

Shape Classification Using Hybrid Regional and Global Descriptor

Cong Lin and Chi-Man Pun

Abstract—In this paper we focused on designing a novel hybrid regional descriptor for shape classification, which includes three novel descriptors regional area descriptor (RAD), regional skeleton descriptor (RSD) and tangent function(TF). The RAD and RSD are based on the primitive skeleton of shape while the tangent function is developed from the contour by finding out the important landmark points and collecting regional information around them. In the matching stage, a customized Optimal Path Searching algorithm is integrated into our setting as distance measure function, which is an extension of efficient dynamic programming algorithm. The proposed shape descriptors are tested on commonly used datasets and the results are analyzed and compared to state-of-the-art methods. The experimental results show that, compared to others, our results are satisfactory.

Index Terms—Shape classification, skeleton, contour, rad, rsd, tangent function, optimal path searching.

I. INTRODUCTION

The shape of an object in an image is usually affected by various factors. Shape variations with occlusion, protrusion, noise and affine deformation on the shape are called non-rigid transforms[1], [2]. A good representation should be rotation, scale and translation (*RST*) invariant and be robust to non-rigid transform.

The shapes classification is a procedure to find out good features/descriptors that can represent the shape and a similarity measure for classification. Many methods and descriptors have been proposed and proved that work well in some extent, which can be roughly classified into three types: 1) Contour based methods. 2) Region based methods. 3) Skeleton based methods.

Contour based methods only use boundaries of the shapes and elaborates it into features vectors. Methods use contour based features are relatively efficient and easily implemented. However, these features are sensitive to noise on boundaries. Typically contour based methods are Curvature Function[1], [3], Wavelet Descriptors[4]-[6] on boundary, Hidden Markov Models [7] etc.

Region based methods use information inside the shape. The whole shape is regarded as a domain and to be transformed into another domain for descriptor extraction, e.g. Radon transform. Some method, like shape context [8] select landmark points on boundary and the descriptors are computed by summarizing region information around the landmark points.

Skeleton based methods aim to acquire topological information (medial axis) from the shape and build shape descriptors based on it. Skeleton are further developed to many methods like Shock Graph [9], Geodesic path matching [10], etc. Many methods fall into this category are required to process alignment problems [10] which resolve to optimal assignment techniques. However, occlusion and protrusion can heavily affect skeletons.

In this paper, we proposed a novel hybrid regional descriptor for shape classification, which includes three novel descriptors regional area descriptor (*RAD*), regional skeleton descriptor (*RSD*) and tangent function. In next section, the framework of the proposed approach and how our features, the *LSD*, *GSD* and tangent function, extracted from the shape are explained. We also introduce the dynamic programming classification algorithm and how the distance function generated the dissimilarity in Section III. Experiment results from the proposed approach done on widely used dataset are presented in Section IV. Analysis on the results and comparison with other approaches are also done in this section. Finally, conclusions are made in Section V.

II. EXTRACTION OF SHAPE DESCRIPTORS

Our approach is influenced by the Ling's Inner-distance [11], Belongie's Shape Context [12] and Alajlan's Triangle-Area Representation [13]. Our work shares similar framework with those previous works. We start with feature extractions on local points and collect surrounding shape information. A dynamic programming based Optimal Path Searching (*OPS*) is employed to find out correspondences of local points and measure total distance of two shapes. And finally the shape is classified into a category with minimal distance. The proposed idea differs from previous work in that we used local regions rather than local points. In our proposed method, we defined three descriptors: 1) regional skeleton descriptor (*RSD*) and 2) regional area descriptor (*RAD*) which used the skeletons and the contours and could be considered as a hybrid descriptors of the skeleton based and the contour based. And 3) tangent function, which belongs to contour based category, picks up a collection of landmarks from the contour according to some rules. Like other shape classification method, it is assumed that all shapes are well segmented and rendered as solid plate.

