Shape Filling Rate for Silhouette Representation and Recognition

An Guocheng Zhang Fengjun Wang Hong'an Dai Guozhong
Intelligence Engineering Lab, Institute of Software Chinese Academy of Sciences, 100190, China
E-mail: anguocheng99@yahoo.com.cn

Abstract

Research on complex shape recognition showed that the shape context algorithm is sensitive to relative position variation of articulation. Aimed at this problem, a shape recognition method is proposed based on local shape filling rate of various object silhouettes. We take each landmark point as a circle center and use r as its radius. Then, under a particular radius, the ratio between the covered silhouette pixels and the total pixels is defined as local shape filling rate. Thus, different radius may form different local shape filling rates. All landmark points with different radius will constitute a characteristic matrix which can effectively reflects the entire statistical property of the object shape. Experiments on a variety of shape databases show that the novel method is insensitive to articulation and less influenced by the number of landmark points, so our algorithm has strong power in describing object details.

1. Introduction

Shape plays an important role in recognition and it is widely used in image retrieval, model registration, and object localization. Sebastian proposed a novel framework for recognition of the shape silhouettes by measure the distance between two shapes as the minimum extent of deformation [1]. But the authors also point out that the approach is not scale invariant. Bai proposed a skeleton graph matching algorithm by comparing the geodesic paths between skeleton endpoints [2]. The performance of this algorithm is limited in the presence of large protrusions and authors thought it can be solved with partial matching. Belongie proposed the shape context descriptor (SC) [3], which is represented by a set number of feature points. This algorithm makes use of the histogram of the relative polar coordinates of all other points to match different shapes. Mori used the shape context to quickly prune a search for similar shapes [4]. But the main shorting of the shape context is sensitive to articulations. The inner-distance is used to build shape descriptors in the framework of shape context by Ling, called inner-distance shape context (IDSC) [5]. This algorithm is robust to articulation and capture part structure. Yang proposed a matching of shapes which is influenced by the other shapes and they demonstrated the influence is beneficial even in the unsupervised setting[6]. But it is limited when three is noise in the database or the local geometry of the shapes is sparse.

Compared with shape context algorithm, inner distance is articulation insensitive and can reflect part structure in a certain way, but experiments showed that the performance of the algorithm is sensitive to the number of the landmark points. If take less landmark points, the reflect ability is lower, while if take more landmark points from the shape contour, the ability of the descriptor will be stronger, but in this case, the algorithm will not meet the application requirement in fasting shape classification speed as it increased the computation burden to the system. Then how to find a novel descriptor which is robust to the number of the landmark points is an important thing for practice applications. In order to solve these shortcomings of shape context and inner-distance shape context, we proposed a method which takes the shape local information as the recognition cue. The novel algorithm maintains the characteristic of articulations insensitiveness and is less influenced by the number of the shape contour landmark points and so our algorithm has good performance results in complex

Supported by the National Grand Fundamental Research 973 Program of China under Grant No.2009CB320804; the National Natural Science Foundation of China under Grant No.60673188, and No.U0735004; the National High-Tech Research and Development Plan of China under Grant No. 2009AA01Z337, No.2008AA01Z303.

1051-4651/10 $26.00 © 2010 IEEE
DOI 10.1109/ICPR.2010.129

shape classification.

2. Shape Descriptor

2.1. Classical descriptors

The shape context [3] was firstly introduced by Belongie. The main idea of shape context is to describe the coarse distribution of different shapes by the relationship of the points which distribute on the shape contours. A shape is represented by n discrete points: x_1 , x_2 , ……, x_n . Shape context histogram h_i under point x_i is computed, which describes the relative coordinates between point x_i and the remaining $n-1$ point:

$$h_i(k) = \#\{x_j : j \neq i, x_j - x_i \in bin(k)\}$$

This histogram bins are distributed in log-polar space. The angle bins n_θ is 12 and the log-spaced radius bins n_r is 5 or 8.

