2D Shape Matching by Contour Flexibility

Chunjing Xu, Jianzhuang Liu, *Senior Member*, *IEEE*, and Xiaoou Tang, *Senior Member*, *IEEE*

Abstract—In computer vision, shape matching is a challenging problem, especially when articulation and deformation of parts occur. These variations may be insignificant for human recognition but often cause a matching algorithm to give results that are inconsistent with our perception. In this paper, we propose a novel shape descriptor of planar contours, called contour flexibility, which represents the deformable potential at each point along a contour. With this descriptor, the local and global features can be obtained from the contour. We then present a shape matching scheme based on the features obtained. Experiments with comparisons to recently published algorithms show that our algorithm performs best.

Index Terms—2D shape, contour flexibility, matching.

1 Introduction

The shape of an object is very important in object recognition. Using the shape of an object for object recognition and image understanding is a growing topic in computer vision and multimedia processing and finding good shape descriptors and matching measures is the central issue in these applications.

Analyzing the silhouette of an object is the first and most important step in shape matching. Based on the silhouette of objects, a variety of shape descriptors and matching methods have been proposed in the literature. In these papers, a silhouette is simplified as a curve (contour) represented by a sequence of landmarks.

Some simple and elegant methods were developed by Kendall [16], Bookstein [7], and Mardia and Dryden [20]. They introduced the shape represented by 2D landmarks and investigated the shape space (which is a Riemannian manifold) and the metric on it. Klassen et al. [17] presented a new differential geometric shape representation by employing the direction and curvature functions of curves. In [10], [25], and [8], moment-based features are used for shape analysis. For these *math-flavored* methods, which focus on global features, the major limitation is that the descriptors are sensitive to local changes. For example, the left two pairs of contours in Fig. 1 may cause mismatching by these methods. In these cases, the significant deformation of the contours introduces adverse information to the global shape descriptors.

To overcome the problems in the global descriptors, contour local features and nonlinear methods are introduced. In [13], Hoffman and Richards noticed that the parts of an object play a key role in recognition. The psychological evidence they showed leads to a decomposition scheme in which the boundary curve of an object is partitioned into parts at negative curvature minima. From the decomposition, a *codon*-based shape representation is proposed. With a similar principle, Siddiqi and Kimia [30] presented a shape matching scheme based on the decomposition of silhouettes. Recently, a similar hierarchical segment-based shape matching algorithm was proposed by McNeill and Vijayakumar in [22]. Other than the representations with the segments of a contour, a skeleton-based method is given in [31] and [29] where the shapes are characterized by shock graphs.

Noticing that contour curvature is one of the intrinsic features to the visual perception of the shape of an object [6], [3], Sebastian et al. [28] proposed a method that aligns curves with length and curvature. Ling and Jacobs [19] developed an inner distance to define an articulation-insensitive descriptor. In [1], the local features like concaveness and convexedness are used to build the representation of a shape. A representation of a shape is given in [14], which focuses on the local structure that is preserved in deformations. To obtain a multiscale representation of a shape, wavelets [9] and Fourier descriptors [18] are used to build hierarchical structures to describe the coarse-to-fine details of an object.

However, the methods focusing on local features suffer more from noise. An example is shown in the rightmost pair in Fig. 1. Therefore, when we design a matching algorithm, we should find not only a good descriptor but also a proper trade-off between the global and local features. In this paper, we propose a new descriptor for closed curves, named *contour flexibility*, which depicts the deformable potential at each point along a curve, and show that both local and global features can be extracted by this descriptor. Based on the obtained features, a scheme is presented for shape matching.

The rest of this paper is organized as follows: In Section 2, we define the contour flexibility and determine an important parameter for it. In Section 3, we discuss how to extract local and global features from a contour with the contour flexibility. A shape matching scheme based on the obtained features is also presented. Section 4 describes the implementation of the contour flexibility and how to use it to obtain the landmarks from a contour. Section 5 gives the experimental results. Finally, we draw conclusions in Section 6.

