Angular Pattern and Binary Angular Pattern for Shape Retrieval

Rong-Xiang Hu, Wei Jia, *Member, IEEE*, Haibin Ling, *Member, IEEE*, Yang Zhao, and Jie Gui

***Abstract*— In this paper, we propose two novel shape descriptors, angular pattern (AP) and binary angular pattern (BAP), and a multiscale integration of them for shape retrieval. Both AP and BAP are intrinsically invariant to scale and rotation. More importantly, being global shape descriptors, the proposed shape descriptors are computationally very efficient, while possessing similar discriminability as state-of-the-art local descriptors. As a result, the proposed approach is attractive for real world shape retrieval applications. The experiments on the widely used MPEG-7 and TARI-1000 data sets demonstrate the effectiveness of the proposed method in comparison with existing methods.**

***Index Terms*— Shape retrieval, angular features, binary pattern, multi-scale representation.**

I. Introduction

With the rapid increase of archived images captured by various of acquisition devices, there is an urgent demand for effective and efficient image retrieval algorithms. For large image databases, textual annotation-based image retrieval is often impractical and inefficient, due to the requirement of intensive human labor. Consequently, content-based image retrieval (CBIR) systems have received great attention in recent years. Instead of using textual annotation, CBIR systems typically rely on automatically extracted visual features for image representation. Shape information often plays a key role in human cognition system due to their strong discriminative power. Despite a great amount of previous research in shape analysis, it remains a challenging task to design a shape-based retrieval system that is both effective (in terms of retrieval accuracy) and efficient (in terms of responding time). As a result, shape-based image retrieval systems are generally less popular than appearance-based ones. In a typical shape-based CBIR system, image segmentation is first usually performed to obtain (binary) shapes, and then shape descriptors are extracted from them for comparison. Normally, in order to achieve desirable retrieval performance, shape descriptors are required to be insensitive to similarity transformation and to tolerate a small degree of non-rigid distortion.

Due to the importance of shape information in image understanding, many shape comparison algorithms have been proposed [1]–[19], which can be roughly divided into two categories: contor-based and region-based. The contor-based methods are much more popular than the region-based ones in the past decade, and they can be further classified into two subcategories: global and local.

In global shape comparison, a shape is usually represented by a feature vector extracted from the whole contor, and shape comparison is conducted by comparing such representation vectors. A classic global shape representation is the curvature scale space (CSS) [3], which has been recommended by the MPEG-7 community as one of the standards. In CSS, the zero-crossings of the contor curvature function are located at different scales. These zero-crossings form a CCS image, and the maxima of such CCS image contors are used for shape matching. Another example of global method is the polygonal multi-resolution and elastic matching (PMEM) [9], in which three primitives of each contor segment are extracted at different scales. Then, the sum of absolute differences (SAD), improved by the elastic matching, is used to measure the similarity between shapes. Another recent global method is the contor points distribution histogram (CPDH) [19], which represents a shape by the spatial distribution of contor points in the polar coordinate system and compares such distributions using the Earth Mover's Distance (EMD) [29].

In local shape comparison, a shape is typically represented by a set of local descriptors such that each descriptor captures only local shape information. Shape similarities are then derived by two layers of comparison: the low-level similarities between local descriptors and the high-level comparison on top of such similarities. A classic example in this category is the shape context (SC) [1], which uses 2-D histograms to capture the spatial context around each landmark point, and compares two shapes by matching two sets of such histograms. SC becomes very popular because of its powerful descriptive ability and is extended in various ways. One such extension is the inner distance shape contexts (IDSC) [5], which replaces the Euclidean distance with the inner distance to achieve robustness against articulation. In addition, in [5] dynamic programming (DP) is used to utilize the continuity

Manuscript received November 20, 2012; revised April 19, 2013 and August 9, 2013; accepted September 30, 2013. Date of publication October 18, 2013; date of current version January 27, 2014. This work was supported by the National Science Foundation of China under Grants 61305006, 61175022, 61100161, and 60705007. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Xilin Chen.

R.-X. Hu and W. Jia are with the Institute of Nuclear Energy Safety Technology, Chinese Academy of Sciences, Hefei 230031, China (e-mail: hurongxiang2008@gmail.com; icg.jiawei@gmail.com).

H. Ling is with the Department of Computer and Information Sciences, Temple University, Philadelphia, PA 19122 USA (e-mail: hbling@temple.edu).

Y. Zhao is with the School of Electronic and Computer Engineering, Peking University, Shenzhen 518055, China (e-mail: zyknight@mail.ustc.edu.cn).

J. Gui is with the Hefei Institute of Intelligent Machines, Chinese Academy of Sciences, Hefei 230031, China (e-mail: guijiejie@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2013.2286330

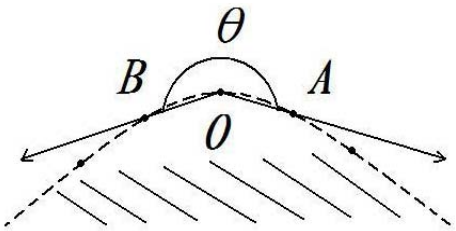


Fig. 1. Illustration of Angular Pattern (AP). Dashed line indicates the shape contor.

constraint on contor points. In [10], hierarchical Procrustes matching (HPM) is proposed to capture shape information across different hierarchies. The shape tree (ST) in [2] also captures hierarchical geometric propensities of a shape, and it utilizes a tree matching method for shape comparison. In [16], a descriptor named contor flexibility (CF) is proposed which describes each contor point by its deformable potential. CF also uses DP for shape matching. Recently, a novel shape descriptor using height functions (HF) is proposed in [23] and DP is again used for matching such descriptors.

Local shape descriptors usually outperform global ones in terms of shape matching accuracy. However, such superiority is typically at a cost of reduced efficiency in terms computation time. By contrast, global methods achieve better run time efficiency thanks to the simplicity in their shape representation and associated shape matching algorithms. Modern CBIR systems often deal with large image datasets, and are therefore in favor of fast query response. Consequently, both effectiveness and efficiency are desired by such systems [20]. In fact, MPEG-7 has set several criteria [21] to evaluate a shape descriptor, including high retrieval accuracy, feature compactness and low computational complexity. Notice that in recent years, the multi-scale framework [2], [9], [10] introduced in many methods has been proved to be able to enhance the performance of the original single-scale method. It is worth to mention that most of the existing methods rely on scale normalization to achieve scale invariance and circular shifting to achieve the rotation insensitivity, both of which request extra computational time.

Instead of designing new shape descriptors and/or shape comparison methods, some recent studies improve shape retrieval by exploring retrieval techniques, such as in [24]–[26]. These methods can be viewed as post-shape-matching enhancement and can be divided into three categories: context-based, knowledge-based and fusion-based. More details about these methods can be found in [27].