A. Regional Skeleton Descriptor

The skeleton is a very useful primitive shape descriptor for representing a shape. However, it is well known that matching two skeletons is a very difficult task. This is mainly because

the skeleton is the representation of the topological structure of the shape and the topological structure changes as the shape deforms. In other words, the dimension of the descriptor is not fixed and changes according to quantity of information the shape provided. And matching descriptors in different structures and different dimensions is not as simple as calculating Euclidean distance between two vectors in the same dimension.

Thus, in our proposed method, we tried to take the advantages of the skeleton which is effective in reflecting the topological structure and tried avoid the disadvantages in its matching process. And we proposed the shape descriptor called regional skeleton descriptor.

The regional skeleton descriptor (*RSD*) describes the local part of shape by computing the perimeter of the area extended from a fixed length segment of contour to the skeleton. Using the procedure below, the regional skeleton feature can be extracted from the local region:

- 1) A graph with multiple base points similar to that used in Inner-distance is built firstly. The points could be selected with a fixed interval on contour. In other words, they are equally distributed along the contour.
- 2) The skeleton of the shape is found out. In the skeleton, each point on skeleton must contain the radius of it maximum disk from the contour.
- 3) Skeleton is searched to find out the closest points on skeleton from each base point. Compute the distance between base points and their correspondent closest points on skeleton using Euclidean distance.
- 4) For each pair of consecutive base points on contour, the distance between their correspondent closest points on skeleton is calculated by finding out the path length on skeleton using depth priority search (*DPS*) algorithm.
- 5) The regional skeleton descriptor is generated by adding up distances from 3) and 4).

Geometrically, the meaning of this descriptor is how much area the part of contour between two base points is expanded into the shape. Given two base points p_1 and p_2 , and their closest points on skeleton sp_1 and sp_2 , the *RSD* of the region between these points can be mathematically defined as:

$$RSD(p_1, p_2) = D(p_1, sp_1) + D(p_2, sp_2) + \text{length}(DPS(p_1, p_2)) \quad (1)$$

where $DPS(\bullet, \bullet)$ is depth priority search function which returns a discrete path. Supposed we have N base points, the global descriptor of shape O is organized into a vector:

$$RSD(O) = \{RSD(p_1, p_2), RSD(p_2, p_3), \dots, RSD(p_{N-1}, p_N), RSD(p_N, p_1)\} \quad (2)$$

B. Regional Area Descriptor

Although the *RSD* could be used to represent a shape as whole, however, we find using *RSD* solely is not enough to justify the distance between shapes. The two local shape descriptors shown in Fig. 1 is a very clear counter-example:

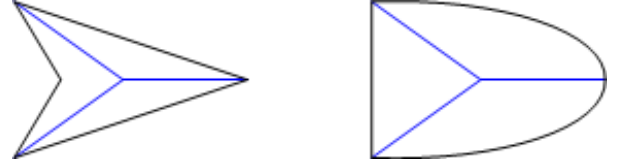


Fig. 1. Two simple different shapes have same skeleton.

Therefore, we try to consider more and turn to the areas of blocks divided by sample points and their nearest skeleton points. The regional area descriptor is generated in the same way as *RSD* that based on local part. The Regional Area Descriptor (*RAD*) adds up the maximal radii information on the skeleton segment which is corresponded to the segment of contour, which is roughly equal to the quantity of area of local part. It could be expressed in mathematical form as:

$$RAD(O) = \{RAD(p_1, p_2), RAD(p_2, p_3), \dots, RAD(p_{N-1}, p_N), RAD(p_N, p_1)\} \quad (3)$$

where

$$RAD(sp_i, sp_{i(\text{mod } N)+1})) = \sum_{j=1}^{\text{length}(sp_i, sp_{i(\text{mod } N)+1}))} \text{radius}(p_j) \quad p_j \in [sp_i, sp_{i(\text{mod } N)+1}]$$

C. Tangent Function

Contour is a very important primitive descriptor for a shape which also could be used as complementary descriptor along with the skeleton based descriptors as show in Fig. 2.



