2009 10th International Conference on Document Analysis and Recognition

A Pixel-level Statistical Structural Descriptor for Shape Measure and Recognition*

Jing Zhang[1], Liu Wenyin[2]
[1]Department of Computer Science, University of South Florida, Tampa, FL, USA
[2]Department of Computer Science, City University of Hong Kong, Hong Kong SAR, PR China
jzhang2@cse.usf.edu, csliuwy@cityu.edu.hk

Abstract

A novel shape descriptor based on the histogram matrix of pixel-level structural features is presented. First, length ratios and angles between the centroid and contour points of a shape are calculated as two structural attributes. Then, the attributes are combined to construct a new histogram matrix in the feature space statistically. The proposed shape descriptor can measure circularity, smoothness, and symmetry of shapes, and be used to recognize shapes. Experimental results demonstrate the effectiveness of our method.

1. Introduction

Shape description and recognition play an important role in the computer vision research field. Along with the extensive studies on the geometrical, topological, and statistical features of shapes, many shape descriptors have been proposed in recent years. Belongie *et al.* [1] used Shape Context to describe the relative spatial distribution of land mark points around feature points. Ling *et al.* [2] proposed an extension of Shape Context. Inner distance, the length of the shortest path between landmark points within the shape silhouette, computed by dynamic programming technique is used for shape description. By using Discrete Curve Evolution pruning, Bai *et al.* [3] utilized the similarity of the shortest paths between each pair of skeleton endpoints to match shapes. Based on the appearance of shapes, Rosin [4] presented ellipticity, rectangularity, and triangularity shape descriptors for shape analysis and discrimination.

A symbol descriptor based on Statistical Integration of Histogram Array (*SIHA*) was proposed by Yang [5]. The skeleton of a symbol is extracted and sampled into N points: $P_0, P_1, \ldots, P_N$. For each triplet containing P_i, P_j, and P_k, angle $\angle P_jP_iP_k$ and length ratio $min(|P_iP_j|/|P_iP_k|,\ |P_iP_k|/|P_iP_j|)$ are the two structural features extracted to indicate the relations between P_i and the other two points. All triplets are used to construct Length-ratio Histogram (*LH*) and Angle Histogram (*AH*). Furthermore, the information contained in the histograms is statistically integrated to build a feature array with a fixed dimension.

Based on the structural features of *SIHA*, we propose a new shape descriptor, Structural Feature Histogram Matrix (*SFHM*). Compared with *SIHA*, *SFHM* has the following improvements: (1) Length ratios and angles are integrated in one histogram matrix so that more structural and statistical information is extracted from shapes. (2) By taking centroid as a fixed reference point to calculate length ratios and angles, *SFHM* can measure circularity, smoothness, and symmetry of shapes. (3) The computational complexity is reduced from $O(N^3)$ to $O(N^2)$.

The rest of this paper is organized as follows. Section 2 presents how to calculate *SFHM*. The properties of the descriptor are explored in Section 3. Section 4 shows experimental results on three data sets and Section 5 shows our conclusions.

2. Descriptor Computation

The computation of the proposed descriptor can be briefly summarized into 3 steps: (1) The centroid of a shape is computed based on the Distance Transform. (2) Two structural attributes, the length ratios and angles of each point on the contour, are calculated by taking the centroid as a fixed reference point. (3) Statistics are conducted on the two attributes to generate the Structural Feature Histogram Matrix (*SFHM*).

2.1. Distance Transform based Centroid

By exploring the properties of Distance Transform (DT), we present a DT based centroid. Compared with the centroid obtained by averaging all points in a shape,

* The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China [Project No. CityU 117907].

978-0-7695-3725-2/09 $25.00 © 2009 IEEE
DOI 10.1109/ICDAR.2009.175

DT based centroid is closer to the centroid given by the human visual system and more robust to noise.