Motivated by the above discussion, in this paper we propose novel global shape representations that are efficient for shape retrieval. The key idea is to capture the angular information among contor points using the proposed angular pattern (See Fig. 1), and then build rich shape descriptors using such patterns. Angular patterns are naturally invariant to similarity transformation, including scaling and rotation, which means neither normalization nor circular shifting is required when comparing related descriptors. In addition to the original angular pattern, another variation named binary angular pattern is proposed inspired by the recent success

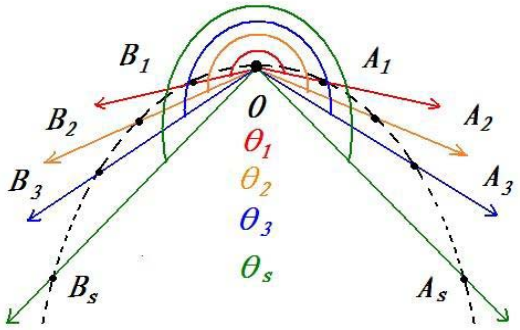


Fig. 2. Illustration of multi-scale AP.

of binary texture descriptors [22]. Furthermore, we organize the angular information in a multi-scale framework to further improve its discriminative power.

The proposed descriptors are tested for shape retrieval tasks using two public benchmark datasets. Being a global shape descriptor, our method runs much faster than popular local shape descriptors with only marginal sacrifice in shape matching accuracy. That said, our approach outperforms previously proposed global shape descriptors in retrieval accuracy. In other words, our method achieves a promising balance between retrieval accuracy and retrieval speed, and is therefore attractive for large scale shape retrieval applications.

II. Angular Pattern and Binary Angular Pattern

A. Angular Pattern

Given n uniformly sampled points on a contor, geometric features (e.g., curvatures and distances to the centroid) are often used for shape description. Neither curvature nor distance to a shape centroid, however, is invariant to scale or can be easily extended to multi-scale [20]. To overcome this limitation, we introduce a novel angular feature for each contor point as illustrated in Fig. 1. It should be noted that the dashed line in the figure is the shape contor and the shadowed region represents the interior part of the shape (assuming that contors are anticlockwise).

Given a point O on the contor, let A and B be two points before and after O respectively on the contor and they have the same distance to O. The anticlockwise angle θ (ranging from 0 to 2π) between vector OA and vector OB can be used to describe the geometric properties of point O. This angular feature can be extended to multi-scale naturally. Specifically, let A_s and B_s points that are s points away from O in the index order for an integer s ranging from 1 to floor((n−1)/2). Then, as shown in Fig. 2, we define θ_s as the angle between OA_s and OB_s. Finally, given a shape with n contor points and a specific s, n such angles can be drawn, and these angles capture the geometric property of the shape at scale s. We call the representation the Angular Pattern (AP) of the shape. Obviously, for a contor with n points, we can extract in total floor((n−1)/2) APs at different scales, and these APs are all invariant to scale and rotation.

To use these angular features for efficient shape retrieval, a histogram of AP, denoted as H, is generated according to the following steps:

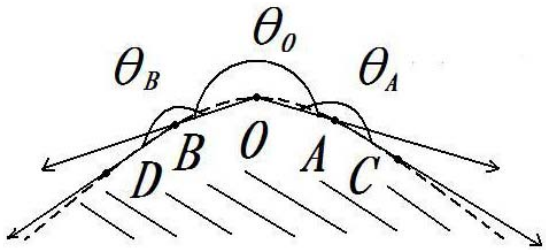


Fig. 3. Illustration of the simplest BAP.

Step 1: Given a contor with n points and a specific s, we extract n angular features of this contor, i.e., $AP_s = \{AP_s(i),\ i = 1, 2, \ldots, n\}$.

Step 2: We uniformly divide $[0, 2\pi]$ into K bins.

Step 3: We construct the histogram $H = \{h(k),\ k = 1, 2, \ldots, K\}$, with each bin as:

$$h(k) = card\{a | a \in AP_s, \frac{k-1}{K} \cdot 2\pi < a \leq \frac{k}{K} \cdot 2\pi\}, \quad k = 1, 2, \ldots, K \tag{1}$$

To compare two AP histograms, namely H_{S1} and H_{S2}, we use the χ^2 distance defined below:

$$D(H_{S_1}, H_{S_2}) \equiv \frac{1}{2} \sum_{1 \leq k \leq K} \frac{[h_{S_1}(k) - h_{S_2}(k)]^2}{h_{S_1}(k) + h_{S_2}(k)} \tag{2}$$

where $h_{S_1}(k)$ and $h_{S_2}(k)$ are the k-th bin of the corresponding histograms, and K is the number of histogram bins. Varying the scale s from 1 to floor((n−1)/2), we can construct floor((n−1)/2) AP histograms, and each of them can be used to retrieval shapes independently. However, using one scale alone may not generate satisfactory performance due to the lack of sufficient discriminant information. We will latter describe how to integrate AP over different scales to enhance the discriminant power.

B. Binary Angular Pattern

By definition, AP does not capture the relations between the adjacent angular features. Recently, in computer vision and pattern recognition, an emerging useful idea is to capture local intensity relations (also called local structural pattern) for applications such as texture classification. A representative example is the Local Binary Pattern (LBP) [22]. In LBP, intensities of pixels within a region are compared with the intensity of the pixel in the center of the region. A pixel is coded with 1 if its intensity is larger than the intensity of the central pixel, and 0 otherwise. By this way, a binary pattern describing the local region is generated and a histogram of different patterns is constructed over the image. The histogram of LBP is proven to be capable of capturing the local intensity context and has produced superior performance in texture classification. Inspired by the idea, we propose the Binary Angular Pattern (BAP) to encode structure information among contor points.

The simplest BAP is illustrated in Fig. 3. Given a point O on the contor and its equal distance point pairs (A, B) and (C, D), a segment containing five points can be constructed.

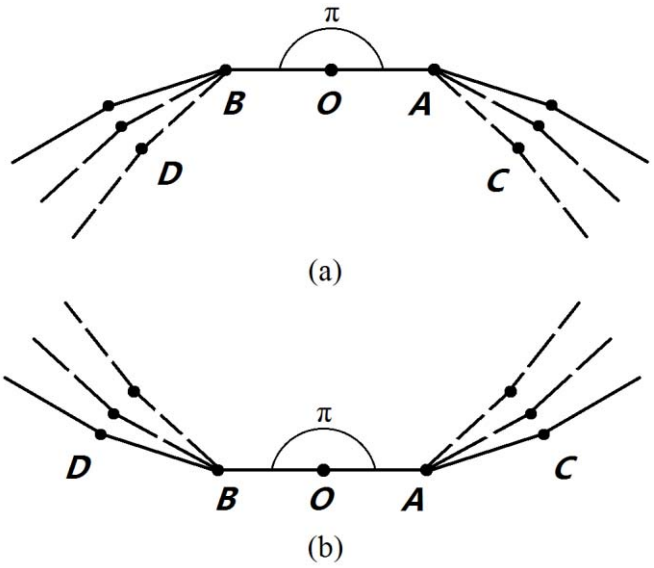


Fig. 4. Illustration of the BAP feature to small degrees of non-rigid distortion which does not affect the binary angular relations: (a) all patterns are [1 1]; (b) all patterns are [0 0].