Fig. 2. Two simple shapes may have same skeleton and same size of local area.

Generally, a segmented shape contains some noise or small dents which contribute little to its shape descriptors and cause high computational cost. The noise and insignificant information on the boundary should be removed and those important points on the large turning corners should be remained. A simplification process, called discrete Curve Evolution Procedure [14], [15], is necessary in order to obtain a simpler shape. Our simplification process differs from [14], [15] in that we only used the turning angles and consider the other two shape descriptors in the whole system will compensate the loss of information in the length of line segment on simplified contour. The basic idea of simplification is: we replace two consecutive line segments on the contour with a single line segment joining their endpoints, if joint 'condition' of two consecutive line segments lower than a pre-defined threshold. The joint 'condition' is a function which is defined to compute the importance of the two consecutive line segments and their angle contributed to the whole shape. And our approach reduces every shape to a fixed number of points. To achieve this, we first define the relevance function [14], [15]:

$$K(s_1, s_2) = \frac{\beta(s_1, s_2) \cdot L(s_1) \cdot L(s_2)}{L(s_1) + L(s_2)} \quad (4)$$

The relevance function defined on two consecutive vectors and outputs a value represent how importance the triangle consist of $\overline{s_1}$ and $\overline{s_2}$ contributes to the whole shape. The higher value of $K(s_1, s_2)$, the more significant. If there is an i , such that it is minimum and number of points is still more than the pre-defined constant, the two vectors are to be replaced by one that connects the endpoints. This operation iteratively traverses the shape boundary until number of points is reduced to our pre-defined value.

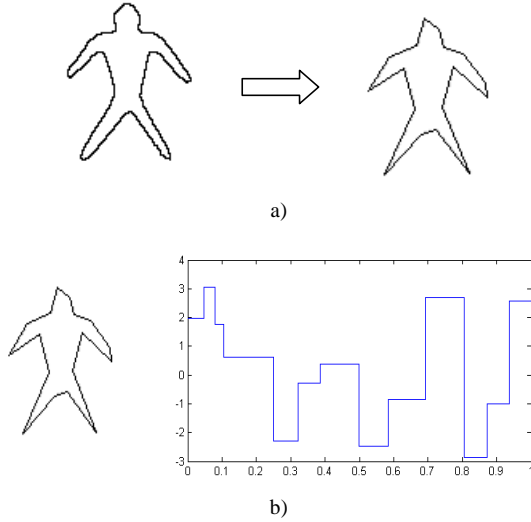


Fig. 3. a) An image boundary and corresponding simplified one. b) A simplified boundary and its tangent function.

Shown in Fig. 3, after simplification, the result generated a simplified shape which has less noise and preserves shape information as well.

The tangent function (TF) as shown in Fig. 4 is also called a turning function which is a multi-valued function defined $T(C): [0, 1] \rightarrow [-\pi, \pi]$ by $T(C)(s) = C'_-(s)$ and $T(C)(s) = C'_+(s)$, where $T(C)(s) = C'_-(s)$ and $T(C)(s) = C'_+(s)$ are left and right derivatives of C . The value of $T(p)$, p is a point on C , is angle between a reference vector and the line segment vector which play on. The tangent function $T(C)(s)$ describe where the turning takes place on the contour, increasing with left turns and decreasing with right turns, is frequently used as a shape descriptor. In order to achieve rotation invariant, we pick up the farthest pixel from centroid in the shape to start simplification.

D. Rotation, Scale, Transition Invariance

The proposed descriptors could achieve RST invariance easy. Firstly, the transition invariant could be naturally satisfied. After the contour and skeleton could be extracted, our descriptors have nothing to do with the position of the shape. As for rotation invariant, our descriptors are organized in vector form in which the local features are in clock-wised order. Therefore, the invariant property could be easily solved by a circular shift on the vector in matching. Finally, by normalizing the global vector or the vectors divided by the sum of local features, the scale invariant can be achieved.