Assume $I(x,y)$ is a point in a shape, $D(x,y)$ is the corresponding point after DT, whose value $|D(x,y)|$ indicates the distance to the closest boundary point from $I(x,y)$. The DT based centroid is computed by:

$$C_{DT}(x,y)=\frac{\sum_k [(x,y)\times |D(x,y)|]}{\sum_k |D(x,y)|},$$

where k is the number of points in the shape.

We can see that C_{DT} is computed by weighted average. The points from the main portions of the shape have bigger weights, while the points or noise from minor portions have smaller weights. Fig 1 shows two examples. 'o' is the centroid obtained by averaging all shape points, 'x' is the DT based centroid.

(a) (b)

Fig. 1 DT based centroid

Generally, human vision takes the disc in Fig. 1-a as the main portion and supposes the centroid to be near the center of the disc. By averaging all points, however, the centroid ('o') is far from the disc center due to the existence of the curves. While the DT based centroid ('x') is much closer to human visual result. Fig 1-b shows another example from Kimia's database [6].

2.2. Structural Feature Histogram Matrix

Based on the obtained centroid, we start to compute the proposed shape descriptor. Assume that $C(x,y)$ is the centroid of a shape, P_i and P_j are two different points on the contour $\{P_0,P_1,\ldots,P_{K-1}\}$, $|P_iC|$ and $|P_jC|$ denote the lengths of the two vectors P_iC and P_jC, θ_{ij} is the angle between them, as shown in Fig. 2, two structural attributes d_{ij} and θ_{ij} are defined as follows:

$$d_{ij}=\min(|P_iC|/|P_jC|,|P_jC|/|P_iC|),\ d_{ij}\in[0,1], \quad (1.a)$$

$$\theta_{ij}=\angle P_iCP_j,\quad \theta_{ij}\in[0,\pi]. \quad (1.b)$$

Clearly, they are invariant to scaling and rotation.

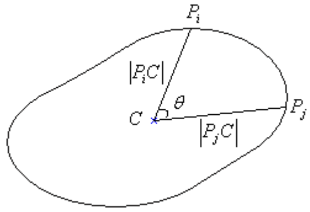


Fig. 2. Two Structural Attributes of *SFHM*

Note that any pair of points (P_i, P_j) out of the K points on the contour has a corresponding pair (d_{ij},θ_{ij}). Hence, there are $K(K-1)/2$ pairs of (d_{ij},θ_{ij}), $i\in[0,K-2]$, $j\in[i+1,K-1]$ for a shape. The set $(D,\Lambda)=\{(d_{ij},\theta_{ij})\mid i\in[0,K-2];j\in[i+1,K-1]\}$ are utilized to describe the shape. First, we transform the shape into a feature space with a new coordinate system, in which θ denotes the X axis, d denotes the Y axis, and every element in the set (D,Λ) is a point in the feature space. Second, since $d_{ij}\in[0,1]$ and $\theta_{ij}\in[0,\pi]$, we divide [0,1] into M equal bins, $[0,\pi]$ into N equal bins. Consequently, the feature space is divided into $M\times N$ blocks shown in Fig. 3.

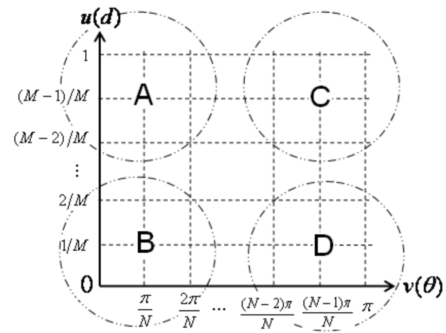


Fig. 3 The coordinate system of the feature space

Third, we calculate the percentage of the contour points in the block at the u^{th} row and the v^{th} column by the formula:

$$q(u,v)=\frac{2}{K(K-1)}\sum_{i=0}^{K-2}\sum_{j=i+1}^{K-1}h(d_{ij},\theta_{ij},x_{u-1},x_u,y_{v-1},y_v),$$

$$u\in[0,M-1],v\in[0,N-1], \quad (2)$$

where,

$$h(d_{ij},\theta_{ij},x_{u-1},x_u,y_{v-1},y_v)=\begin{cases}1 & d_{ij}\in[x_{u-1},x_u),\theta_{ij}\in[y_{v-1},y_v)\\ 0 & Else\end{cases}$$