Obviously, three angular features, i.e., θ_O, θ_A and θ_B can be calculated. Then, θ_A and θ_B are compared with θ_O. If the angle θ_A or θ_B is larger than θ_O, the corresponding point is coded with 1, otherwise, it will be coded with 0. Consequently, for this segment with five points, a two-bit binary pattern can be generated, which is robust to a small degree of non-rigid distortion. As shown in Fig. 4(a) and (b), for $\theta_O = \pi$, there are two cases. First, θ_A and θ_B are greater than θ_O, and in this situation the BAP of point O is always coded as [1 1] for all three deformed contors depicted in Fig. 4(a). Similarly, the BAP of point O in Fig. 4(b) is [0 0] for all three contors for the second case in which both θ_A and θ_B are smaller than θ_O.

Since AP is multi-scale, the proposed simplest BAP can also be extended to multi-scale. If the pattern is constructed with O, A_s, B_s, C_s and D_s as defined before, then the computed two-bit binary pattern will capture another scale of angular relations. Finally, floor((n−1)/2) scales of BAP with two-bit pattern can be extracted for a contor containing n points.

To use these angular relations for efficient shape retrieval, a histogram of BAP is generated. Since there are only two bits in the simplest BAP pattern as illustrated in Fig. 3, the histogram may contain only 4 bins corresponding to all the 4 patterns ([0 0], [0 1], [1 0] and [1 1]). Obviously, the number of bins is too small to provide sufficiently discriminative information; thus, we introduce BAP features with more bits. For the sake of convenience, we call a pattern BAPmP if it contains m bits, where m angular features of neighbor points are compared with the angular feature of the central point respectively and result in m-bit binary pattern. As illustrated in Fig. 5, patterns of 4-bit, 6-bit, 8-bit, 10-bit and 12-bit can be extracted. Consequently, the histograms of these patterns contain 16, 64, 256, 1024 and 4096 bins, respectively. To index bins of the histogram, each bit is assigned with a weight as shown in Fig. 5, and the weighted sum of values of bits is

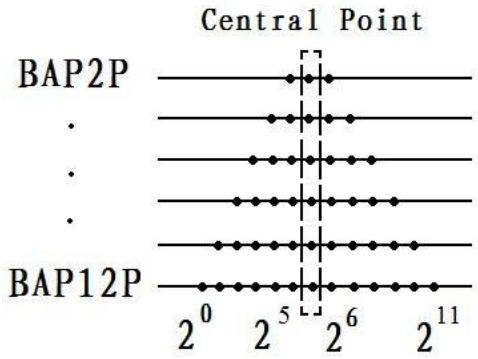


Fig. 5. Illustration of BAPmP while $s = 1$.

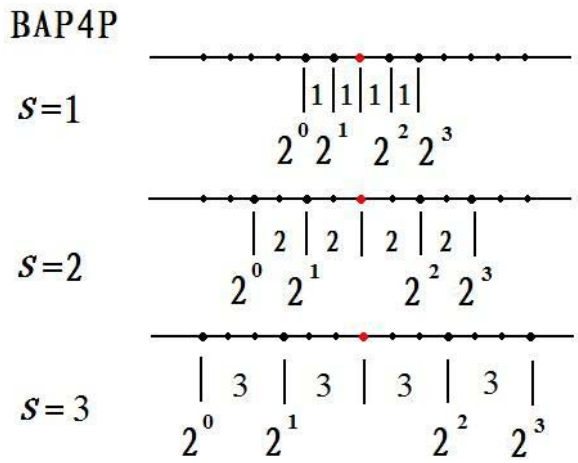


Fig. 6. Illustration of BAP4P in two scales $s = 1$, $s = 2$ and $s = 3$.

the index. It should be noted that m is an even number due to the bilateral symmetry of the points around the central point. Additionally, patterns larger than 12 bits are not considered in this paper, because in common shape representation there are at most a few hundreds of points, and a histogram of larger pattern would contain more than 10 thousand bins and would not be robust enough. It is worth mentioning that extracting BAP feature of m bits will intrinsically involve m + 3 points, i.e. one central point and m/2 + 1 points on each side of the central point for extracting basic AP feature.

The BAPmP can also be extended to a multi-scale representation. In particular, the BAPmP with a specific scale s is constructed as follows: Firstly, given a central point, the AP features are extracted from points whose distances to the central point are $t \times s$, where t is a positive integer and varies from 1 to $m/2$. Secondly, these AP features are compared with the AP feature (also with scale s) of the central point to form the m bits binary pattern. This way, m-bit binary patterns of all points on the contor can be extracted with the scale s.

An example including three scales ($s = 1$, s = 2 and $s = 3$) of BAP4P is shown in Fig. 6. In the figure, the red points are the central point whose BAP4P feature will be extracted, and black dots are neighboring contor points. For the case of $s = 1$, the AP features with scale 1 are firstly extracted from four larger black points around central red point. Obviously, the distances from these four points to central point are 1 and 2, respectively, since $t = \{1, 2\}$ and $s = 1$. Then, these four AP features are compared with the AP feature (with scale 1) of the red central point to form the 4-bit binary pattern. For the case $s = 2$, the AP features with scale 2 are firstly extracted from the four larger black points, from which the distances to the red central point are 2 and 4 (here, $t = \{1, 2\}$ and $s = 2$), respectively. Obviously, those smaller black points are not involved in the computation. After comparing the four AP features of the larger black points with the AP feature of the red central point, the 4-bit binary pattern of BAP4P with scale 2 are obtained. Similarly, for the case $s = 3$, the AP features with scale 3 are firstly extracted from the four larger black points whose distances to the red central point are 3 or 6 (here, $t = \{1, 2\}$ and $s = 3$), respectively. These AP features are then compared with the AP feature of the red central point to form the 4-bit binary pattern of BAP4P with scale 3. In all scales, the 4-bit binary patterns are weighted by 2^0, 2^1, 2^2 and 2^3, respectively.

Following the above example, the multi-scale representations of BAPmP features for other m can be extracted similarly. Histograms of such BAPmP features can then be calculated and used for shape retrieval. In particular, after constructing these histograms, the χ^2 distance is used for comparing them and consequently for shape matching. It is worth noting that for the scale s of BAPmP, $m/2 \cdot s$ is upper-bounded by floor((n−1)/2), since the number of bins of the BAPmP pattern shall not exceed the number of contor points. Using BAPmP features alone may not generate satisfactory results for shape retrieval, due to the discarding of more precise values of angular features. In the next section, we will describe how to integrate multi-scale AP and BAP together to address this issue. It is worth noting that the proposed BAP is intrinsically invariant to scale and rotation.