III. CLASSIFICATION AND DISTANCE MEASURE

The Optimal Path Searching (OPS), which is extended from dynamic programming, was employed in our matching stage. There are several advantages of OPS used in pair-wised matching. The most important one among them is that dynamic programming could handle two descriptor vectors with difference dimensions. The local point on one shape is not necessary matched to a fixed point on another shape, but could be matched to the neighbors of a point on the other shape. This is quite useful in handling deformed shapes which global similarity is exist while some significant differences on local parts. The total distance of two shape is computed during OPS as cost of the optimal path

To obtain the optimal path, searching all the possible paths brutally is a conventional approach. However, the time and space complexities are prohibitive. To solve this problem, we adopt the dynamic programming algorithms based on Bellman's principle to reduce the time and computational cost. In other words, the Bellman's principle tells us that the optimal path is the concatenation of its two optimal sub-paths. Therefore, to search the optimal path, we could cast the problem into searching its two optimal sub-paths.

The proposed descriptors have different capability in classifying shapes because of different natures and different origins. A combination might be a good idea to take all use their advantages. In order to avoid complicated combination of distance measures, we use a simple linear combination of the distances from three shape descriptors.

IV. EXPERIMENTAL RESULTS

In this section, we presented our experimental results on the 99shapes dataset and the articulated dataset (as shown in Fig. 4) using the three shape descriptors respectively and their combination as well. We also compared our best results with other the-state-of-art published papers. An analysis of both advantages and disadvantages of our idea are also given in this section.

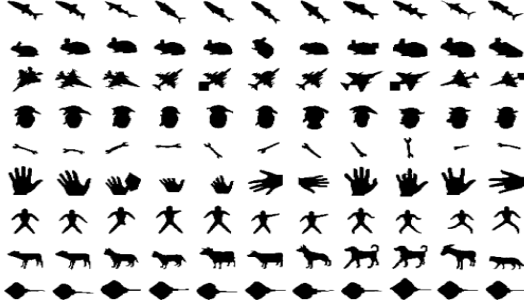
A. Classification

Our classification employed the traditional distance measure based approach which determined by the minimal distance between sample shape and shapes in the dataset. The distance is the dissimilarity of two shape and distance functions are usually defined for the feature that proposed. The distance measure function takes in two descriptors vectors from a pair of shapes respectively and calculates dissimilarity. The larger is the output distance, the less similar are the two shapes. Every shape in dataset is labeled with a known class. A sample from dataset is to be matched with every shape in the same dataset except for itself and the distances are also calculated. We find out the shape in dataset with minimal distance and consider the sample and this retrieved shape are in the same class.

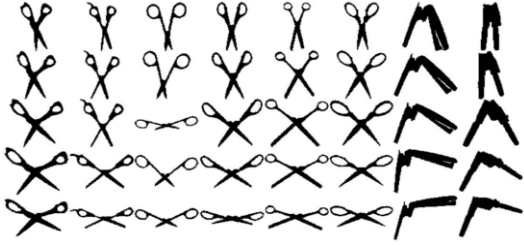
B. Datasets for Experiments

In order to evaluate which approach is better in this field, some standardized databases and benchmarks for experiment are designed. Many databases are designed for different kinds of methods, owing that different methods may have different

assumptions on shapes. A commonly used database is 99shapes, by Kimia *et al.* It contains ninety nine planar shapes which classified into nine classes, with eleven shapes in each. Shapes in the same class are in different variant form, including occluded, noised, rotated, etc. Other databases including MPEG-7 Shape Dataset [5], Articulated Dataset, Swedish Leaf Dataset and Brown Dataset are used to have further experiments. Similar to [16], Precision and Recall is used for benchmark for the reason of fair comparisons.



(a)



(b)

Fig. 4. Common used test material (a) Image dataset of 99 testing shapes provided by Kimia *et al.* (b) Image dataset of articulated testing shapes.

C. Results and Discussion

Table I shows the optimal result from test on 99shape dataset. The numbers of points we sampled from the shapes are 50, 50 and 25 for *RSD*, *RAD* and *TF* respectively.

