Image and Vision Computing 29 (2011) 286–294

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

A novel contour descriptor for 2D shape matching and its application to image retrieval

Xin Shu [a,b,c], Xiao-Jun Wu [a,b,*]

[a] Department of Computer Science and Technology, Jiangnan University, Wuxi 214122, China
[b] Key Laboratory of Advanced Process Control for Light Industry(Ministry of Education), Jiangnan University, Wuxi 214122, China
[c] School of Computer Science & Engineering, Jiangsu University of Science & Technology, Zhenjiang 212003, China

ARTICLE INFO

Article history:
Received 2 December 2009
Received in revised form 27 September 2010
Accepted 15 November 2010
Available online 29 November 2010

Keywords:
Shape matching
Shape retrieval
Contour points distribution histogram (CPDH)
Earth mover's distance (EMD)

ABSTRACT

We suggest a novel shape contour descriptor for shape matching and retrieval. The new descriptor is called contour points distribution histogram (CPDH) which is based on the distribution of points on object contour under polar coordinates. CPDH not only conforms to the human visual perception but also the computational complexity of it is low. Invariant to scale and translation are the intrinsic properties of CPDH and the problem of the invariant to rotation can be partially resolved in the matching process. After the CPDHs of images are generated, the similarity value of the images is obtained by EMD (Earth Mover's Distance) metric. In order to make the EMD method used effectively for the matching of CPDHs, we also develop a new approach to the ground distance used in the EMD metric under polar coordinates. Experimental results of image retrieval demonstrate that the novel descriptor has a strong capability in handling a variety of shapes.

1. Introduction

The ever growing number of images generated everyday by electronic devices has motivated researchers to develop sophisticated algorithms for the retrieval of images from large databases based on their content rather than their textual annotations alone. Among other generic image features that are used to achieve this objective, like color and texture, shape is considered a very important visual feature in object recognition and retrieval [1,2]. Using the shape of an object for object recognition and image retrieval is an important topic of computer vision and multimedia processing. Finding good shape descriptors and similarity measures are the central issues in these applications [3]. However, shape representation and description is still a difficult task. According to Kim [4], a good shape representation should be compact and retains the essential characteristics of the shape. Meanwhile, invariant to rotation, scale, and translation are also required since such invariance is consistent with the human vision system. Besides, a good method should deal with the challenges like noise, distortion and occlusion since they change a shape in a more complex way.

Many shape representation and analysis methods have been proposed during the past decades [5]. The existing shape representation and description techniques can be generally classified into two categories: contour-based methods and region-based methods. The contour-based methods, such as the curvature scale space [6], which is based on the computation of a similarity measure on the best possible correspondence between maximal convex/concave arcs are contained in simplified versions of the boundary contours. Fourier descriptors [7,8], which are easy to be implemented are based on the well-developed theory of Fourier analysis. Contour flexibility [3], which represents the deformable potential at each point along a contour, and shape context [9], is a descriptor developed for finding correspondences between point sets. Contour-based methods also include robust symbolic representation [10], distance sets [11], and elastic matching [12], etc. These methods extract information from the boundary of a shape only and ignore the rich information contained in the shape region. On the other hand, the region-based methods consider the global information of all the pixels within a shape. Some methods in this category are termed moment analysis [13,14], generic Fourier descriptor [15] etc. There are also hybrid methods that combine both contour-based and region-based methods together, such as rolling penetrate descriptor [16].

We are interested in the first one, i.e. contour-based shape techniques, because an increasing attention has been paid to this topic. But most of approaches in this field need to find the correspondence between contour points from two shapes respectively. This work suffers from not only time-consuming but also missing correspondences, in most cases, because the correspondences are obtained only by local (part) contour information.

In this paper, we propose a new descriptor for closed curves, called contour points distribution histogram (CPDH), which belongs to the contour-based methods and depicts the deformable potential at each point along a curve. Because of the peculiarity of the CPDHs, a novel ground distance calculation technique is developed, in the EMD scheme, for shape matching.

* Corresponding author. School of Information Engineering, Jiangnan University, Wuxi 214122, China. Tel.: +86 510 85912139.
E-mail addresses: ecsishu@yahoo.com.cn (X. Shu), wu_xiaojun@yahoo.com.cn (X.-J. Wu).

0262-8856/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2010.11.001

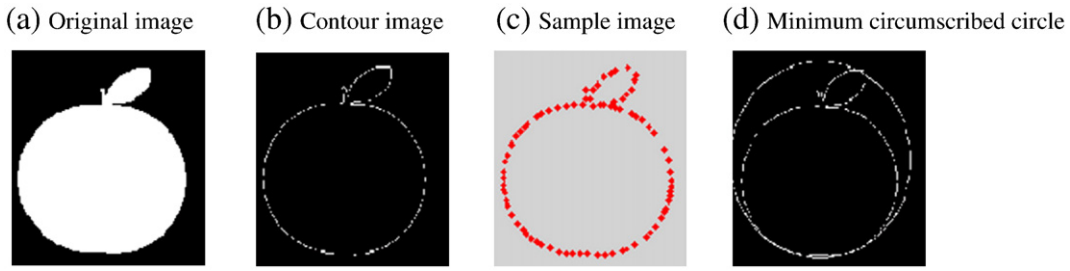


Fig. 1. Procedures and results of contour points extracted and sampled.

The rest of the paper is organized as follows. In Section 2, the proposed novel shape contour descriptor is discussed in detail. The similarity between CPDHs measured by the EMD approach is discussed in Section 3 and Section 4 includes computational complexity of the proposed approach. The experimental results are presented in Section 5. Finally, Section 6 concludes the paper.

2. CPDH descriptor

In this section we describe how to extract the contour point's distribution histogram which is invariant with respect to the scale of the shape. Analyzing the contour of an object is the first and most important step in shape matching. Based on the contour of the object, a variety of shape descriptors and matching methods have been proposed in the literature. The most similar works to ours can be found in Ref. [9]. In the following we will outline the proposed contour descriptor CPDH. Generating the CPDH includes two steps: contour points extraction and the CPDH construction under polar coordinates.