LBP is known to be very effective for texture and appearance representation. How will AP and BAP perform for shape matching? To answer this question, two simple examples of AP and BAP2P are illustrated below. Generally speaking, shape consists of three types of elements: flat area, outward (convex) strands and inward (convace) strands, and the proposed AP and BAP can provide certain geometry constrains to depict these elements. The first type is shown in Fig. 4, which can be well described by combining AP with BAP2P and is robust to small non-rigid shape distortion which does not change angular relations. For the second type, the examples of outward strands are shown in Fig. 7. Two cases are given to illustrate the geometric constrains of AP and BAP2P, where both BAP2P features of point O are [1 1]. In Fig. 7(a), θ_O is fixed to $\frac{4}{3}\pi$. If θ_A and θ_B vary simultaneously, to maintain the same BAP2P feature, the diamond shape outward strands are expected. While in Fig. 7(b), θ_A, θ_A and θ_B vary simultaneously, geometric constrains of AP and BAP2P will also result in certain outward strands. Similarly, in Fig. 8 two cases of inward strands are given, which belong to the third type, and both cases are under the geometric constrain that BAP2P of point O is [0 0]. From the above examples, we expect that different combination of AP and

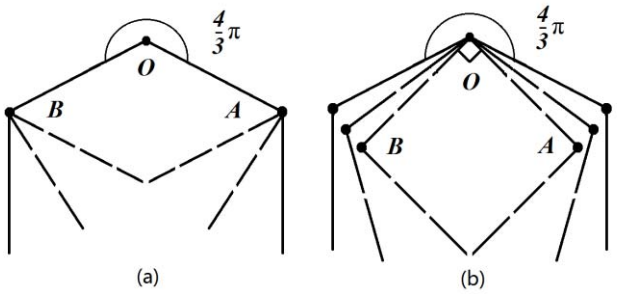


Fig. 7. Illustration of the AP and BAP2P to describe certain outward strands: (a) θ_0 does not change; (b) θ_0 changes.

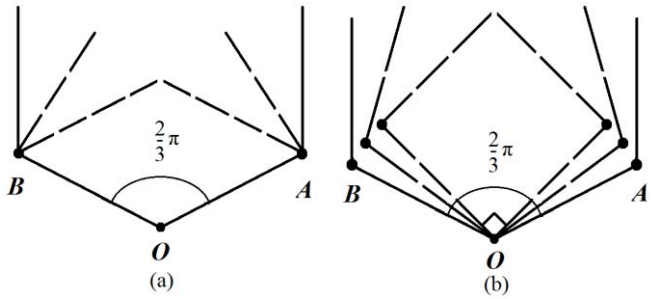


Fig. 8. Illustration of the AP and BAP2P to describe certain inward strands: (a) θ_0 does not change; (b) θ_0 changes.

BAP can accommodate different geometric constrains and hence describe different structures of the shape.

C. Multi-Scale Integration

We propose a score-level fusion for multi-scale integration of AP and BAP. Given M shapes, for any specific scale of AP and BAP, a distance matrix D of size $M \times M$ is computed with each element measuring the distance between two of the M shapes. The matrix is then normalized to a matrix D_N with zero-mean and regularized standard deviation, i.e., z-score normalization.

$$D_N(i, j) = \frac{D(i, j) - mean(D)}{std(D)}, \quad i, j = 1, 2, \ldots, M \quad (3)$$

Assuming there are in total L normalized distance matrices from AP and BAP denoted as D_{Nl}, $l = 1, 2, \ldots, L$ corresponding to L scales of AP and BAP, it is desirable to select some of them to form a new distance matrix, denoted as D^*, that best retrieves the known shapes. Then, as the combination of D_{Nl} is determined, so is the selection of scales of AP and BAP. We use the Sequential Forward Selection (SFS) to solve this problem. All D_{Nl} are considered as candidates at the beginning and none is selected. The entries of D^* are initialized to zeros. In each iteration, the candidate distance matrices are added to D^* in turn, and the one that achieves the best retrieval performance is chosen. The scale at which a candidate corresponding to the best performance is accepted as a newly selected scale. The selected candidate matrix is then removed from the candidate set. This process continues until no new candidate is accepted, which means that integrating one more scale will hurt the retrieval performance. Finally the output is the recorded indexes of scales. SFS may not find the best subset of scales of AP and BAP, but it is effective and efficient. In practice, the multi-scale AP and BAP can achieve a much higher retrieval rate than the single-scale counterparts. Notice that in real applications, the multi-scale subset of AP and BAP is trained from known shapes and then fixed for retrieval, so the SFS will not affect the retrieval efficiency of the proposed method.

III. Experiments and Discussions

We first test the single-scale AP and BAP on the MPEG-7 dataset to show that they are unable to provide satisfactory performances. After that, the multi-scale versions are tested to demonstrate that integrated patterns are much more effective and are comparable with several state-of-the-art approaches on the MPEG7 dataset. Then, the proposed approach is tested on the TARI-1000 dataset, to reassure the effectiveness of the proposed approach. Finally, we validate the robustness of the proposed approach to rotation.

For evaluation we use the widely used criterion named the Bull's eye retrieval rate (BER), which is the ratio of the total number of correct matches to the maximum number of correct matches. Specifically, each image is used as a query, and the number of images belonging to the same class is counted in the 40 images that are most similar to the query one. In all experiments, a five-fold cross validation is conducted: the shapes of each class are orderly divided into five separate splits, and for each fold one split is used for validation and the rest for training. The SFS is applied on training subsets to choose scales for each pattern. Then samples from each validation subset are tested for BER on the whole dataset using the scales selected from the corresponding training subset. The final BER is the average BERs over all samples.

For constructing the proposed shape descriptors, 200 points are uniformly sampled along each shape contor, and the number of possible scales of AP and BAPmP are determined accordingly. There are 99 scales of AP, 99 scales of BAP2P, 49 scales of BAP4P, 33 scales of BAP6P, 24 scales of BAP8P, 19 scales of BAP10P and 16 scales of BAP12P. In total, there are 339 scales of AP and BAPmP. The number of bins of the histogram of AP is experimentally fixed to 24.

A. Experiments on MPEG-7 Dataset

The proposed method is firstly tested on the widely used MPEG-7 dataset (the MPEG-7 CE-Shape-1 Part B). The dataset contains 70 shape classes, and each class has 20 different shapes. Thus, this dataset contains 1400 samples in total. Since there are 20 shapes in one class, the maximum number of correct matches for a single query image is 20, and the total number of correct matches is $1400 \times 20 = 28000$. Notice that for each training subset using five-fold cross validation there are only 16 shapes in one class, the maximum number of correct matches for a single query image is 16, and the total number of correct matches is $1400 \times 16 = 22400$. Some examples of this dataset are shown in Fig. 9.

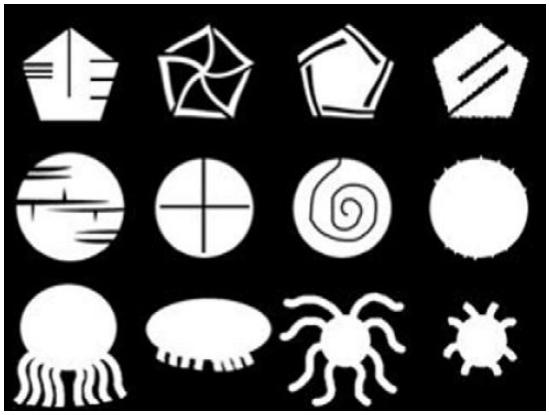


Fig. 9. Some examples of MPEG-7 dataset.