The shape context algorithm has good descriptive power to the object shape, but for complex shapes with articulation parts, the performance of the shape context algorithm drops in a certain degree. To clarify this problem, we give a brief introduction on shape context. In Fig.1, there are two part structures on a circle body, the solid line represents the out contour of the shape. The dashed line represents a relative displacement of the two part structures. When the part structures are changing, which we may see in Fig.1, the relative distance and angle between point P and the other points change correspondingly, thus the shape context of point P altered. Note that the distance and angle between point P and point Q differs from the distance between point P′ and point Q′ , while for shape matching, the ideal shape descriptor should have the ability of insensitive to the part structure shift.

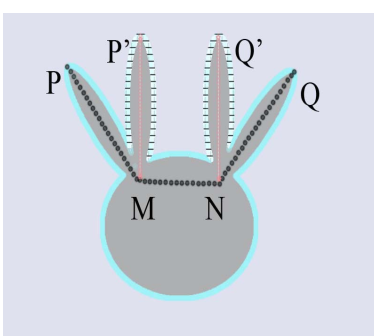

Figure 1. Shape with articulated parts

In order to solve the above problem, Ling proposed an inner distance shape context algorithm for shape matching [5]. Its main idea is to divide the object into several parts, and name each joint as a local part. For each part, Euclidean distance is still used to measure the distance of different points. The distance of different part points is computed through other joint part points. For example, the inner distance of point P and Q is computed as $d = PM + MN + NQ$. We can easily see that the relative displacement of different parts does not affect the distance of shape points as $d' = P'M + MN + NQ'$. Using this distance measure method, the shape descriptor has strong adaptiveness for the displacement of part structure of shapes. However, through experiments, we discovered that it is easy influenced by the number of the shape landmark points. When the number of the landmark points is few, the algorithm performance obviously drops.

In order to effectively describe part structure and articulation in complex shape recognition, we propose a novel algorithm, shape filling rate, which is insensitive to the number of contour landmark points. First, we take the each landmark point as a circle center and draw circles with different radius. For each circle, the ratio between the coverage portion pixels and the total pixels in the contour is computed as local shape filling rate. We may get a series of data by computing the ratios under one landmark point which we named as the feature vector to reflect the shape local characteristic. Then different landmark points will form different point feature vectors which are combined as a feature matrix to reflect the whole shape structure, say, not only articulation. In this way, we can measure the similarity of different shape feature matrixes by using the statistical property of the shape to realize image retrieval, recognition and etc. Simultaneously the proposed algorithm is less influenced by the parameter setting, that to say, number of landmark points, so the performance of our algorithm is very good.

2.2. Shape filling rate

In the process of object shape recognition, shape context and inner distance shape context have a similarity, in brief, only focus on the relative distance of different landmark points, but they neglected the local characteristic in shape structure. Therefore, we introduce a shape descriptor by using the shape local information, which we call it as shape filling rate (SFR). Our algorithm focuses on different local shape information on different object structures. Shape statistical property computed by the whole local shape information is used as a weighing standard to realize object recognition.

Taking the landmark point i as a circle centre o, we may draw some circles with different radius, shown in Fig.2. The coverage pixels which lie on the circle and belong to the shape under radius r is represented as n_i^r. The total pixel number is represented as N_i^r. A ratio is computed as:

$$fr_i(r) = \frac{n_i^r}{N_i^r}, \quad r \le R$$

where R is the biggest circle radius within the bounded shape structure. In practical application, as the distance between the object and the camera is changeable, shape structure could be different and this will bring us trouble in the setting of the parameter R. Aim to solve this problem, we usually normalize the shape before computing the feature matrix, in this way, we can easily set R within the bounded area.