2 Contour Flexibility

As usual, we assume that the contour of a 2D object is a simple closed curve and the area enclosed by the curve is topologically homeomorphic to a disk. We also assume that the centroid of the curve has been moved to the origin of the 2D coordinate system.

Given a curve $\mathcal{L}$, define two functions of distance transform κ_+ and κ_- on $\mathbb{R}^2$ as

$$\kappa_+(x) = d(x, \mathbb{R}^2 \setminus D), \quad x \in \mathbb{R}^2, \tag{1}$$

$$\kappa_-(x) = d(x, D), \qquad x \in \mathbb{R}^2, \tag{2}$$

where $d(\cdot,\cdot)$ is the euclidean distance between two sets and D is the domain bounded by $\mathcal{L}$. When x is outside D, $\kappa_+(x) = 0$ and $\kappa_-(x) > 0$; when x is inside D, $\kappa_-(x) = 0$ and $\kappa_+(x) \geq 0$. With κ_+ and κ_-, we have the following definition of the contour flexibility:

Definition 1. *For a point p on a contour $\mathcal{L}$ and a given radius r, the interior flexibility ω_+ and exterior flexibility ω_- at p are defined as*

$$\omega_+(p,r) = \frac{\int_{C^+_{p,r}} \kappa_+(x)dx}{\int_{C^+_{p,r}} dx}, \tag{3}$$

$$\omega_-(p,r) = \frac{\int_{C^-_{p,r}} \kappa_-(x)dx}{\int_{C^-_{p,r}} dx}, \tag{4}$$

where $C^+_{p,r}$ and $C^-_{p,r}$ are the connected components containing p in the sets $\{x \in D \,|\, \|x-p\| \leq r\}$ and $\{x \in \mathbb{R}^2 \setminus int(D) \,|\, \|x-p\| \leq r\}$, respectively, and $int(\cdot)$ is the interior of a set. The contour flexibility ω at p is defined as

$$\omega(p,r) = \min(\omega_+(p,r), \omega_-(p,r)), \tag{5}$$

where r is called the bendable size.

• *The authors are with the Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong. E-mail: {cjxu6, jzliu, xtang}@ie.cuhk.edu.hk.*

Manuscript received 24 July 2007; revised 17 Apr. 2008; accepted 21 July 2008; published online 29 July 2008.
Recommended for acceptance by S. Carlsson.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-2007-07-0449.
Digital Object Identifier no. 10.1109/TPAMI.2008.199.

Fig. 1. Three examples that may cause mismatching.

Three examples of $C^+_{p,r}$ and $C^-_{p,r}$ are given in Fig. 2. For a well-tuned bendable size r, $\omega(p,r)$ provides the information about how extensively the neighborhood of p is connected to the main body and about the deformation tolerance of an object at p. It can be observed that the smaller the $\omega(p,r)$ is, the more flexible (bendable) the neighborhood of p is. Fig. 3a shows one contour superimposed by a number of "+"; the density of "+" denotes the values of $\omega(p,r)$ along the contour. The segments marked by sparse "+" are composed of the points with small $\omega(p,r)$.

Now, we discuss how to find a proper bendable size r. In order to obtain a good descriptor of a shape which can extract more information from the contour, the difference between the values of flexibility at limb-like parts and the main body should be large. When r is very large or smaller than the size of the limb-like parts, $\omega(p,r)$ tends to be the same everywhere, which cannot reflect the flexibility of different segments of the contour. Our experiments show that, when r is taken as the average width of the limb-like parts of an object, the difference is statistically large enough and, therefore, $\omega(p,r)$ can well describe the flexibility of the contour. Next, we develop a scheme to find the limb-like parts of an object.