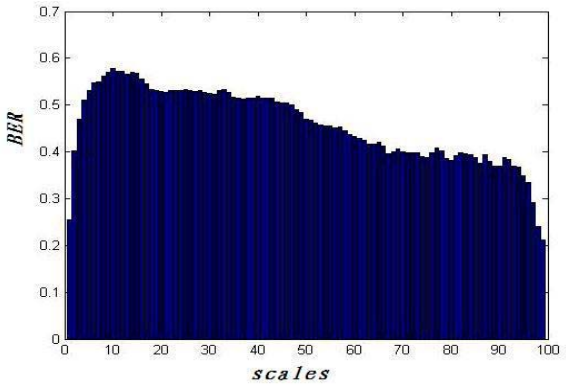


Fig. 10. BERs of 99 scales of AP.

1) Experiments of Single-Scale AP and BAP: We first the effectiveness of each scale of AP and BAP. In Fig. 10, the BER of each scale of AP is given, and the best performance of 57.90% is achieved when $s = 10$. In Fig. 11, the performances of BAPmP features are given. The BERs of all scales of BAP2P, BAP4P, BAP6P, BAP8P, BAP10P and BAP12P are shown in Fig. 11(a-f), showing that the performances of single-scale BAPmPs are still unsatisfactory. BAP2P can only achieve an average BER around 20%, and BAP4P a bit higher than 40%. Other BAPmP features can achieve more acceptable BERs around 60% at most scales.

The highest BER, 65.70%, of all APs and BAPs is achieved by BAP10P when $s = 10$. In all AP and BAP the scale of $s = 10$ provides the top-level performance. Consequently, the scale $s = 10$ is regarded as the best scale for AP and BAP. From the experimental results, it can be seen that the single-scale AP and BAP fail to provide satisfactory performances.

As mentioned previously, a histogram of AP has 24 bins, and histograms of BAP2P and BAP4P only have 4 bins and 16 bins, respectively. For there are 200 points to be distributed to these bins, which is large enough to make the histograms robust. How are the situations for BAP6P, BAP8P and BAP10P? Here we give some examples extracted from shapes in the dataset. As shown in Fig. 12, a histogram of BAP6P is well distributed, and there are about 10 occurrences for noticeable bins. While for BAP8P, there are 5 to 10 occurrences for noticeable bins. As for BAP10P, there are about 5 occurrences for noticeable bins, which is still robust to some extent. The situation of BAP12P is not shown here, since the number of occurrences for noticeable bins will drop to less than 3, making the descriptor unstable for shape matching. It can be seen that the patterns of straight zeros and the patterns of straight ones both have more occurrences in all cases, which means that these patterns are more common for all shapes.

2) Experiments of Multi-Scale AP and BAP: To demonstrate the effectiveness of multi-scale representation, an integration of AP and BAPmP of their own scales is firstly tested. The BERs are reported in TABLE I. It can be seen that after integration all patterns achieve significantly improved BERs, and AP is the best one among them. Among all BAPmP features, BAP6P and BAP8P achieve the top-level performances. The performances of BAP2P and BAP4P are relatively weak due to the low dissimilarity of their histograms, while BAP10P and BAP12P suffer from the instability due to too many bins in their histograms.

Secondly, the integration of AP with one BAPmP is tested to determine the supplementation of each BAP pattern to AP. The BERs are shown in Table II. It can be seen that all integrated patterns achieve more satisfactory BERs than using an individual pattern, and BAP8P demonstrates its best supplementation to AP with the highest BER.

The third experiment is to test the performance of the integration of some subsets from all 339 scales of AP and BAPmP features. The results are shown in TABLE III. Symbol "$\surd$" and "$\times$" indicate if a certain pattern is included in the candidate set or not, respectively. The best performance is achieved when BAP10P and BAP12P are excluded, and the second best performance is achieved by using all patterns. It can also be seen that AP+BAP6P+BAP8P achieve a BER only 0.01% lower than that the combination with additional BAP10P and BAP12P. These observations confirm that though BAP10P and BAP12P are discriminative they are not very robust. So these two patterns can be neglected for real application. On the contrary, though BAP2P and BAP4P are not very discriminative as TABLE I shows, they are, however, robust and provide good supplementation to other patterns, as demonstrated by the results from the third to the fifth rows in TABLE III.

In Fig. 13(a), the top three performances in TABLE III are described in detail, where the BERs versus the number of selected scales are plotted. It can be seen that the BERs reach 80% with less than 5 scales and 85% with less than 10 scales. After that, steady but slow growths are observed. For the best performance, the first 5 selected scales for each training subset all include 3 scales of AP, which are shown in Fig. 13(b). The specific scales selected from different training subsets are very similar. Considering that adjacent scales have similar retrieval performances, as shown in Figures 11 and 12, we conclude that the scales selected by SFS are stable to a small range and AP plays the key role among all patterns.

3) Comparison With Existing Methods: In this subsection, the proposed method is compared with some existing methods.

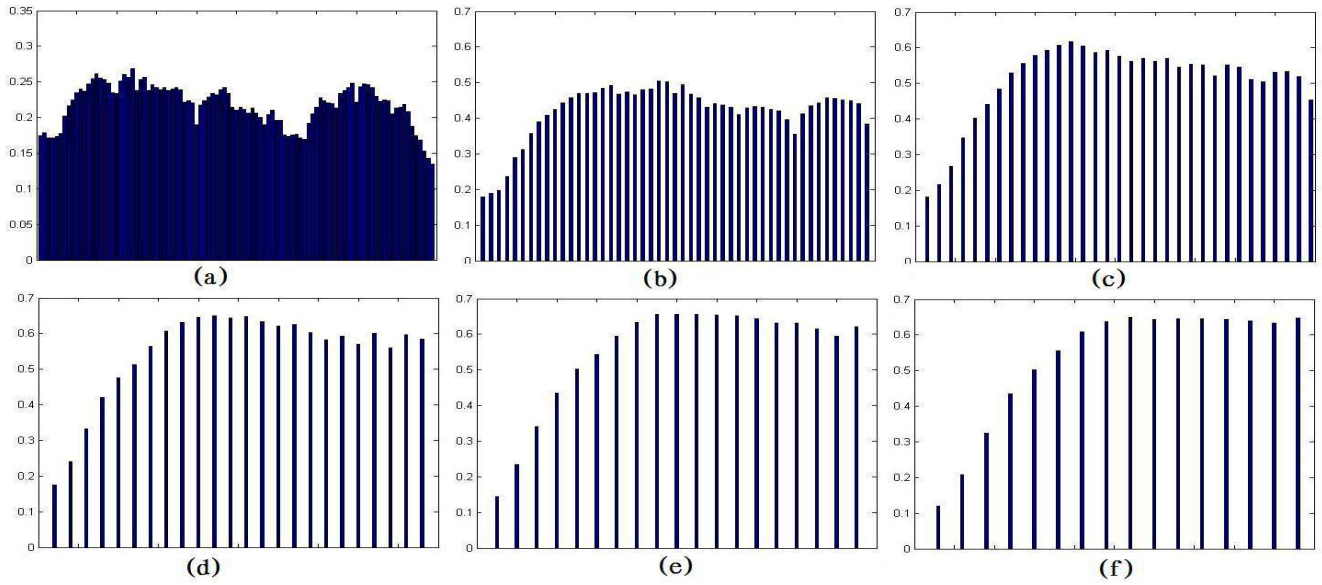


Fig. 11. BERs of all scales of BAP: (a), (b), (c), (d), (e) and (f) correspond to BAP2P, BAP4P, BAP6P, BAP8P, BAP10P and BAP12P, respectively.