2.1. Contour points extraction

Different alignment methods of the points' coordinates on an object contour form different shapes. That is to say shapes are determined by the distribution of points on the object contour. So, we can use different distributions of the contour points to describe its shape. In this paper we first use the standard Canny operator to detect the object boundary. But after this operation, too many points of the contour are obtained. The algorithm in Ref. [9] is used to sample the points, which are obtained by the Canny operator, on the object boundary. Then the resultant points on the contour can be presented as $P=\{(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)\},(x_i,y_i)\in R^2$, where n denotes the number of points on the contour. The results of the contour points extraction are summarized in Fig. 1: (a) is the original image with an apple shape; (b) is the contour of the apple detected by the Canny operator; (c) is the sampled result using the method in Ref. [9]; and (d) is the sampled contour and its minimum circumscribed circle, which is used for the CPDH construction.

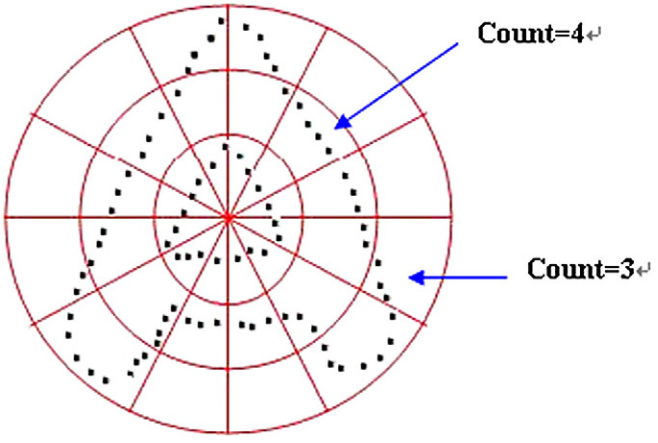


Fig. 2. Distribution of points on object contour under polar coordinates.

2.2. Construction of CPDH

After extracting and sampling the points on the object boundary, we can construct the CPDH. Firstly, we denote the object centroid as the centre and the maximum value of distance between the centre and the point on the contour is selected as radius to build a minimum circumscribed circle. An example is shown in Fig. 1(d). Then we identify the distribution over the relative position as a robust and compact, yet highly discriminative descriptor. We divide the region of the minimum circumscribed circle into several bins using some concentric circles and equal interval angle. The details of the partition are clearly presented in Fig. 2. Each bin of the minimum circumscribed circle can be described as a triple $H_i=(\rho_i,\theta_i,n_i)$, where ρ_i denotes the radius of the concentric circles θ_i denotes the angle space and n_i denotes the number of points located in the bin r_i. We call the triples of all the bins CPDH (Contour Points Distribution Histogram).

Fig. 3 demonstrates two examples of the CPDHs extracted from the original shapes. In Fig. 3, (a) and (d) are the original shape images respectively, (b) and (e) are the sampled contour points' distribution diagrams respectively, and (c) and (f) are the corresponding CPDHs respectively. From Fig. 3 we can see, that the two CPDHs are very similar, despite the distortions of the tails of the two birds, because the similarity of the major sampled contour points' distributions of the two images is large. Another reason is that the CPDHs merely take account of the number of the points in each bin instead of the points' accurate location. This loose measurement makes the CPDHs have the ability to deal with the distortions to some extent. Different colors of the bins in Fig. 3(c) and (f) represent different number points located in the bins.

As mentioned above, the whole algorithm of building the CPDH can be summarized as follows:

Algorithm 1. Construction of CPDH

Step 1: Input a binary shape image.
Step 2: Extract object contour points with Canny operator.
Step 3: Sample out N points on the contour with x and$y$ coordinates: $P=\{(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)\},(x_i,y_i)\in R^2$.
Step 4: Calculate the centroid of the shape (x_c,y_c).
Step 5: Set the centroid as the origin and translate P into polar coordinates, $P=\{(r_1,\rho_1),(r_2,\rho_2),\cdots,(r_n,\rho_n)\},(r_i,\rho_i)\in R^2$, where $r_i=((x_i-x_c)^2+(y_i-y_c)^2)^{1/2}$ is the distance between the points (x_i,y_i) and (x_c,y_c), $\theta_i=\arctan((y_i-y_c)/(x_i-x_c))$ is the angle between ρ_i and x-axis.
Step 6: Get minimum circumscribed circle C with the centre (x_c,y_c) and radius ρ_{max}, where $\rho_{max}=\max\{\rho_i\}$, $i=1,2,\cdots,n$.
Step 7: Partition the area of C into $u\times v$ bins with u bins for ρ_{max} and v bins for θ.

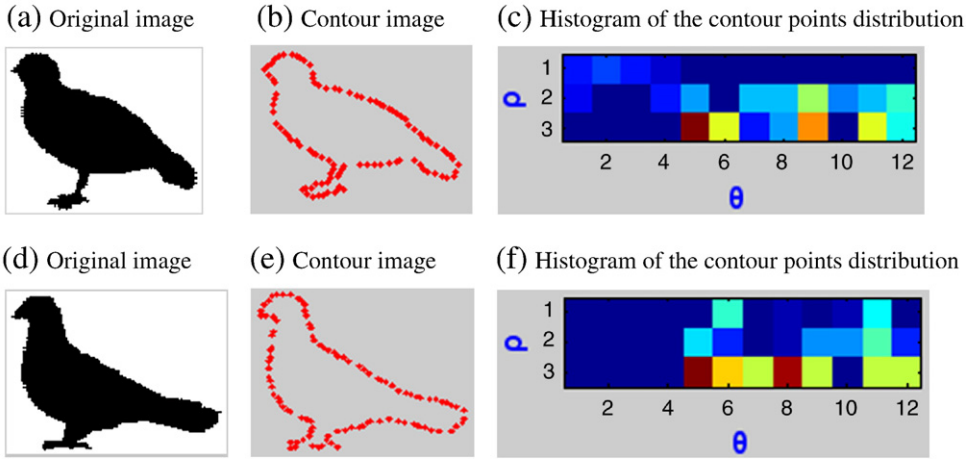


Fig. 3. Histogram of the contour points distribution.

Step 8: Construct the CPDH of the shape image by counting the number of points which are located in every bin.
Step 9: Output CPDH.