Hence, there is

$$\sum_{u=0}^{M-1}\sum_{v=0}^{N-1}q(u,v)=1.$$

Finally, a $M\times N$ histogram matrix Q is constructed based on all $q(u,v)$, $u\in[0,M-1]$, $v\in[0, N-1]$.

$$Q=\begin{bmatrix} q(0,0) & q(0,1) & \cdots & q(0,N-1)\\ q(1,0) & q(1,1) & \cdots & q(1,N-1)\\ \vdots & \vdots & & \vdots\\ q(M-1,0) & q(M-1,1) & \cdots & q(M-1,N-1)\end{bmatrix}$$

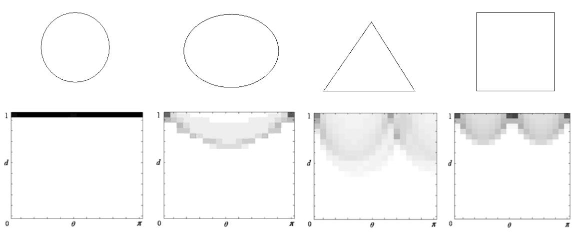
From the definitions in (1) and (2), we can see that Q is a Histogram Matrix in the Structural Feature space. Hence, Q is referred to as *SFHM*, which is the proposed descriptor for shape measure and recognition.

3. Descriptor Properties and Applications

In this section, we investigate the properties of the proposed shape descriptor, including periodicity for regular polygons, circularity, smoothness, symmetry, and its application in shape recognition.

3.1. Periodicity for Regular Polygons

Structural feature histogram matrices Q (20×20) of circle, ellipse, regular triangle, and square are illustrated as intensity images in Fig. 4. The gray level of each block corresponds to its element value. Dark means a large value and bright means a small value. It is easy to notice that the profiles of nonzero elements in their *SFHM* are periodical, with 1, 1.5 and 2 periods for ellipse, regular triangle, and square, respectively. Let us take the square for instance to explain this scenario. When $\theta \in \{\pi/4, 3\pi/4\}$, the length ratios between any two points are the smallest so that peaks occur. When $\theta \in \{0, \pi/2, \pi\}$, the length ratios are biggest (nearly 1) so that valleys occur. Hence the profile of nonzero elements of regular rectangle has 2 periods. This scenario also exists for other regular polygons, e.g., 2.5 periods for pentagon, 3 periods for hexagon, and so on.



(a) (b) (c) (d)

Fig. 4. Four symbols and their corresponding structural feature histogram matrices ((a) Circle, (b) Ellipse, (c) Regular Triangle, (d) Square).

3.2. Circularity

According to (1.a), we know that if the Structural Feature Histogram Matrix Q of a shape has more d_{ij} whose values are close to 1, then more points on the contour have similar distances to the centroid. In other words, this shape is closer to a circle. Based on this property, we can compute the circularity (C_{ir}) of a shape.

Assume the size of Q of a shape is $m \times n$ with the distance $d=1$ at 1st row and $d=0$ at m^{th} row. (1) Assign the elements in i^{th} row of Q a weight i so that a bigger d has a smaller weight and a smaller d has a bigger weight. (2) Let C_{ir} equal to the weighted summation of all elements in Q. The pseudo code is:

```
Cir =0;
FOR i=1:m
    Cir=Cir+(summation of all element on ith row)*i;
END
```

Because all elements in the *SFHM* of circles are at the first row where $d=1$ (Fig. 4-a) and the summation of all elements is 1. Hence, circle has smallest circularity, $C_{ir}=1$. A shape with C_{ir} value closer to 1 is more similar to a circle.