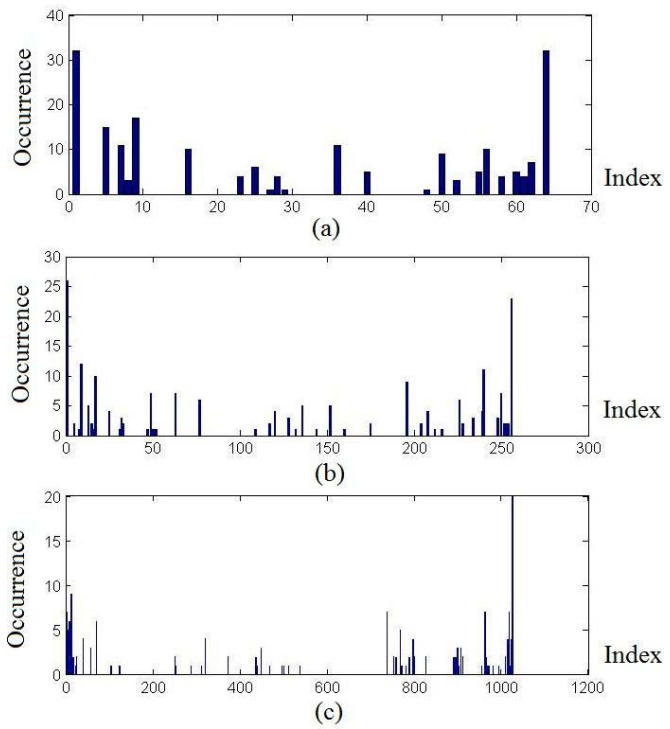


Fig. 12. Examples of histogram of BAP*m*P: (a) BAP6P, (b) BAP8P, (c) BAP10P.

TABLE I

BERs USING AP AND BAP*m*Ps ALONE ON MPEG-7

Pattern	AP	BAP2P	4P	6P	8P	10P	12P
BER (%)	**81.20**	70.20	74.70	78.26	78.55	76.31	74.34

TABLE II

BERs FOR INTEGRATING AP AND ONE BAP*m*P ON MPEG-7

Pattern AP +	BAP2P	4P	6P	8P	10P	12P
BER (%)	84.00	85.40	85.59	**86.53**	86.22	86.15

TABLE III

BERs *s* FOR INTEGRATING SUBSET OF ALL 339 SCALES ON MPEG-7

AP	BAP2P	4P	6P	8P	10P	12P	BER (%)
√	√	√	√	√	×	×	**87.04**
√	√	√	√	√	√	×	86.88
√	√	√	√	√	√	√	87.03
√	×	√	√	√	√	√	86.85
√	×	×	√	√	√	√	86.65
√	×	×	√	√	×	×	86.64

TABLE IV

COMPARISON WITH EXISTING METHODS ON MPEG-7

	IDSC[1]	ST[2]	HF[23]	CSS[3]	PMEM[9]	Ours
BER (%)	85.40	87.70	**90.35**	81.12	84.33	87.04
Category	Local	Local	Local	Global	Global	Global

The performances and categories of the methods are shown in TABLE IV. It can be seen that HF has achieved the highest BER so far, and PMEM has achieved the highest BER among all previously proposed global methods. The proposed method achieves a comparable BER as ST and is higher than PMEM. It is worth mentioning that PMEM needs to use the circular shifting procedure to acquire invariance to rotation, so it is less efficient than the proposed method.

To demonstrate the drastic difference in efficiency between global and local methods, we compare the time of applying DP in HF and IDSC with the time of computing χ^2 distance in the proposed method, both implemented in C and can be downloaded from the authors' homepages. DP in HF and IDSC costs about 116 ms and 35 ms, respectively, while the proposed

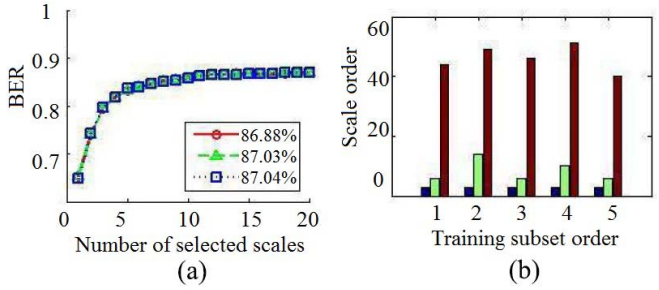


Fig. 13. (a) BERs versus the number of selected scales of top 3 performances; (b) The three selected scales of AP pattern in the first five selected scales from each training subset of the best performance.

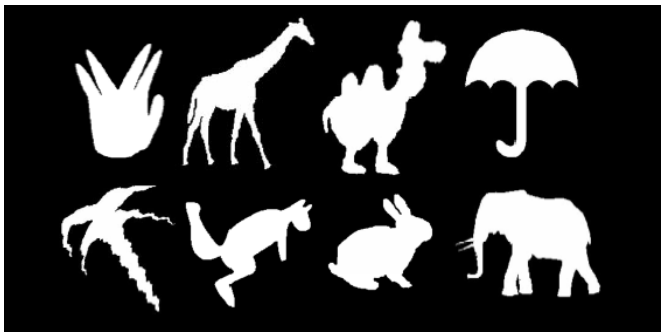


Fig. 14. Some examples of TARI-1000 dataset.

TABLE V
BERs FOR INTEGRATING AP AND ONE BAP*m*P ON TARI-1000

Pattern AP +	BAP2P	4P	6P	8P	10P	12P
BER (%)	90.13	91.37	**92.46**	92.17	91.80	91.48

method costs only about 0.05 ms, which shows the efficiency of the proposed method.

B. Experiments on TARI-1000 Dataset

The TARI-1000 dataset contains 1,000 silhouette images grouped into 50 classes with 20 images per class. So the total number of correct matches is $1000 \times 20 = 20000$, and for each training subset using 5-fold cross validation the total number of correct matches is $1000 \times 16 = 16000$. The dataset is designed to have more articulation changes within each class than the MPEG-7 dataset, and consequently IDSC achieves better results on this dataset [26]. Examples of this dataset are shown in Fig. 14.