Definition 2. *Let $z(t) = (x(t), y(t))$, $0 \le t < 1$, be the arc-length parameterization of a contour $\mathcal{L}$. For any two points $p_1 = z(t_1)$ and $p_2 = z(t_2)$ on $\mathcal{L}$, the enclosure ratio $\epsilon(p_1,p_2)$ or $\epsilon'(t_1,t_2)$ is defined as*

$$\begin{aligned}\epsilon(p_1,p_2) &= \epsilon'(t_1,t_2)\\ &= \begin{cases} \frac{\|z(t_1)-z(t_2)\|}{\min(|t_1-t_2|,1-|t_1-t_2|)} & \text{if } t_1 < t_2\\ 1 & \text{otherwise.}\end{cases}\end{aligned} \tag{6}$$

It is not difficult to prove that $\epsilon(p_1,p_2)$ is a continuous function if $z(t)$ is smooth. The definition of $\epsilon(p_1,p_2)$ has the property that the local minima of $\epsilon(p_1,p_2)$ often correspond to limb-like parts. For example, all of the 15 local minima[1] of $\epsilon(p_1,p_2)$ are shown on the contour in Fig. 3b, where the pairs labeled with 1, 2, 5, 8, and 11 are the places we should use to define a proper bendable size r. In Fig. 3b, we also observe that not every local minimum corresponds to a limb-like part. Thus, we need to develop a method to eliminate the pairs we do not want to use.

Definition 3. *Given a set $\mathcal{S}$ of pairs that are the local minima of $\epsilon'(t_1,t_2)$, a pair $(t_a,t_b) \in S$ is called a prime pair when there is no another pair $(t_i,t_j) \in S$ such that $t_a \le t_i \le t_j \le t_b$ and $\mathcal{L} \bigcap l_{a,b} = \{z(t_a), z(t_b)\}$, where $l_{a,b}$ is the straight line connecting $z(t_a)$ and $z(t_b)$.*

In this definition, the condition $\mathcal{L} \bigcap l_{a,b} = \{z(t_a), z(t_b)\}$ indicates that the line $l_{a,b}$ does not intersect with the contour except at $z(t_a)$ and $z(t_b)$. We only use the prime pairs to compute the bendable size r. Here, we define the prime pair according to our intuition on the limb-like parts of an object, which has also been supported by our extensive experiments. For example, after removing the nonprime pairs in Fig. 3b, the prime pairs for the contour are shown in Fig. 3c, where all of the pairs except the one labeled with 10 denote the widths of the limb-like parts.

Definition 4. *Given the set of prime pairs of $\epsilon'(t_1,t_2)$ for a contour, $\mathcal{S} = \{(t_1^k, t_2^k) | k = 1,2,\ldots,M\}$, the bendable size r is defined as*

1. These local minima are computed from the samples of the contour.

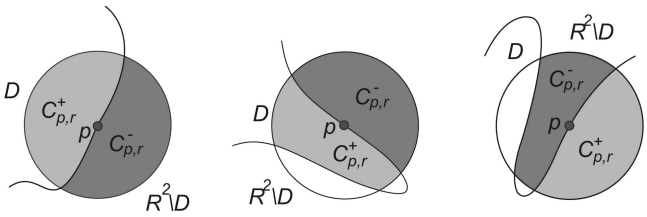


Fig. 2. Three examples of $C^+_{p,r}$ and $C^-_{p,r}$.

$$r = \frac{1}{M}\sum_{k=1}^{M} \|z(t_1^k) - z(t_2^k)\|. \tag{7}$$

Some points are worth mentioning here:

- A small change of r (compared to the size of the object) results in an insignificant change of $\omega(p,r)$ since $\omega(p,r)$ is defined as one of the two means of κ_- and κ_+ in a certain domain ($C^-_{p,r}$ or $C^+_{p,r}$).
- The bendable size r and the contour flexibility $\omega(p,r)$ are proportional to the scale of the contour. However, they are invariant to the translation, rotation, and the choice of starting point for the parameterization of the contour.

3 A Scheme for Shape Matching

A simple closed planar 2D curve can be approximated by a sequence of samples (landmarks) on it and the landmarks can be obtained by uniform or nonuniform sampling on the contour. We denote a sequence of landmarks from a contour by $\Phi = \{z^k | k = 1,2,\ldots,n\}$ in what follows.

