it is matched with all other images and the top 40 most similar candidates are counted. At most, 20 of the 40 candidates are correct hits. The score of the test is the ratio of the number of correct hits

Fig. 5. Samples of shapes from the MPEG7 part-B database.

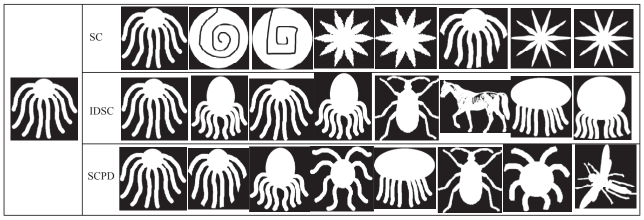

Fig. 6. An retrieval example for comparing SC, IDSC and SCPD on the MPEG7 database. The left column shows the query shape and the three right rows show the top to eight matches, from top to bottom: SC, IDSC and SCPD.

of all images to the highest possible number of hits (which is 20×1400).

In our experiment, we divide the local coordinate space evenly into 360 and each fragment is quantified to 360 dimensional vector (corresponding to 360 points). Table 1 lists the reported results from different algorithms. It shows that our algorithm outperforms all the alternatives and it can accurately extract important matching fragments. The reason is that the proposed shape descriptor SCPD is partial based, therefore it performs well on inter-class invariance where local shape could remain. Moreover, it is invariant to RST, which also makes the shape retrieval algorithm more robust to plane transformation. To help understand this performance, we did another experiment in the same setting and compared with the SC and IDCS descriptor. Fig.6) shows some detailed retrieval results for one image, where we can see that the SCPD is more effective at capturing local shape structure, i.e., it retrieve objects with similar local features.

Table 1. Retrieval rate (Bullseye) of different methods for the MPEG7 Part-B shape database.

Method	Retrieval rate
D-shape [7]	83.68%
Generative Model [17]	80.03%
SC [1]	76.51%
ID [10]	85.40%
IDSC [10]	85.4%
SCPD	**88.26**%

5. CONCLUSION

In this paper, we have presented a new shape descriptor self-closed partial descriptor (SCPD), that has very good performance in the presence of inter-class variation. This shape descriptor has three benefits. First, it is partial-based and enables our method to retriave part occlusion and distortion objects. Second, each SCPD has equal length by using a semicircle to connect fragment to be a closed curve, which is convenient the partial shape matching. Third, local scales can be calculated based on SCPD, therefore local scale approximation scheme can be used to remove invalid fragment matchings. The performance of the proposed SCPD is examined on the MPEG7 part-B shape database and compared with the state-of-the-art descriptors to justify the effectiveness of the proposed method for shape retrieval.

Acknowledgements

This work has been funded by Natural Science Foundation of China (Grant No. 60871078 and 61105013).

6. REFERENCES

[1] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *TPAMI*, pages 509–522, 2002.
[2] Y. Cao, Z. Zhang, I. Czogiel, I. Dryden, and S. Wang. 2d non-rigid partial shape matching using mcmc and contour subdivision. In *CVPR*, 2011.
[3] A. Di Lillo, G. Motta, and J. Storer. A rotation and scale invariant descriptor for shape recognition. In *ICIP*, pages 257–260, 2010.
[4] A. El-ghazal, O. Basir, and S. Belkasim. A novel curvature-based shape fourier descriptor. In *ICIP*, pages 953–956, 2008.
[5] V. Ferrari, F. Jurie, and C. Schmid. From images to shape models for object detection. *IJCV*, 87(3):284–303, 2010.
[6] V. Ferrari, T. Tuytelaars, and L. Van Gool. Object detection by contour segment networks. In *ECCV*, pages 14–28, 2006.
[7] A. Fornés, S. Escalera, J. Lladós, and E. Valveny. Symbol classification using dynamic aligned shape descriptor. In *ICIP*, pages 1957–1960, 2010.
[8] A. Foulonneau, P. Charbonnier, and F. Heitz. Multi-reference shape priors for active contours. *IJCV*, 81(1):68–81, 2009.
[9] C. Gu, J. Lim, P. Arbeláez, and J. Malik. Recognition using regions. In *CVPR*, 2009.
[10] H. Ling and D. Jacobs. Shape classification using the inner-distance. *TPAMI*, 29(2):286–299, 2007.
[11] C. Lu, L. Latecki, N. Adluru, X. Yang, and H. Ling. Shape guided contour grouping with particle filters. In *CVPR*, 2009.
[12] T. Ma and L. Latecki. From partial shape matching through local deformation to robust global shape similarity for object detection. In *CVPR*, 2011.
[13] D. Martin, C. Fowlkes, and J. Malik. Learning to detect natural image boundaries using local brightness, color, and texture cues. *TPAMI*, 26(5):530–549, 2004.
[14] S. Ravishankar, A. Jain, and A. Mittal. Multi-stage contour based detection of deformable objects. In *ECCV*, pages 483–496, 2008.
[15] K. Riesen, M. Neuhaus, and H. Bunke. Bipartite graph matching for computing the edit distance of graphs. *Graph-Based Representations in Pattern Recognition*, pages 1–12, 2007.
[16] P. Srinivasan, Q. Zhu, and J. Shi. Many-to-one contour matching for describing and discriminating object shape. In *CVPR*, 2010.
[17] Z. Tu and A. Yuille. Shape matching and recognition–using generative models and informative features. *ECCV*, pages 195–209, 2004.