Firstly, the integration of AP with one BAP*m*P is tested to study the supplementation of each BAP pattern to AP. The BERs are shown in the column labeled 'Original' in Table V. It can be seen that BAP6P demonstrate the best supplementation to AP with the highest BER, while BAP8P performs the second best. This is a bit different from the results on the MPEG-7 dataset, but still BAP6P and BAP8P achieve the top-level performances due to their strong discriminability and robustness.

The second experiment is to test the performance of the integration of some subsets from all 339 scales of AP and BAP*m*P.

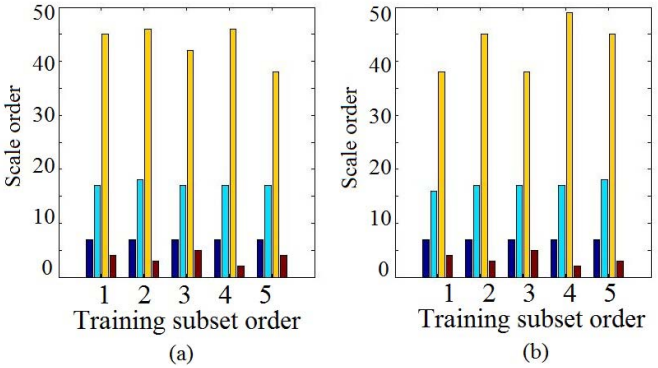


Fig. 15. The four selected scales of AP pattern in the first five selected scales from each training subset of (a) the best performance and (b) the worst performance.

The results are shown in TABLE VI. The best performance is achieved when all patterns are used while AP+BAP6P+BAP8P achieves the worst one, which is mainly in accordance with the results on the MPEG-7 dataset. It is worth noting that, however, BAP10P and BAP12P seem to play a more important role on this dataset, and without them the performances drop significantly. Since TRAI-1000 has more articulation changes, it is reasonable to conclude that BAP10P and BAP12P are more useful for comparing articulated shapes due to their higher discriminability.

For the best performance of using all patterns and the worst performance of using AP+BAP6P+BAP8P, the first 4 of the 5 selected scales for each training subset are shown in Fig. 15(a) and (b). It can be seen that the specific scales selected from different training subsets are still very similar. This confirms again that the stability of the scale selection by SFS.

Next, the proposed method is compared with some existing methods. No global methods have reported results on this newly extended dataset, so the proposed method is mainly compared with some reported local methods. In addition, we conduct experiment of a classic global method, Fourier shape Descriptor (FD), and all the performances and categories of the methods are shown in TABLE VII. It can be seen that IDSC achieves the highest BER because its insensitivity to articulation. The proposed method is able to achieve a comparable result between SC and a recently proposed local method in [28].

Finally, to validate that AP and BAP are robust to rotation, scale and noise, we conducted experiments on the rotated, scaled and noisy versions of the TARI-1000 dataset, respectively. The images are rotated by $\pi/8$, $\pi/6$, $\pi/4$, $\pi/3$ and $3\pi/8$, scaled to half and double sizes, and all cases are implemented with bilinear interpolation. The noises are randomly added to the coordinates of contor points within 0.5%, 1%, 2%, and 5% of the width and length of the shape. The BER performances of all these cases are listed in Table VI. It can be seen that the proposed features are very robust to the rotation and scale. To our surprise, the retrieval performances on the rotated dataset are better, which may due to the smooth effect of the bilinear interpolation. For the noisy cases, random noise of 5% clearly degrades the performances, while in other cases the proposed features are still robust.

TABLE VI
BERs s FOR INTEGRATING SUBSET OF ALL 339 SCALES ON TARI-1000

AP	2P	4P	6P	8P	10P	12P	Original	Rotation					Scale		Noise			
								$\pi/8$	$\pi/6$	$\pi/4$	$\pi/3$	$3\pi/8$	0.5	2	0.5%	1%	2%	5%
√	√	√	√	√	×	×	92.42	92.45	92.81	92.78	92.56	92.54	92.66	92.48	92.43	92.58	92.30	91.31
√	√	√	√	√	√	×	92.45	92.62	**92.85**	92.86	92.65	92.60	**92.84**	92.53	92.45	92.68	92.32	**91.34**
√	√	√	√	√	√	√	**92.62**	**92.71**	**92.85**	**92.87**	**93.02**	**92.87**	**92.84**	**92.76**	**92.52**	**92.71**	**92.37**	**91.34**
√	×	√	√	√	√	√	92.46	92.65	92.73	92.78	92.87	92.66	92.81	92.68	92.43	92.66	92.36	**91.34**
√	×	×	√	√	√	√	92.43	92.61	92.71	92.78	92.80	92.66	92.76	92.65	92.32	92.64	92.34	91.32
√	×	×	√	√	×	×	92.35	92.42	92.69	92.76	92.54	92.45	92.64	92.48	92.25	92.53	92.30	91.30

TABLE VII
COMPARISON WITH EXISTING METHODS ON TARI-1000

	COP [28]	SC [26]	IDSC [26]	FD[20]	Ours
BER (%)	92.18	94.18	**95.33**	75.67	93.02
Category	Local	Local	Local	Global	Global

IV. CONCLUSION

In this paper, we proposed a global shape descriptor for shape retrieval. The descriptor is based on the multi-scale integration of two angular features, namely angular patterns (AP) and binary angular patterns (BAP), both are invariant to scale and rotation. To integrate different scales of AP and BAP, the z-score normalization is applied for distance matrices computed at each scale, and the normalized distance matrices are summed together to create a new distance matrix for shape retrieval. The Sequential Forward Selection (SFS) scheme is used to determine a set of scales that effectively capture discriminative information. The experimental results on the MPEG-7 and TARI-1000 datasets demonstrate the effectiveness and efficiency of the proposed method.

In the future, we are interested in further improving the proposed approach in several ways. First, different integration methods may be applied, such as a different normalization procedure can be used instead of the z-score. Second, a more elaborate scale selection method other than the SFS may find better set of scales of AP and BAP. Third, a cross-bin histogram distance can be used instead of the bin-to-bin χ^2 distance to measure the similarity between shapes more robustly.

REFERENCES

[1] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 4, pp. 509–522, Apr. 2002.

[2] P. F. Felzenszwalb and J. D. Schwartz, "Hierarchical matching of deformable shapes," in *Proc. IEEE CVPR*, Jun. 2007, pp. 1–8.

[3] F. Mokhtarian, S. Abbasi, and J. Kittler, "Efficient and robust retrieval by shape content through curvature scale space," in *Proc. IDMS*, 1997, pp. 51–58.

[4] L. J. Latecki and R. Lakamper, "Shape similarity measure based on correspondence of visual parts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 10, pp. 1185–1190, Oct. 2000.

[5] H. Ling and D. W. Jacobs, "Shape classification using the inner-distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 2, pp. 286–299, Feb. 2007.

[6] T. B. Sebastian, P. N. Klein, and B. B. Kimia, "On aligning curves," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 1, pp. 116–125, Jan. 2003.