The circularity values for four shapes in Fig. 4 are listed in TABLE 1.

TABLE 1. Circularity of shapes in Fig. 4

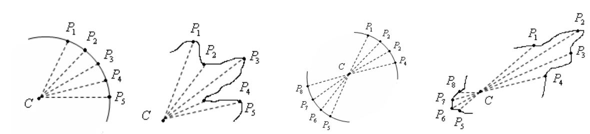
Shape	Fig. 4-a	Fig. 4-b	Fig. 4-c	Fig. 4-d
Circularity	1	11.505	21.748	12.664

3.3. Smoothness and Symmetry

Moreover, the proposed shape descriptor can be used to measure the smoothness and symmetry of shapes. Assume the top-left, bottom-left, top-right, and bottom-right corners of Q are region A, B, C and D, respectively, as shown by the dashed circles in Fig. 3. V_A, V_B, V_C, and V_D are the sets of point pairs (P_i, P_j) corresponding to the elements (d_{ij}, θ_{ij}) in the four regions. For (P_i, P_j) in V_A, vectors $|P_iC|$ and $|P_jC|$ have similar lengths and small angles between them. Therefore, V_A indicates smoothness of the contour. The larger the element values in region A, the smoother the shape. For vector pairs in V_c, they have similar lengths as well but angles between them are much bigger, nearly π. Therefore, they indicate symmetry of the contour. The larger the element values in region C, the more symmetric the shape. Similarly, region B and region D indicate roughness and asymmetry of the shape. The relations are summarized in TABLE 2.

TABLE 2. Relations between Structural Feature Histogram Matrix and Shape

	Ratio	Angle	Properties	Example
Reg. A	large	small	smoothness	Fig. 5-a
Reg. B	small	small	roughness	Fig. 5-b
Reg. C	large	large	symmetry	Fig. 5-c
Reg. D	small	large	asymmetry	Fig. 5-d



(a) (b) (c) (d)

Fig. 5. Four examples illustrating the properties of region A, B, C and D in Fig. 3

The element values in region A and C are utilized to measure smoothness (S_m) and symmetry (S_y). A smooth and symmetry shape has big S_m and S_y values. Meanwhile, we can know the smoothness and symmetry of a shape from the intensity image of its *SFHM* directly. For example, the top-left and top-right corners of *SFHM* (20×20) in Fig. 4-c have brighter color than others, which means the two corners have smaller values. Hence the triangle is the roughest and most asymmetrical in the four shapes. TABLE 3 shows smoothness and symmetry of the shapes in Fig. 4 when let $S_m = q(1,1)$ and $S_y = q(1,20)$.

TABLE 3. Smoothness and Symmetry of The Four Shapes in Fig. 4

Shape	Fig. 4-a	Fig. 4-b	Fig. 4-c	Fig. 4-d
Smoothness	0.05	0.0492	*0.0275*	0.0361
Symmetry	0.05	0.0468	*0.0030*	0.0171

3.4. Recognition

We use the distance between two feature histogram matrices to recognize shapes. Let Q_i and Q_j be two matrices, the distance between them is defined as:

$$Dist = \sum_{x,y} abs(q_i(x,y) - q_j(x,y)) \cdot$$

The smaller the *Dist*, the more similar the two shapes.

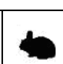
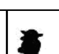
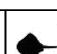
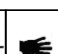
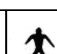
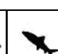
4. Experimental Results

In this section, we will show and analyze the experimental results based on the above properties of the proposed descriptor. In our experiments, Kimia's 99 shape database [6], Asian and Tari's 56 shape database [7], and GREC 2003 symbol database [8] were used to test our descriptor and the size of structural feature histogram matrix was set to 100×100.

4.1. Circularity

To evaluate *SFHM*'s performance on measuring circularity of shapes, we used Kimia's database with 99 images from nine classes. By using the algorithm described in Section 3.2, we computed the average circularity values C_{ir} of each class, shown in TABLE 4.