3.1 Local Features with the Contour Flexibility

From the theories of Attneave [3] and Biederman [6], the valuable features for shape recognition most likely come from the corners (high curvature points) or the parts of an object where changes occur. For example, in Fig. 4a, from the marked key points, the shape of the dog can be somehow restored. In Fig. 4b, it can be observed that only local segments related to the corners are enough for a successful recognition of the object by humans. It is not surprising that the contour flexibility exhibits the capability to extract such interesting parts from the contour because, in most cases, changes correspond to articulated high flexible parts of an object.

Let Φ be a sequence of landmarks sampled uniformly from a curve and $\Omega = \{\omega^1,\ldots,\omega^n\}$ be the sequence composed of the values of the contour flexibility at each landmark of Φ, i.e.,

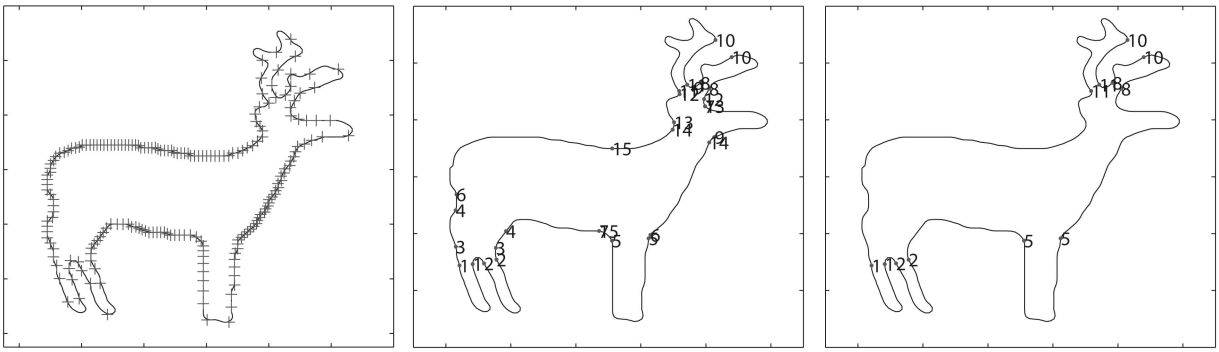
$$\omega^i = \omega(z^i), \quad i = 1,2,\ldots,n. \tag{8}$$

In Fig. 5, we can observe that the main local minima of Ω coincide with the interesting points. With this idea, the matching of the local features between two contours can be formulated as finding the matching of fluctuation trends between two sequences, which is given as follows:

Let $\Omega_1 = \{\omega_1^k\}$ and $\Omega_2 = \{\omega_2^k\}$ be two sequences of flexibility for two contours. As a classical method, dynamic time warping [27] is often used to match two 1D signals. The application of this technique to shape matching also can be found in [4]. The basic idea is to find a warping function $\eta(k) = (i(k), j(k))$ such that the following energy function is minimized:

$$E_\eta = \frac{1}{L_\eta}\sum_{k=1}^{n} d\left(\omega_1^{i(k)}, \omega_2^{j(k)}\right), \tag{9}$$

where $i(k)$ and $j(k)$ are monotonously increasing functions and L_η is the length of the warping path of η for normalization. The



(a) (b) (c)

Fig. 3. (a) The distributions of $\omega(p,r)$ on a contour. The denser the "+," the larger the $\omega(p,r)$. (b) Point pairs corresponding to the local minima of $\epsilon(p_1,p_2)$. (c) Prime pairs.

optimization of this problem can be solved by dynamic programming. The warping distance between Ω_1 and Ω_2 is defined as

$$D(\Omega_1,\Omega_2)=\min_{\eta} E_\eta. \tag{10}$$

In Figs. 6a, 6b, and 6c,[2] we show three gestures and their corresponding flexibility values (signals). The two parts between the lines denote the significant difference between signals in Figs. 6a and 6b, which comes from the different positions of the thumbs. We can observe that signals in Figs. 6a and 6b can be matched better than signals in Figs. 6a and 6c or Figs. 6b and 6c. This is consistent with the 2D shapes of gestures shown in the right of the signals.