3. Similarity of the CPDHs

Similarity measure scheme is quite significant to the recognition and retrieval results. Several measures of similarity between histograms have been defined. The commonly used histogram similarity measures that are used for image recognition and retrieval are Minkowski distance [17], histogram intersection distance [17], quadratic distance [18], edit distance [19], χ^2 statistics distance and Earth Mover's Distance (EMD) [20]. Minkowski distance and histogram intersection distance belong to bin-by-bin dissimilarity measures [21], in which only pairs of bins in the two histograms that have the same index are matched. The similarity between two histograms is a combination of all the pairwise comparisons. Quadratic distance belongs to cross-bin similarity measures which does not enforce a one-to-one correspondence between mass elements in the two histograms [21]. The Edit distance between the two histograms is the number of operations required to transform one of them into the other. The χ^2 statistics distance measures how unlikely it is that one distribution was drawn from the population represented by the other. The Earth Mover's Distance (EMD) is based on the minimal cost that must be paid to transform one distribution into the other. It is more robust than the histogram matching techniques, in that it can operate on variable-length representations of the distributions that avoid quantization and other binning problems of typical histograms [21].

In this paper, considering the particularity of the CPDHs, we adopt the Earth Mover's Distance to measure the similarity between the CPDHs with a novel ground distance metric. This framework is discussed at length as follows:

3.1. CPDHs' distance with EMD

The EMD is proposed by Rubner et al. [21] to measure the dissimilarity between signatures that are compact representations of distributions. Intuitively, given two CPDHs, which are considered as two distributions, one can be regarded as a mass of earth properly spread in space, the other as a collection of holes in that same space. We can always assume that there is at least as much earth as needed to fill all the holes to capacity by switching what we call earth and what we call holes if necessary. Then, the EMD measures the least amount of work needed to fill the holes with earth.

In this paper, the EMD is formalized as the following linear programming problem: Let $P=\{(p_1,w_{p_1}),...,(p_m,w_{p_m})\}$ be the first CPDH with m clusters, where p_i is the earth representative and w_{p_i} is the weight of the cluster, i.e. number of points in the bin r_i; $Q=\{(q_1,w_{q_1}),...,(q_n,w_{q_n})\}$ be the other CPDH with n clusters, where q_i and w_{q_i} have the same meanings as p_i and w_{p_i} respectively. For those CPDHs that have equal number of bins, m is equal to n in our experiments. $D=[d_{ij}]$ is the ground distance matrix where d_{ij} is the ground distance between clusters p_i and q_j.

The following task is to find a flow $F=[f_{ij}]$, where f_{ij} is the flow between p_i and q_i, that minimizes the overall cost

$$WORK(P,Q,F)=\sum_{i=1}^{m}\sum_{j=1}^{n}d_{ij}f_{ij} \tag{1}$$

Subject to the following constraints:

$$f_{ij}\geq 0 \quad 1\leq i\leq m, 1\leq j\leq n \tag{2}$$

$$\sum_{j=1}^{n}f_{ij}\leq w_{p_i} \quad 1\leq i\leq m \tag{3}$$

$$\sum_{i=1}^{m}f_{ij}\leq w_{q_i} \quad 1\leq j\leq n \tag{4}$$

$$\sum_{i=1}^{m}\sum_{j=1}^{n}f_{ij}=min\left(\sum_{i=1}^{m}w_{p_i},\sum_{j=1}^{n}w_{q_j}\right). \tag{5}$$

Constraint (2) allows moving earth from P to Q and not vice versa. Constraint (3) limits the amount of earth that can be sent by the clusters in P to their weights. Constraint (4) limits the clusters in Q to receive no more earth than their weights and constraint (5) forces to move the maximum amount of earth possible. This amount is called total flow. Once the transportation problem is solved, and the optimal flow F is found, the EMD is defined as the resulting work normalized by the total flow:

$$EMD(P,Q)=\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}d_{ij}f_{ij}}{\sum_{i=1}^{m}\sum_{j=1}^{n}f_{ij}}. \tag{6}$$

3.2. Calculation of ground distance

In general, the ground distance d_{ij} can be any distance, i.e. Euclidean distance. But the Euclidean distance only reflects the straight-line distances between bins and does not take the transport problem into account. Considering the similarity between the CPHDs as the toll flow in a transport problem, we hope that the earth in a CPDH is moved to another CPDH along two directions, i.e. hoop direction and radial direction, respectively. This adjacent moving mechanism not only conforms to the peculiarity of the CPDHs but also fits the visual perception of the human. We propose a novel metric for the ground distance d_{ij} in the context of the CPDH. We use a diagram (Fig. 4) to explain the details of the ground distance calculation and we call it shift distance.

Firstly, according to the contour points' distribution diagram (Fig. 2), mark a serial number for each bin r_i in anti-clockwise direction from inside to outside. Fig. 4 denotes the marked result clearly. Then, how to define the ground distance between the bins in Fig. 4? An example is given as follows. If the earth is moved from one bin to another in the same concentric circle, the ground distance d_{ij} is defined as the interval value times the distance coefficient, namely, $d_{ij}=|r_i-r_j|\times\Delta_1$, for example, $d_{1,3}=|r_1-r_3|\times\Delta_1=2\times1=2$, $d_{1,12}=|r_1-r_{12}|\times\Delta_1=1\times1=1$, where Δ_1 denotes the distance coefficient between the same concentric circles. If the earth is moved between bins, which are in different concentric circles, the ground distance d_{ij} is defined as sum of the corresponding distance in the same concentric circle and the interval distance between different concentric circles, that is $d_{ij}=|r_i-r_{j-12\times\alpha}|\times\Delta_1+\alpha\times\Delta_2$, where α denotes the interval number of different concentric circles and Δ_2 denotes the distance coefficient between the two different concentric circles. In our experiments, we consider the earth moving along two directions having equal importance, Δ_1 and Δ_2 are given the same unit value of 1. For example, $d_{1,25}=|r_1-r_{25}|=|r_1-r_{25-12\times2}|\times\Delta_1+2\times\Delta_2=0\times1+2\times1=2$, $d_{1,27}=|r_1-r_{27}|=|r_1-r_{27-12\times2}|\times\Delta_1+2\times\Delta_2=2\times1+2\times1=4$. This ground distance calculation method is not only very simple but also very effective in measuring the similarity between CPDHs.