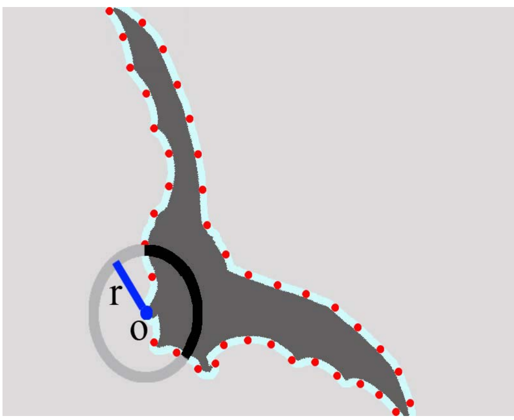

Figure 2. Computation of SFR

We also use χ^2 test statistic to measure the distance between two shapes as,

$$c(i,j) = \frac{1}{2}\sum_{r=1}^{R}\frac{[fr_{p,i}(r) - fr_{q,j}(r)]^2}{fr_{p,i}(r) + fr_{q,j}(r)} \tag{1}$$

where i and j are the point indexes of the reference shape p and the candidate shape q. $fr_{p,i}$ and $fr_{q,j}$ are the shape filling rate histograms of the reference shape p and candidate shape q.

Invariance characteristics of the shape filling rate descriptor are clearly showed in the following introduction. Firstly, we take the local shape information as the shape filling rate, in this way, we avoid influence by using coordinate information, so it has the translation invariant. Second, our algorithm has rotation invariant: for partial structure rotation, we do not use the relationship of different landmark points, so it has local rotation invariant; for the whole object rotation, we can compute the distance of the reference shape and candidate shape with different beginning landmark point, and select the minimum distance as the final distance. Third, the shape filling rate is robust to slight shape changing. We can take normalized shape to deal with shape changing in large-scale. The whole system flow is as,

Algorithm 1 SFR-based Recognition

Step1 Extract shape contour of the reference shape p, and carry on equal-space sampling;

Step2 Take landmark point i as a center of circle o. Under radius r, compute the total pixels n_i^r which are on the circle in the shape domain and compute the total pixels N_i^r, and computer $fr_{p,i}(r)$;

Step3 Take landmark point j, compute $fr_{q,j}(r)$ of the candidate shape q according to Step 1 and Step 2;

Step4 Compute the distance of the reference shape and the candidate shape according to formula (1);

Step5 Choose the minimum distance candidate shape as the result.

Only calculating the distance between the reference shape and the candidate shape to carry on the image retrieval is not enough, say, the performance is not always ideal if disturbed by the existing noise and alike. Thus we also need to obtain the corresponding relationships between the landmark points of the reference and candidate shapes to strengthen the understanding of the image. To realize it, we need a point matching algorithm. Belongie used a bi-partite graph matching algorithm for point matching [3] and Ling used dynamic programming (DP) algorithm [5]. We also use DP for point matching.

3. Experiments

First, we test the shape filling rate algorithm with the articulated shape data set of Ling [5]. There are 40 shapes from eight objects, each of which has 5 shapes articulated to different degrees. We compare the performance of SC, IDSC, and the SFR in the following experiment. The number of landmark points is n, which is set as 200 for SC and IDSC and the bin number is $n_\theta = 12$, $n_r = 5$. For the SFR, parameters are set as $n = 50$ and $R = 19$. We also force the contours to start from the bottom left points and set $k = 1$ for DP matching as Ling [5]. The retrieval result is summarized as the number of correct object shown in Table 1.

Table 1. Retrieval result on articulate data set

Algorithm	1th	2th	3th	4th
SC+DP	20/40	10/40	11/40	5/40
IDSC+DP	40/40	34/40	35/40	27/40
SFR+DP	**40/40**	**36/40**	**34/40**	**30/40**

Here we give a short comparison on the performance under different number of landmark points and we may see which algorithm is good and less influenced by the number of the landmark points. For the performance comparison, we use sum of the first, second, third, and fourth most similar matches coming from the correct object. In Fig.3, y-axis represents the total number of correct matches and x-axis represents the number of landmark points, and the curves represent the results. We can conclude that the SFR algorithm under less landmark points is robust in recognition compared with other two algorithms.