The flexibility sequence can be a more useful tool to overcome the difficulty of matching deformed parts of objects. Compared with curvature, it is more stable to deformation. An example is given in Figs. 6f and 6g, where the two parts between the lines correspond to the tails of the rays. The part in Fig. 6g is longer than that in Fig. 6f because the second ray has a longer tail. The warping can well handle this situation and a matching between the two parts can be established. However, the curvature will suffer from this situation because its changes can be large when the tail bends.

3.2 Global Matching with Contour Flexibility

Local features can often reflect the shape of an object. However, the methods focusing on local features may fail to match objects of the same class when the contours have significant deformation or noise. For example, the first and the last pairs of images shown in Fig. 1 are hard to match by these methods. Therefore, it is unsuitable to enitrely rely on the matching on local features. In this case, the global shape can play a key role for matching these contours. In this section, we develop a method in favor of the global shapes of objects with the contour flexibility.

In what follows, the sequence Φ of the landmarks from a contour is also treated as a vector $\Phi=(z^1,z^2,\ldots,z^n)^T$. Each landmark is considered as a complex number $z^k=x^k+jy^k$. The vector is normalized to unit length. In addition, without loss of generality, Φ is assumed to satisfy $\sum_{k=0}^{n} z^k=0$.[3] Given two sequences Φ_1 and Φ_2, the Procrustean distance [16] between the two sequences

$$d(\Phi_1,\Phi_2)=\cos^{-1}(|\langle\Phi_1,\Phi_2\rangle|) \tag{11}$$

is used to measure the distance between the two sequences, where $\langle\cdot,\cdot\rangle$ denotes the inner product of two complex vectors.

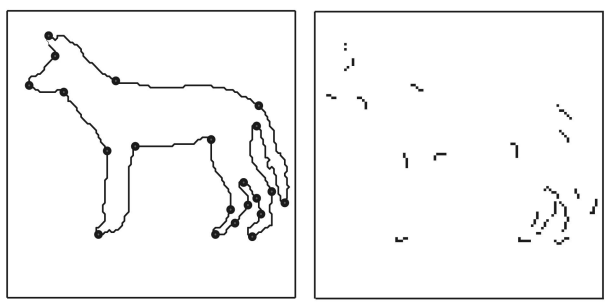
2. The gestures in the figure are from the GESTURE database, the dogs are from MPEG-7, and the rays are from the MARINE database. All three benchmark databases are described in Section 5.

3. With $z^k\leftarrow z^k-\frac{1}{n}\sum_{k=0}^{n} z^k$, we can always achieve this purpose.

The landmarks of a contour can be generated by uniform-length or nonuniform-length sampling. The traditional method is to sample the contour uniformly (i.e., to sample at a constant speed). This method gives the same weight to different landmarks for matching, which is not suitable for matching flexible parts of two contours. Our strategy is to give large weights to more inflexible landmarks and smaller weights to more flexible landmarks. This is equivalent to using more samples for the matching on the segments of a contour which are more inflexible. The contour flexibility $\omega(p,r)$ defined in (5) is suitable for this purpose. Let

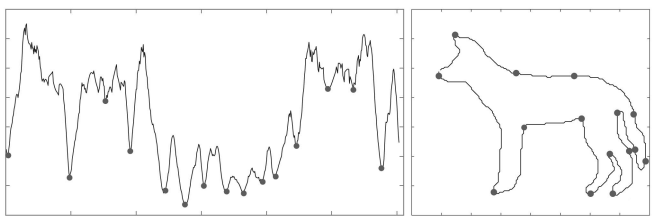
$$\gamma(t)=\frac{\int_0^t \omega(z(u),r)du}{\int_0^1 \omega(z(u),r)du}, \tag{12}$$

where r is the bendable size for the arc-length parameterization of a contour $z(t)$, $0\le t<1$. When we sample the contour with speed $d\gamma/dt$, we obtain a sequence that is dense along the inflexible segments and sparse along the flexible segments. A straightforward method is to select the set $\{z(\gamma(\frac{k}{n}))|k=1,2,\ldots,n\}$. Let $\bar{\Phi}_1$ and $\bar{\Phi}_2$ be two sequences obtained from two shapes by the nonuniform



(a) (b)

Fig. 4. (a) A dog with high curvature points marked. (b) The dog composed of only local segments on the high curvature points.