3.3. Shape distance

In this section, we make a precise definition of the shape distance. In order to resolve the rotation invariant problem to some extent, we adopt the circular shift (clockwise or anti-clockwise direction) and mirror (horizontal direction and vertical direction respectively) matching scheme to measure the distance between CPDHs. Then we choose the minimum one among the three as the final shape distance. The distance between shapes can be formalized as follows:

$$Dist(P,Q)=\ min(EMD_{CS}(P,Q),EMD_{HM}(P,Q),EMD_{VM}(P,Q)) \tag{7}$$

where, $EMD_{CS}(P,Q)$ denotes the minimum value among the circular shift matching (12 times), $EMD_{HM}(P,Q)$ denotes the minimum value of the horizontal mirror matching and $EMD_{VM}(P,Q)$ denotes the minimum value of the vertical mirror matching respectively. The algorithm for measuring the similarity between CPDHs can be summarized as follows:

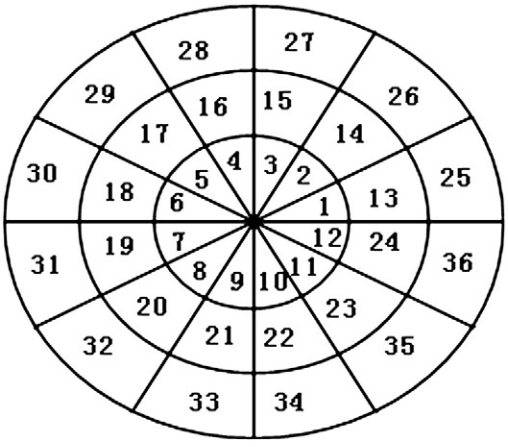


Fig. 4. Diagram used for computing the ground distance.

Algorithm 2. Similarity measure between the CPDHs

Step 1: Input two CPDHs: we denote $CPDH_1$ as the feature of the query image P and $CPDH_2$ as the feature of the target image Q

Step 2: Calculate $EMD_{CS}(P,Q)$

For i = 1 to 12

1) Calculate the ground distance matrix $D=[d_{ij}]$ between $CPDH_1$ and $CPDH_2$.
2) Calculate the distance between $CPDH_1$ and $CPDH_2$ using Eq. (6), and the distance stored in $dist(i)$.
3) Shift the $CPDH_1$ one bin on clockwise or anti-clockwise direction and keep the $CPDH_2$ unchanged.

End

Step 3: Get $EMD_{CS}(P,Q)=\min_{i=1\sim12}[dist(i)]$

Step 4: Calculate $EMD_{HM}(P,Q)$

1) Mirror the original $CPDH_1$ on horizontal direction and keep the $CPDH_2$ unchanged.
2) Calculate the ground distance matrix $D=[d_{ij}]$ between $CPDH_1$ and $CPDH_2$.
3) Calculate the distance between $CPDH_1$ and $CPDH_2$ using Eq. (6), and the distance stored in $EMD_{HM}(P,Q)$.

Step 5: Calculate $EMD_{VM}(P,Q)$

1) Mirror the original $CPDH_1$ on vertical direction and keep the $CPDH_2$ unchanged.
2) Calculate the ground distance matrix $D=[d_{ij}]$ between $CPDH_1$ and $CPDH_2$.
3) Calculate the distance between $CPDH_1$ and $CPDH_2$ using Eq. (6), and the distance stored in $EMD_{VM}(P,Q)$.

End

Step 6: Get $Dist(P,Q)=min(EMD_{CS}(P,Q),EMD_{HM}(P,Q),EMD_{VM}(P,Q))$.

Step 7: Output Dist(P,Q) (the final similarity between $CPDH_1$ and $CPDH_2$).

In the following we give an example to illustrate the advantage of this shift and mirror matching method. Fig. 5 shows two hammers and their CPDH (with 200 sample points and $5\rho\times12\theta=60$ bins). If we don't use the shift and mirror matching, their distance value is 1.205. However, if we use this shift and mirror matching mechanism, the distance value of EMD_{CS}, EMD_{HM} and EMD_{VM} are 0.115, 0.315 and 1.185 respectively. Then we choose the minimal one (0.115) as distance between the two hammers according to Eq. (7). That is to say, this shift and mirror matching method can make the distance of similar images more reasonable.

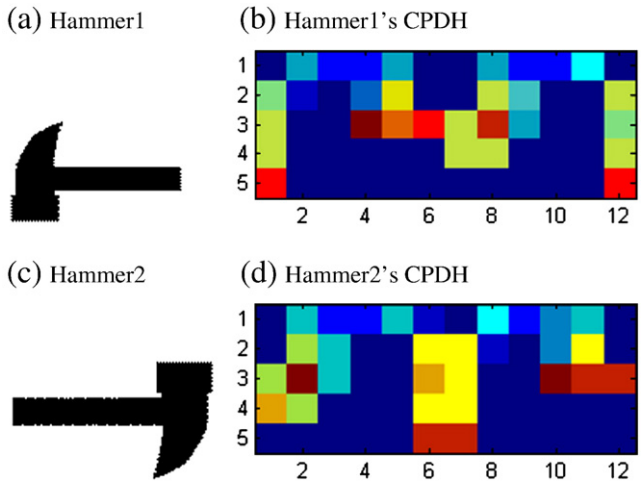


Fig. 5. Two hammers and their CPDHs.

4. Computational complexity

In this section we analyze the complexity for the CPDH representation extraction stage and its similarity matching stage independently, as they usually perform separately.

The CPDH extraction mainly involves contour extraction, points sampling, counting and calculating the distance between the sampled points and the object centroid. Therefore, the complexity of the extraction stage is $O(n)$, where n denotes the number of points sampled from the contour in the CPDH representation. All of these works can be done offline.

The complexity of matching algorithm is mainly determined by the EMD algorithm, which is a transportation problem, as our main intention is to find the optimal matching between two descriptors and show the maximum retrieval capabilities of the proposed shape representation. The original EMD algorithm has complexity of $O(n^3 \log n)$[21], where n is the number of bins used for the CPDH construction. The entire matching algorithm has the complexity of $O((k_1+k_2)n^3 \log n)$ since the circular shift and mirror matching are involved, where $k_1=12$ are the circular shift times and $k_2=2$ denotes twice mirrors (horizontal direction and vertical direction respectively). If the optimal EMD is used, the complexity can be reduced to $O(n^2)$.

5. Experimental results and analysis

In this section we will demonstrate and compare the performance of our method with some typical methods. Most publicly obtained benchmark shape databases will be considered and used for the comparison.

Table 1
Comparison of retrieval for different algorithms tested on the Kimia-25 database.