TABLE 4. Averaged Circularities of Nine Classes

Class									
C_{ir}	21.4	22.0	25.5	26.2	29.1	29.6	35.4	38.7	42.8

We can see that the results are reasonable and consistent with the judgment of human beings. The wrenches and sharks in narrow shapes, and the objects with protrusion, such as leg, limb, and finger, are more dissimilar to circle than other objects.

4.2. Smoothness and Symmetry

In our experiments, the size of *SFHM* was 100×100 and the smoothness (S_m) was calculated by using $S_m = q(1,1) + 0.9 \times q(1,2) + 0.9 \times q(2,1) + 0.8 \times q(2,2)$.

In Fig 6, the nine classes from Kimia database are shown in descending order based on their smoothness values from top to bottom and the 11 shapes of each class are shown in descending order based on their smoothness values from left to right. We can see that the results of our method are compatible with human visual system. "Bunny" is the smoothest class and "Person" is the roughest class. The persons with fewer limbs (9th row) and hands with fewer fingers (6th row) are smoother than other shapes in the class. The double open-end wrenches (4th row) and the two dogs with big tails (7th row) are less smooth than other shapes in the class. The averaged smoothness values of the nine classes are shown in TABLE 5.

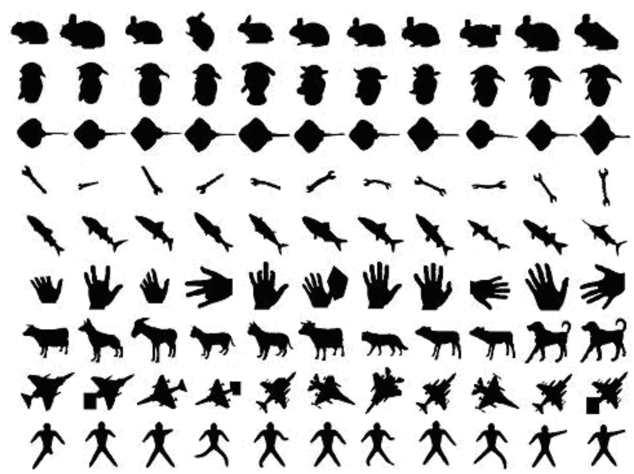


Fig. 6 Results of Smoothness on Kimia's Database

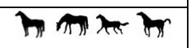
TABLE 5. Averaged Smoothness of Classes in Fig. 6

Row	1st	2nd	3rd	4th	5th	6th	7th	8th	9th
S_m ($\times 10^{-3}$)	4.3	4.8	5.2	5.3	7.3	7.7	9.4	9.5	9.8

Asian and Tari's 56 shape database which contains more symmetrical shapes were used to evaluate the performance of our descriptor on measuring symmetry of shapes. Note the measured symmetry is centroid based. The symmetry (S_y) was calculated by $S_y = q(1,100) + 0.9 \times q(1,99) + 0.9 \times q(2,100) + 0.8 \times q(2,99))$.

The experimental results are shown in TABLE 6 with their averaged symmetry values in descending order. Clearly, the first two classes with biggest S_y values are much more symmetrical than other classes.

TABLE 6. Results of Symmetry on Asian and Tari's Database

S_y $(\times 10^{-4})$	Class	S_y $(\times 10^{-4})$	Class
2.060		1.397	
1.110		1.048	
0.892		0.837	
0.717		0.700	
0.695		0.648	
0.599		0.580	
0.574		0.418	

4.3. Recognition

TABLE 7 shows the recognition accuracies of *SFHM* on Kimia 99 shape database. The first column indicates the number of top 1 to top 10 closest matches (99 is the best result for each of them). Its overall recognition rates are higher than Shape Contexts [1] but lower than other methods which use much more complex algorithms. E.g., Path Similarity [3] uses discrete curve evolution and optimal subsequence bijection to extract and match skeleton; Inner Distance [2] uses dynamic programming technique. Compared with them, *SFHM* is easier to implement by computing only length ratios and angles among points on the contour. Hence, its recognition rate is acceptable if it slightly drops.