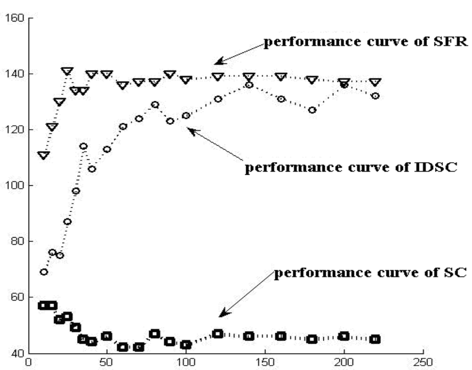

Figure 3. Performance curves of SC, IDSC, and SFR

We use two shape databases of Kimia [1] to test our algorithm. The parameters are set as $n = 100$, $n_\theta = 12$, $n_d = 5$ for SC and IDSC and $n = 50$, $R = 49$ for SFR. Without using matching algorithm, test on each descriptor clearly show their performance in Table 2.

Table 2. Retrieval result on Kimia data set 1

Algorithm	1th	2th	3th
Belongie[3]	25/25	24/25	22/25
IDSC+DP	25/25	24/25	25/25
SFR+DP	**25/25**	**25/25**	**24/25**
SC	23/25	19/25	9/25
IDSC	22/25	17/25	9/25
SFR	**25/25**	**24/25**	**22/25**

The proposed algorithm has strong power to descript the shape local information, yet there is one exception, that if the shape database is big, the relation of the landmark points will play an important role, thus we propose a combination of SC and SFR to enhance the recognition ability. The total distance of the reference shape and the candidate shape is computed as,

$$Dis = \alpha Dis_{SC} + \beta Dis_{SFR}$$

where α, β are adjustment parameters which are set as $\alpha = 3$, $\beta = 1$. These algorithms are tested on the Kimia data set 2, and without using matching algorithm either. The landmark points are set as 100 and 25. The other parameters are set as $n_\theta = 12$, $n_d = 8$ for SC and IDSC and $R = 49$ for SFR. The results are shown in Table 3.

Table 3. Retrieval result on Kimia data set 2

Algorithm	1th	2th	3th	4th	5th	6th	7th	8th	9th	10th	Total
IDSC 100	99	97	97	97	97	96	93	89	84	72	921
SC 100	99	97	97	97	96	96	94	90	87	82	935
SFR+SC 100	99	97	97	97	96	95	97	92	89	82	939
IDSC 25	95	92	91	82	84	84	71	69	55	58	781
SC 25	99	96	95	94	86	89	83	78	80	60	860
SFR+SC 25	99	96	97	93	92	94	84	83	84	65	887

4. Conclusion

In this paper, we have presented a novel object shape recognition algorithm based on shape filling rate. The novel algorithm is easy to apply and the results demonstrate that the proposed algorithm is robust towards variations in landmark points and articulation.

References

[1] Sebastian T B, Klein P N, Kimia B B. Recognition of shapes by editing their shock graphs [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2004, 26(5): 550-571.

[2] Bai X, Latecki L J. Path similarity skeleton graph matching [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2008, 30(7): 1282-1292.

[3] Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts [J]. IEEE Trans. on Pattern Analysis and machine Intelligence, 2002, 24(4): 509-522.

[4] Mori G, Belongie S, Malik J. Efficient shape matching using shape contexts [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2005, 27(11): 1832-1837.

[5] Ling H B, Jacobs D W. Shape classification using the inner-distance [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2007, 29(2): 286-299.

[6] Yang X W, Tezel S K, Latecki L J. Locally constrained diffusion process on locally densified distance spaces with applications to shape retrieval[C]. IEEE Conference on Computer Vision and Pattern Recognition, Miami, 2009: 357-364.