Method	Top1	Top2	Top3
SC [9]	25	24	22
Sharvit et.a [22]	23	21	20
MDS + SC + DP [23]	23	20	19
IDSC + DP [23]	25	24	25
Gdalyahu and Weinshall [24]	25	21	19
CPDH + EMD(with Eculidean distance)	25	24	22
CPDH + EMD(with shift distance)	25	24	24

5.1. Kimia's-25 shape data set

Kimia-25 provided by Sharvit et al. [22] consists of 25 shapes from six different categories as shown in Fig. 6(a). In this database a comparison between different algorithms [9,22–24] for shape retrieval is obtained by stating the number of first, second, and third nearest neighbors that fall into the correct category. A comparison between results obtained with different algorithms is shown in Table 1. It is clear that our approach outperforms the other four algorithms followed by IDSC + SC [23]. The proposed algorithm uses 200 sample points with $5\rho \times 12\theta = 60$ bins and the same parameters for the next experiments unless otherwise mentioned.

5.2. Kimia's-99 shape data set

This database consists of 99 shapes from nine categories previously tested by several references [9,23,25,26]. Examples of these shapes are shown in Fig. 6(b). A comparison of performances between different algorithms is summarized in Table 2, showing the number of top 1–10 closest matches. The best possible result for each category is 99. As shown in Table 2, our algorithm can also get better retrieval results than that of Ref. [9]. Although the performance of ours is lower than that of Refs. [23,25,26], the framework of ours is more simple and concise than theirs.

Furthermore, our method only takes the contour information into consideration and most of the others not only use the contour information but also integrate other information of the shape. The simplicity of our method makes it more feasible than others in the practical applications.

5.3. MPEG-7 shape data set (1400 shapes)

In this section, a set of simple geometric features are used to further improve the discrimination ability of the distance, which

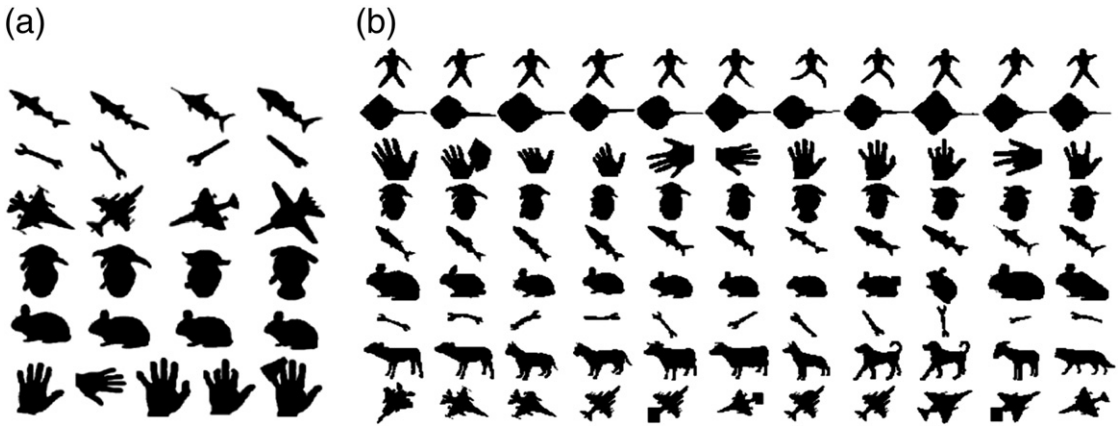


Fig. 6. (a) Kimia's data set 1, 25 instances from six categories. (b) Kimia's data set 2, 99 instances from nine categories.

Table 2
Comparison of retrieval rates for different algorithms tested on the Kimia-99 database.

Method	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	Total
SC [9]	97	91	88	85	84	77	75	66	56	37	756
Gen. model [25]	99	97	99	98	96	96	94	83	75	48	885
Shock edit [26]	99	99	99	98	98	97	96	95	93	82	956
IDSC + DP [23]	99	99	99	98	98	97	97	98	94	79	958
CPDH + EMD (with Euclidean distance)	96	94	94	87	88	82	80	70	62	55	808
CPDH + EMD (with shift distance)	98	94	95	92	90	88	85	84	71	52	849

Fig. 7. Examples of shapes in the MPEG-7 shape database. One object for each one of the 70 categories is shown.

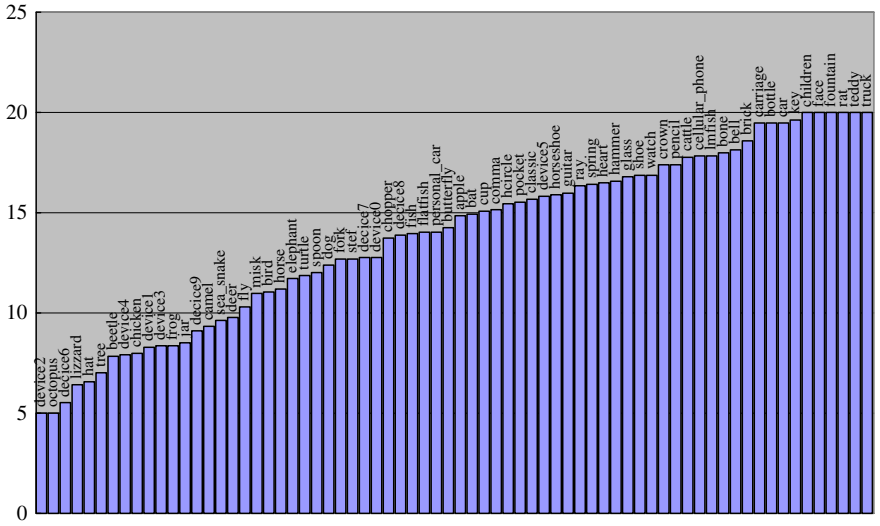


Fig. 8. Correct retrieval per category obtained by the proposed algorithm for the MPEG-7 shape database.

includes Eccentricity (E) and Solidity (S). These features include considerable information about the global properties of a shape [30]. The final distance in this experiment between shapes P and Q is given as

$$D_f(P,Q) = \mathrm{Dist}(P,Q) + \alpha_e|E_P - E_Q| + \alpha_s|S_p - S_Q| \quad (8)$$

where E_P and S_P are the Eccentricity and Solidity of shape P respectively, the same notations are adopted for shape Q, α_e and α_s are the associated weights. In this experiment, the values of weight are $\alpha_e = 0.8$ and $\alpha_s = 0.6$. It should be noted that the choice of the weights is determined by the cross-validation method. The optimal values of α_e and $\alpha_s$are specific to applications.