TABLE I: RETRIEVAL RESULT OF COMBINED DESCRIPTORS ON 99 SHAPE DATASET FOR EACH CLASS

Rank	1 st	2 nd	3 rd	4 th	5 th	6 th	7 th	8 th	9 th	10 th
Quadrupeds	11	11	11	11	8	8	6	5	6	3
Humans	11	11	11	11	11	11	11	11	11	10
Airplanes	11	11	8	5	6	6	2	4	4	2
Grebes	11	11	11	11	10	11	10	8	10	3
Fish	11	11	11	11	10	11	10	7	6	5
Hands	11	11	11	11	9	8	6	8	4	1
Rays	11	11	10	11	11	11	9	7	5	3
Rabbits	11	11	11	10	10	10	11	5	8	5
Wrenches	11	11	11	11	11	11	11	11	11	11

For the articulated dataset, 45, 35 and 45 points are samples for *RSD*, *RAD* and *TF*. Retrieval result on articulated dataset was presented in Table I. We have noticed that result on 99shape from *RAD* is slightly better than *RSD*, while on articulated dataset *RSD* performs slightly better than *RAD*. During the above experiment, we tried to normalize the

descriptors and found that experiment on 99 shapes received little influence from normalization while result from articulated dataset has some improvement.

TABLE II: COMPARISON OF RESULTS OF 99 SHAPE DATASET WITH SHAPE CONTEXT [16]

Method	Shape context	Ours
1 st	97	99
2 nd	91	99
3 rd	88	95
4 th	85	93
5 th	84	85
6 th	77	87
7 th	75	74
8 th	66	68
9 th	56	63
10 th	37	45

TABLE III: COMPARISON OF RESULTS OF ARTICULATED DATASET WITH STATE-OF-THE-ART

Descriptor	1 st	2 nd	3 rd	4 th
L_2 (baseline)	25	15	12	10
SC + DP[16]	20	10	11	5
MDS + SC + DP[16]	36	26	17	15
Contour Distance[10]	36	31	28	23
Center of mass-based[10]	25	12	9	11
Our method	39	36	30	22

From the Table I, our algorithm has a almost 100% correct classified rate for Human and Wrench. We noticed that the Airplanes class is of the lowest correct rate except for the top 3 ranks. And the hit rate declined rapidly which make it singled out from the Table I. The matching distance in this class is carefully investigated and the distance revealed that our descriptors have difficulties in classifying sampled with a large part of protrusion in proportion to its main part. Another reason leads to the low correct rate in Airplanes class is the different numbers of sharp protrusions in different samples in same class and the numbers are various. Since the sharp protrusions influence the branches in skeleton, different numbers of protrusion affect the skeletons' stability within the class and diversify the descriptors of this class.

Comparison with Shape Context is listed in Table II. Compared to Shape Context our proposed shape descriptors have the advantageous property of scale and rotation invariant. Our method enjoy a high hit rate from 1st to 5th rank however, unfortunately, it slightly falls behind from 6th to 10th rank.

A comparison of the results on articulated dataset with other state-of-the-art methods shows that our method is highly articulation insensitive. The articulated dataset is a challenging dataset for shape classification owing that most samples in the dataset are segmented from real images of same type of object, for instance, the scissors and samples in the same class are from the same object but with some degree

of articulation at some branch points. Results data listed in Table III shows that the proposed method outperformed others on this dataset.

V. CONCLUSIONS

In this paper, we proposed hybrid regional and global descriptor for shape classification. By combining the descriptors, the drawbacks of each descriptor are overcome. In the matching stage, a dynamic programming based Optimal Path Searching is integrated into our method. The matching algorithm successfully handled the challenge of matching two different descriptors in different lengths. The proposed methods are tested on the widely used 99shape and articulated dataset. Experimental results are satisfactory. Data of result is presented in table form and compared to other the-state-of-art methods in literature as well.