TABLE 7. Recognition Rate of Kimia's Database

Alg.	Shape Context [1]	**Our Method**	Gen. Model [9]	Path Similarity [3]	IDSC +DP [2]
1st	97	**96**	99	99	99
2nd	91	**96**	97	99	99
3rd	88	**91**	99	99	99
4th	85	**89**	98	99	98
5th	84	**84**	96	96	98
6th	77	**84**	96	97	97
7th	75	**75**	94	95	97
8th	66	**77**	83	93	98
9th	56	**56**	75	89	94
10th	37	**55**	48	73	79

Moreover, in contrast to other methods in TABLE 7, *SFHM* can be used to recognize symbols by using symbol skeletons to replace shape contours. Meanwhile, because the skeletons usually contain more points than contours, our descriptor can achieve better performance. TABLE 8 shows the number of misclassification over 3150 samples in the Rotation, Scaling and Combination dataset and the Deformation and Degradation dataset of GREC2003. Our shape descriptor outperformed other two methods.

TABLE 8. Recognition performance on GREC2003

Algorithm	**Our Method**	Kernel Density [10]	SIHA [5]
Number of misclassification	**11**	14	39

5. Conclusions

We proposed a pixel-level statistical structural descriptor for shape measure and recognition in this paper. By integrating the length ratios and angles among all points on a shape contour into a Structural Feature Histogram Matrix (*SFHM*), the proposed descriptor extracts structural and statistical information from shapes effectively. Based on the properties of our descriptor, it can be used to measure circularity, smoothness, and symmetry of shapes easily. In addition, the distance between two *SFHM* can be utilized for shape recognition. The experimental results on three data sets demonstrate the validity of our method.

References:

[1] S. Belongie, J. Puzhicha, and J. Malik, "Shape Matching and Object Recognition Using Shape Contexts," *IEEE Trans. Pattern Analysis and Machine Intelligence,* 24(4):509-522, Apr. 2002.

[2] H. Ling and D.W. Jacobs, "Shape Classification Using Inner-Distance", *IEEE Trans. Pattern Analysis and Machine Intelligence,*, 29(2):286-299, Feb. 2007.

[3] X. Bai and L.J. Latecki, "Path Similarity Skeleton Graph Matching", *IEEE Trans. Pattern Analysis and Machine Intelligence*, 30(7):1-11, Jul., 2008.

[4] P.L. Rosin, "Measuring Shape: Ellipticity, Rectangularity, and Triangularity", Machine Vision and Applications, 24: 172-184, 2003

[5] S. Yang, "Symbol Recognition via Statistical Integration of Pixel-level Constraint Histograms: A New Descriptor", *IEEE Trans. Pattern Analysis and Machine Intelligence*, 27(2):278-281, 2005.

[6] T.B. Sebastian, P.N. Klein, and B.B. Kimia, "Recognition of Shapes by Editing Their Shock Graphs", *IEEE Trans. Pattern Analysis and Machine Intelligence*, 26(5):550-571, May 2004.

[7] C. Aslan and S. Tari, "An Axis Based Representation for Recognition", *Proc. Int'l Conf. Computer Vision*, 1339-1346, 2005.

[8] http://www.cvc.uab.es/grec2003/SymRecContest/index.

[9] Z. Tu and A. Yuille, "Shape Matching and Recognition: Using Generative Models and Informative Features", *Proc. European Conf. Computer Vision*, 3:195-209, 2004.

[10] W. Zhang, L. Wenyin, and K. Zhang, "Symbol Recognition with Kernel Density Matching", *IEEE Trans. Pattern Analysis and Machine Intelligence,* 28(12):2020-2024, Dec. 2006.