The MPEG-7 test database consists of 70 types of objects each having 20 different shapes, for a total of 1400 shapes (see Fig. 7 for some typical images). The database is challenging due to the presence of examples that are visually dissimilar from other members of their class and examples that are highly similar to members of other classes. For the retrieval test, we used the standard test bull's eye score [23], in which each shape is used as a test query. Retrieval is considered as correct if it is in the same class as the query. The number of correct retrievals in the top 40 ranks is counted, including the self-match. Retrieval rate for each method is reported as a percentage of the maximal possible number of correct retrievals, i.e. 28,000 (1400×20 correct retrievals). The parameters in this experiment are 200 sample points with $6\rho \times 12\theta = 72$ bins. Fig. 8 shows the average scores per category counting in top 40. Table 3 shows the reported results of different approaches on this database. Since SC [9] and IDSC [23] are closely related to the proposed method of the CPDH, we compare CPDH with SC [9] and IDSC [23] (simple shape context distance measure D_{sc} is used). The average bulls-eye score with SC is 64.59% and IDSC is 68.83%, while CPDH gets a higher score of 71.89%. Our algorithm is implemented in Matlab on a regular Pentium IV, 2.66 GHz PC. A single pairwise match takes 14 ms on average. Our algorithm is much faster than the shape context algorithm and

Table 3
Comparison of retrieval rates for different algorithms tested on the MPEG-7 database.

Algorithm	CSS[6]	Visual parts [27]	SC [9]	Aligning curves [28]	SSC [29]	General model [25]	IDSC [23]	Ours
Score%	75.44	76.45	76.51	78.16	79.92	80.03	85.40	76.56

Table 4
Comparison of retrieval rates for different algorithms tested on the subset of the MEPG-7 shape database.

Method	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	Total
SC [9]	214	209	205	197	191	178	161	144	131	101	78	1809
CPDH + EMD (with Euclidean distance)	214	215	209	204	200	193	187	180	168	146	114	2030
CPDH + EMD (with shift distance)	215	215	213	205	203	204	190	180	168	154	123	2070

those using statistical optimization approaches. For example, with 100 landmarks, the SC algorithm takes about 1.2 s for a single comparison of two shapes, the IDSC reports 0.31 s on a 2.8 GB PC. Though the performance of the proposed algorithm is not at par with the other recently proposed shape matching algorithms, it is a very practical one due to its simplicity and rapidity.

5.4. Comparison with the shape context

Although there are many shape representation and analysis methods based on the contour information, the most similar to ours is the shape context [9]. The CPDH, is evolved from the SC, but has the merits of itself and is more simple and robust than the SC. So we compare their performance through an experiment which is conducted on a subset of MPEG-7 database. This sub database contains 216 images from 18 classes [23]. For each shape, we check whether the 11 closest matches are in the same class as the query. In Table 4, the number of correct matches in each rank is summarized. It is clear that the CPDH is better than the SC. In Sections 5.4, 5.5 and 5.6, the shape distance is as usual calculated by Eq. (7).

5.5. Comparison with some other typical methods

In order to further explain the performance of the CPDH, we use the common performance measures, i.e., average precision and recall of the retrieval as the evaluation measures. Precision P is defined as the ratio of the number of retrieved relevant shapes r to the total number of retrieved shapes n, i.e., $P=r/n$. Recall R is defined as the ratio of the number of retrieved relevant images r to the number m of relevant shapes in the whole database, i.e. $R=r/m$. Fig. 9 expresses the comparison results of some different distance metrics used for the similarity measure of the CPDH. From Fig. 9, it is clear that the EMD is superior to the other distance metrics for the CPDHs' similarity. Fig. 10 expresses the comparison results of some other typical shape analysis methods: Fourier descriptors (FD), Curvature Scale Space (CSS), and Boundary moments (BM). CSS has been selected as the criterion shape descriptor by MPEG-7. From Fig. 10, it is clear that the performance of CPDH + EMD is still very competitive.

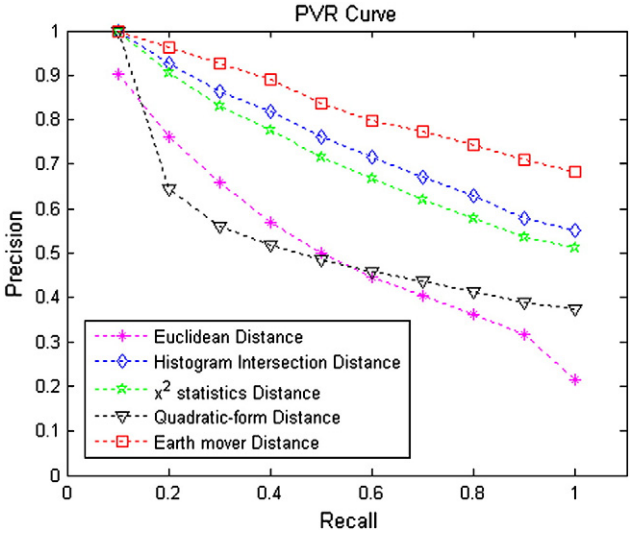


Fig. 9. Average retrieval precision–recall curve for shape retrieval with different distance metrics in the MPEG-7 shape database.

5.6. Rotation invariant test

As described in the Abstract, the invariance to scale and translation are the intrinsic properties for the CPDH and the invariance to rotation can be partially resolved in the matching scheme, i.e., the circular shift and mirror match. For this purpose, we select 10 images randomly from the MPEG-7 shape database as benchmark images and rotate them to form 12 new images with 5°, 15°, 30°, 45°, 60°, 75° in clockwise and anti-clockwise directions respectively. Including the 10 benchmark images, the rotation shape database has 130 images. Table 5 lists the retrieval results with the 10 benchmark images as query images. There are only 3 images, with gray background in the table grids, which do not belong to the query kinds in the all front 12 retrieved images.