[7] T. Adamek and N. E. O'connor, "A multiscale representation method for nonrigid shapes with a single closed contor," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 5, pp. 742–753, May 2004.

[8] B. J. Super, "Learning chance probability functions for shape retrieval or classification," in *Proc. IEEE Conf. CVPR*, Jun. 2004, pp. 93–96.

[9] E. Attalia and P. Siy, "Robust shape similarity retrieval based on contor segmentation polygonal multiresolution and elastic matching," *Pattern Recognit.*, vol. 38, no. 12, pp. 2229–2241, Dec. 2005.

[10] G. McNeill and S. Vijayakumar, "Hierarchical procrustes matching for shape retrieval," in *Proc. IEEE Conf. CVPR*, Jun. 2006, pp. 885–894.

[11] C. Scott and R. Nowak, "Robust contor matching via the order-preserving assignment problem," *IEEE Trans. Image Process.*, vol. 15, no. 7, pp. 1831–1838, Jul. 2006.

[12] B. J. Super, "Retrieval from shape databases using chance probability functions and fixed correspondence," *Pattern Recognit. Artif. Intell.*, vol. 20, no. 8, pp. 1117–1138, 2006.

[13] N. Alajlan, M. S. Kamel, and G. H. Freeman, "Geometry-based image retrieval in binary image databases," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 6, pp. 1003–1013, Jun. 2008.

[14] M. R. Daliri and V. Torre, "Robust symbolic representation for shape recognition and retrieval," *Pattern Recognit.*, vol. 41, no. 5, pp. 208–220, May 2008.

[15] L. Lin, K. Zeng, X. Liu, and S. C. Zhu, "Layered graph matching by composite cluster sampling with collaborative and competitive interactions," in *Proc. IEEE Conf. CVPR*, Jun. 2009, pp. 1351–1358.

[16] C. Xu, J. Liu, and X. Tang, "2D shape matching by contor flexibility," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 1, pp. 180–186, Jan. 2009.

[17] H. Ling, X. Yang, and L. J. Latecki, "Balancing deformability and discriminability for shape matching," in *Proc. Eur. Conf. Comput. Vis.*, 2010, pp. 411–424.

[18] K. Nasreddine, A. Benzinou, and R. Fablet, "Variational shape matching for shape classification and retrieval," *Pattern Recognit. Lett.*, vol. 31, no. 12, pp. 1650–1657, Sep. 2010.

[19] X. Shu and X.-J. Wu, "A novel contor descriptor for 2D shape matching and its application to image retrieval," *Image Vis. Comput.*, vol. 29, no. 4, pp. 286–294, Mar. 2011.

[20] D. Zhang, "Review of shape representation and description techniques," *Pattern Recognit.*, vol. 34, no. 1, pp. 1–19, Jan. 2004.

[21] L. J. Latecki, R. Lakamper, and U. Eckhardt, "Shape descriptors for non-rigid shapes with a single closed contor," in *Proc. IEEE Conf. CVPR*, Jun. 2000, pp. 424–429.

[22] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.

[23] J. Wang, X. Bai, X. You, W. Liu, and L. J. Latecki, "Shape matching and classification using height functions," *Pattern Recognit. Lett.*, vol. 33, no. 2, pp. 134–143, Jan. 2012.

[24] X. Yang, S. Koknar-Tezel, and L. J. Latecki, "Locally constrained diffusion process on locally densified distance spaces with applications to shape retrieval," in *Proc. IEEE Conf. CVPR*, Jun. 2009, pp. 357–364.

[25] X. Bai, X. Yang, L. J. Latecki, W. Liu, and Z. Tu, "Learning context sensitive shape similarity by graph transduction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 5, pp. 861–874, May 2010.

[26] X. Bai, B. Wang, and X. Wang, "Co-transduction for shape retrieval," in *Proc. Eur. Conf. Comput. Vis.*, 2010, pp. 328–341.

[27] R. Hu, W. Jia, Y. Zhao, and J. Gui, "Perceptually motivated morphological strategies for shape retrieval," *Pattern Recognit.*, vol. 45, no. 9, pp. 3222–3230, Sep. 2012.

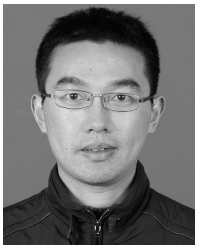
[28] Y. Zhou, J. Wang, Q. Zhou, X. Bai, and W. Liu, "Shape matching using points co-occurrence pattern," in *Proc. ICIG*, Aug. 2011, pp. 344–349.
[29] H. Ling and K. Okada, "An efficient Earth mover's distance algorithm for robust histogram comparison," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 5, pp. 840–853, May 2007.

Rong-Xiang Hu received the B.Sc. degree in computer science from the Hefei University of Technology, Hefei, China, in 2004, and the Ph.D. degree in pattern recognition and intelligence system from the University of Science and Technology of China, Hefei, in 2012. He is currently an Assistant Professor with the Institute of Nuclear Energy Safety Technology, Chinese Academy of Sciences, Beijing, China. His research interests include pattern recognition, machine learning, and image processing.

Wei Jia received the B.Sc. degree in informatics from Central China Normal University, Wuhan, China, in 1998, the M.Sc. degree in computer science from the Hefei University of Technology, Hefei, China, in 2004, and the Ph.D. degree in pattern recognition and intelligence system from the University of Science and Technology of China, Hefei, in 2008. He is currently an Associate Professor with the Institute of Nuclear Energy Safety Technology, Chinese Academy of Sciences. His research interests include biometrics, pattern recognition, and image processing.

Haibin Ling received the B.S. degree in mathematics and the M.S. degree in computer science from Peking University, Beijing, China, in 1997 and 2000, respectively, and the Ph.D. degree in computer science from the University of Maryland, College Park, MD, USA, in 2006. From 2000 to 2001, he was an Assistant Researcher with the Multimodel User Interface Group, Microsoft Research Asia. From 2006 to 2007, he was a post-doctoral scientist at the University of California Los Angeles. He joined Siemens Corporate Research as a Research Scientist. Since 2008, he has been an Assistant Professor with Temple University. His research interests include computer vision, medical image analysis, human computer interaction, and machine learning. He received the Best Student Paper Award at the ACM Symposium on User Interface Software and Technology in 2003.

Yang Zhao received the B.E. degree from the Department of Automation, University of Science and Technology of China, in 2008, and the Ph.D. degree from the Department of Automation, University of Science and Technology of China, in 2013. From September 2013, he is a post-doctoral researcher with the School of Electronic and Computer Engineering, Peking University, China. His research interests include pattern recognition and image processing.



Jie Gui received the B.Sc. degree in computer science from Hohai University, Nanjing, China, in 2004, the M.Sc. degree in computer applied technology from the Anhui Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Hefei, China, in 2007, and the Ph.D. degree in pattern recognition and intelligent systems from the University of Science and Technology of China, Hefei, in 2010. Since 2010, he has been an Assistant Professor with the Hefei Institute of Intelligent Machines, Chinese Academy of Sciences. His research interests include machine learning, pattern recognition, and image processing.