ACKNOWLEDGEMENT

The authors would like to thank the referees for their valuable comments. This research was supported in part by Research Committee of the University of Macau and the Science and Technology Development Fund of Macau SAR (Project No. 034/2010/A2 and 008/2013/A1).

REFERENCES

- [1] L. Latecki, R. Lakamper, and U. Eckhardt, "Shape descriptors for non-rigid shapes with a single closed contour," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2000, vol. 1, pp. 424-429.
- [2] T. Schoenemann and D. Cremers, "Matching non-rigidly deformable shapes across images: A globally optimal solution," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, CVPR 2008, pp. 1-6.
- [3] P. J. V. Otterloo, *A Contour-Oriented Approach to Shape, Analysis*, Englewood Cliffs, NJ: Prentice-Hall International (UK) Ltd, 1991.
- [4] D. Huang and G. Lu, "Review of shape representation and description techniques," *Pattern Recognition*, vol. 37, 2004.
- [5] Y. W. Chen and Y. Q. Chen, "Invariant description and retrieval of planar shapes using radon composite features," *IEEE Transactions on Signal Processing*, vol. 56, pp. 4762-4771, 2008.
- [6] W. I. Yao and C. M. Pun, "Invariant shape representation by radon and wavelet transforms for complex inner shapes," in *Proc. International Conference on Information and Automation*, 2009, pp. 1144-1149.
- [7] M. Bicego and V. Murino, "Investigating hidden markov models' capabilities in 2d shape classification," *IEEE Transactions on Pattern Analysis And Machine Intelligence*, vol. 26, pp. 281-286, Feb. 2004.
- [8] Y. Zheng and D. Doermann, "Robust point matching for two-dimensional nonrigid shapes," in *Proc. Tenth IEEE International Conference on Computer Vision*, vol. 2, 2005, pp. 1561-1566.
- [9] T. B. Sebastian, P. N. Klein, and B. B. Kimia, "Recognition of shapes by editing their shock graphs," *IEEE Transactions on Pattern Analysis And Machine Intelligence*, vol. 26, pp. 550-571, May 2004.
- [10] X. Bai and L. J. Latecki, "Path similarity skeleton graph matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, July 2008.
- [11] H. Ling and D. W. Jacobs, "Shape classification using the inner-distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, pp. 286-299, 2007.
- [12] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 509-522, 2002.
- [13] N. Alajlan, I. E. Rube, M. S. Kamel, and G. Freeman, "Shape retrieval using triangle-area representation and dynamic space warping," *Pattern Recognition*, vol. 40, pp. 1911-1920, 2007.
- [14] L. J. Latecki and R. Lakämper, "Convexity rule for shape decomposition based on discrete contour evolution," *Computer Vision and Image Understanding*, vol. 73, pp. 441-454, 1999.
- [15] L. J. Latecki and R. Lakämper, "Shape similarity measure based on correspondence of visual parts," *IEEE Transactions on Pattern Analysis And Machine Intelligence*, vol. 22, pp. 1185-1120, 2000.
- [16] D. Sharvit, J. Chan, H. Tek, and B. Kimia, "Symmetry-based indexing of image databases," *J. Vis. Commun. Image Represent*, vol. 12, 2004.
- [17] O. Soderkvist., "Computer vision classification of leaves from swedish trees," Master: Linköping University, 2001.

Chi-Man Pun received his B.Sc. and M.Sc. degrees in software engineering from the University of Macau in 1995 and 1998 respectively, and Ph.D. degree in computer science and engineering from the Chinese University of Hong Kong in 2002. He is currently an associate professor at the Department of Computer and Information Science of the University of Macau. He has investigated several funded research projects and published more than ninety refereed scientific papers in international journals, books and conference proceedings. Dr. Pun has also been invited to serve as general/program and session chairs for several international conferences, reviewer and/or committee member for many international journals. His research interests include digital image processing; digital watermarking; multimedia security; pattern recognition and computer vision; intelligent systems and applications. He is also a Senior Member of the IEEE and a professional member of the ACM.