6. Conclusion

We have presented a new shape representation CPDH and the matching methods based on the distribution of points on the contour of objects under polar coordinates. CPDH is intrinsically insensitive to scale and translation and the rotation invariance can be partially obtained through the circular shift and mirror matching scheme. The simplicity of the computations of the shape representation along with the flexibility of the EMD similarity measure is another contribution of this paper. We conducted several experiments on the benchmark databases. The results of the experiments show that the proposed method is effective for shape matching and retrieval. Although the performance of our approach is not the best one, it might be one of the most feasible approaches in practical applications. The deficiency of the CPDH is that it only fits for the shapes with a single closed contour. The CPDH cannot deal with those images with multiple objects such as trademark images. Future work will focus on building more elaborate shape descriptors and the experiments on the 3D shape datasets.

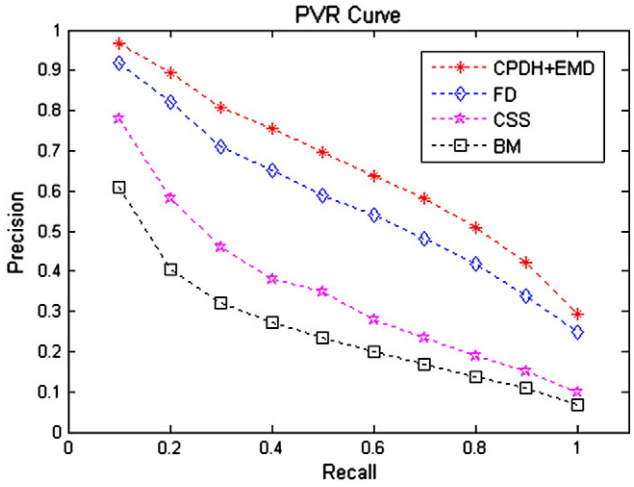
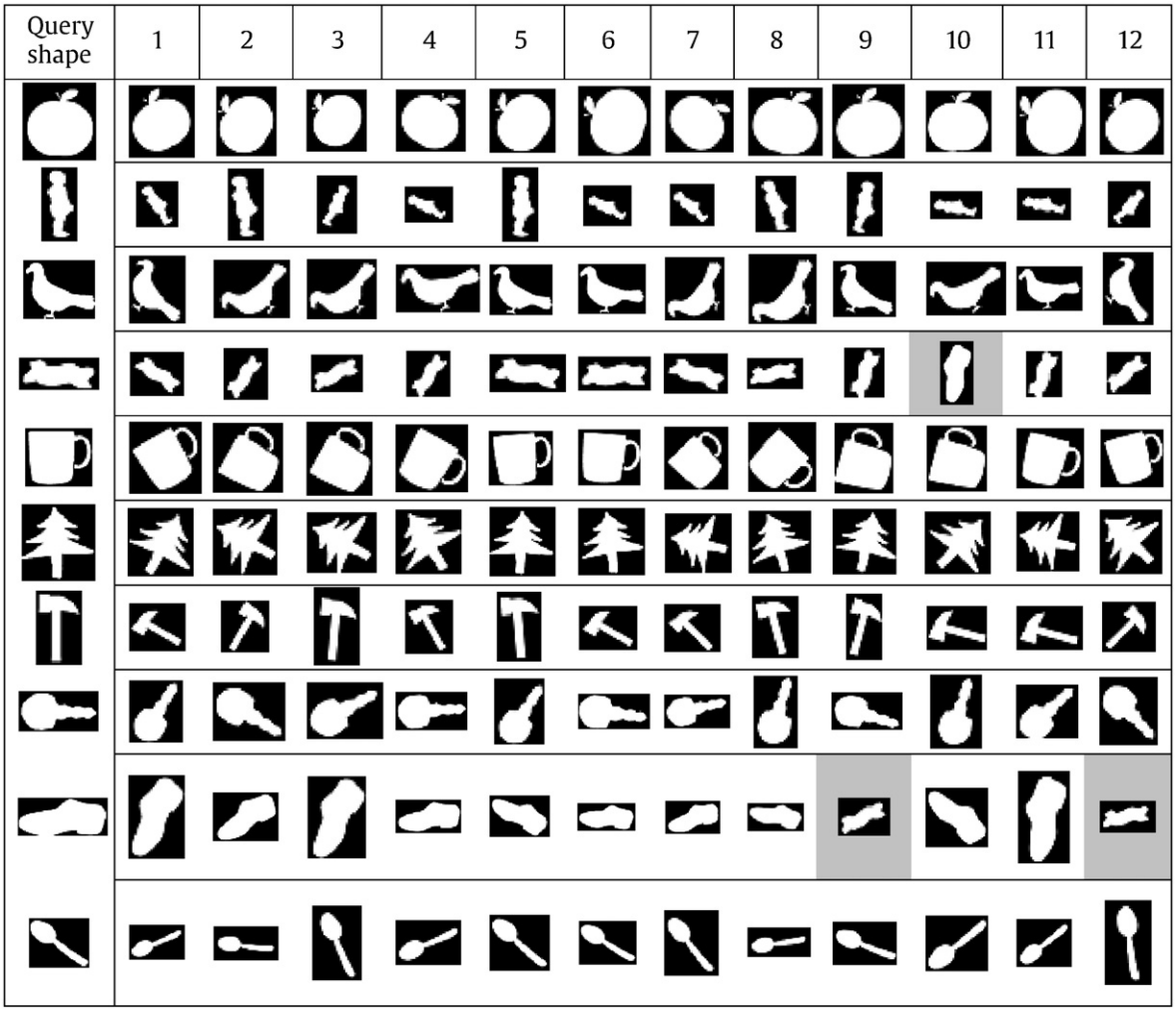


Fig. 10. Comparison of PVR curves for several typical methods in the MPEG7 shape database.

Table 5
Rotation invariant query results on the rotation shape dataset. Left column shows query shapes and the right rows show the first 12 ranked nearest neighbors for each query shape, respectively.



Acknowledgements

The authors would like to thank the anonymous reviewers for their constructive advices. This work is supported by National Natural Science Foundation of PR China (Grant Nos. 60572034, 60973094) and Natural Science Foundation of Jiangsu Province (Grant No. BK2006081) and Natural Science Foundation of Jiangsu Provincial Universities (Grant No.10KJB520006). The authors thank Dr. Yossi Rubner for making the EMD codes available on the web.

References

[1] L. Da, F. Costa, R.M. Cesar Jr., Shape Analysis and Classification: Theory and Practice, CRC Press, 2000.
[2] S. Loncaric, A survey of shape analysis techniques, Pattern Recognition 31 (1998) 983–1001.
[3] C.J. Xu, J.Z. Liu, X. Tang, 2D shape matching by contour flexibility, IEEE Transactions on Pattern Analysis and Machine Intelligence 31 (1) (2009) 180–186.
[4] H. Kim, J. Kim, Region-based shape descriptor invariant to rotation, scale and translation, Signal Processing: Image Communication 16 (2000) 87–93.
[5] D.S. Zhang, G.J. Lu, Review of shape representation and description techniques, Pattern Recognition 37 (2004) 1–19.
[6] F. Mokhtarian, F. Abbasi, J. Kittler, in: SmeuldersA.W.M. , JainR. (Eds.), Efficient and Robust Retrieval by Shape Content through Curvature Scale Space, Image Databases and Multi-Media Search, 1997, pp. 51–58.
[7] Y. Rui, A. She, T.S. Huang, A modified Fourier descriptor for shape matching in MARS, Image Databases and Multimedia Search 8 (1998) 165–180.
[8] I. Bartolini, P. Ciaccia, M. Patella, WARP: accurate retrieval of shapes using phase of Fourier descriptors and time warping distance, IEEE Transactions on Pattern Analysis and Machine Intelligence 27 (2005) 142–147.
[9] S. Belongie, J. Malik, J. Puzicha, Shape matching and object recognition using shape contexts, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (4) (2002) 509–522.
[10] M.R. Daliri, V. Torre, Robust symbolic representation for shape recognition and retrieval, Pattern Recognition 41 (2008) 1782–1798.
[11] C. Grigorescu, N. Petkov, Distance sets for shape filters and shape recognition, IEEE Transactions on Image Processing 12 (10) (2003) 1274–1286.
[12] Emad Attalla, Pepe Siy, Robust shape similarity retrieval based on contour segmentation polygonal multiresolution and elastic matching, Pattern Recognition 38 (2005) 2229–2241.
[13] W.Y. Kim, Y.S. Kim, A region-based shape descriptor using Zernike moments, Signal Processing: Image Communication 16 (2000) 95–102.
[14] S.X. Liao, M. Pawlak, On image analysis by moments, IEEE Transactions on Pattern Analysis and Machine Intelligence 18 (3) (1996) 254–266.
[15] D.S. Zhang, G.J. Lu, Shape-based image retrieval using generic Fourier descriptor [J], Signal Processing: Image Communication 17 (2002) 825–848.
[16] Y.W. Chen, C.L. Xu, Rolling penetrate descriptor for shape-based image retrieval and object recognition, Pattern Recognition Letters 30 (2009) 799–804.
[17] M.J. Swain, D.H. Ballard, Color indexing, International Journal of Computer Vision 7 (1) (1991) 11–32.
[18] W. Niblack, R. Barber, W. Equitz, M.D. Flickner, E.H. Glasman, D. Petkovic, P. Yanker, C. Faloutsos, G. Taubin, Y. Heights, Querying images by content, using color, texture, and shape, In SPIE Conference on Storage and Retrieval for Image and Video Databases 1908 (1993) 173–187.

[19] A. Marzal, E. Vidal, Computation of normalized edit distance and applications, IEEE Transactions on Pattern Analysis and Machine Intelligence 15 (9) (1993) 926–932.
[20] Y. Rubner, L.J. Guibas, C. Tomasi, The earth mover's distance, multi-dimensional scaling, and color-based image retrieval, Proceedings of the ARPA Image Understanding Workshop, 1997, pp. 661–688.
[21] Y. Rubner, C. Tomasi, L.J. Guibas, The earth mover's distance as a metric for image retrieval, International Journal of Computer Vision 40 (2) (2000) 99–121.
[22] D. Sharvit, J. Chan, H. Tek, B. Kimia, Symmetry-based indexing of image database, Journal of Visual Communication and Image Representation 9 (4) (1998) 366–380.
[23] H.B. Ling, D.W. Jacobs, Shape classification using the inner-distance, IEEE Transactions on Pattern Analysis and Machine Intelligence 29 (2) (2007) 286–299.
[24] Y. Gdalyahu, D. Weinshall, Flexible syntactic matching of curves and its application to automatic hierarchical classification of silhouettes, IEEE Transactions on Pattern Analysis and Machine Intelligence 21 (12) (1999) 1312–1328.
[25] Z. Tu, A. Yuille, Shape matching and recognition-using generative models and informative features, Proceedings of European Conference on Computer Vision III (2004) 195–209.
[26] T.B. Sebastian, P.N. Klein, B.B. Kimia, Recognition of shapes by editing their shock graphs, IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (5) (2004) 550–571.
[27] L.J. Latecki Test, Shape similarity measure based on correspondence of visual parts, Intelligence 22 (10) (2000) 1185–1190.
[28] T. Sebastian, P. Klein, B. Kimia, On aligning curves, IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (1) (2003) 116–125.
[29] J. Xie, P. Heng, M. Shah, Shape matching and modeling using skeletal context, Pattern Recognition 41 (5) (2008) 1756–1767.
[30] N. Alajlan, I. Elrube, M.S. Kamel, G. Freeman, Shape retrieval using triangle-area representation and dynamic spacewarping, Pattern Recognition 40 (7) (2007) 1911–1920.

Xin Shu received B.E. and M.E. degrees both in Computer Science and Technology from Jiangsu University of Science and Technology (JUST), P.R. China in 2002 and 2005, respectively. He is currently a Ph.D. candidate of Jiangnan University, Wuxi, P.R. China. His main research interests include computer vision, shape matching and image/video retrieval.

Xiao-Jun Wu received his B.S. degree in mathematics from Nanjing Normal University, Nanjing, P.R. China in 1991 and M.S. degree in 1996, and Ph.D. degree in Pattern Recognition and Intelligent System in 2002, both from Nanjing University of Science and Technology, Nanjing, P.R. China, respectively. He was a fellow of the United Nations University, International Institute for Software Technology (UNU/IIST) from 1999 to 2000. He won the most outstanding postgraduate award by Nanjing University of Science and Technology. From 1996 to 2006, he taught in the School of Electronics and Information, Jiangsu University of Science and Technology where he was an exceptionally promoted professor. He joined the School of Information Engineering, Jiangnan University in 2006 where he is a professor. He has published more than 100 papers. He was a visiting researcher in the Centre for Vision, Speech, and Signal Processing (CVSSP), University of Surrey, UK from 2003 to 2004. His current research interests are pattern recognition, computer vision, and intelligent systems.