(a) (b)

Fig. 5. (a) The local minima of flexibility on the sequence Ω from the dog. (b) The local minima shown on the contour.

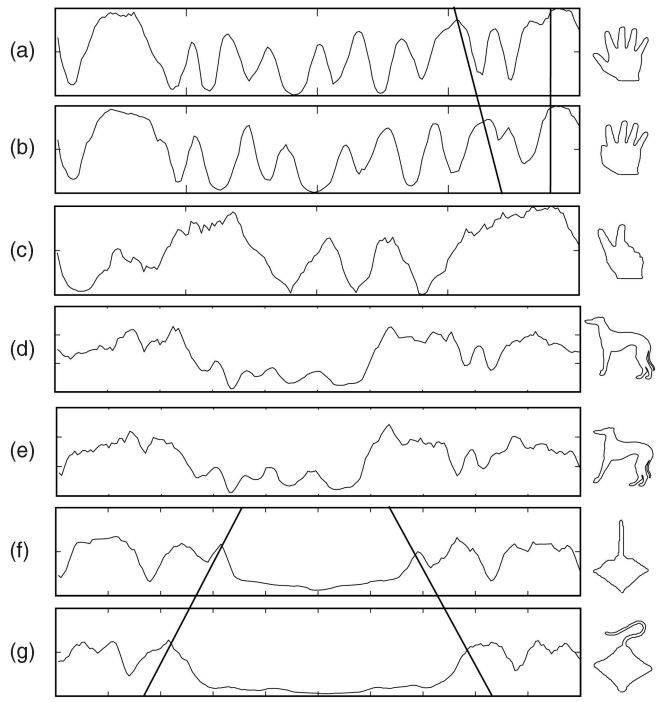


Fig. 6. The sequence of some shapes. (a) Gesture #0. (b) Gesture #9. (c) Gesture #11. (d) Dog #11. (e) Dog #11 after 20° skew. (f) Marine #802. (g) Marine #821.

sampling method discussed above. We use the Procrustean distance $d(\bar{\Phi}_1, \bar{\Phi}_2)$ between them for the global matching of the two shapes.

With the nonuniform sampling method, there are more landmarks obtained from the less deformable parts of an object (e.g., the body of an animal) and fewer landmarks on the parts that can be easily deformed (e.g., the limbs). Therefore, the main bodies of two objects will dominate the matching, and the matching contributed by limb-like parts is lessened. In Fig. 7, we show two examples of matching with both uniformly and nonuniformly sampled sequences. We can see that the matching with the uniformly sampled sequence is disturbed by the deformation of the articulated parts, while the matching with the nonuniformly sampled sequences is better.

3.3 Matching with Local and Global Features

In Section 3.2, we assume that the correspondence between the leading landmarks of two sequences $\bar{\Phi}_1$ and $\bar{\Phi}_2$ is known. Actually, it is determined by a cyclic permutation of one of the two sequences. This correspondence is obtained when the distance between $\bar{\Phi}_1$ and $\bar{\Phi}_2$ achieves the minimum, which is defined by

$$\hat{d}(\bar{\Phi}_1, \bar{\Phi}_2) = \cos^{-1}\left(\max_{s\in\mathbf{N}} \left|\left\langle \bar{\Phi}_1, \sigma^s(\bar{\Phi}_2)\right\rangle\right|\right), \tag{13}$$

where σ^s is a cyclic permutation[4] with offset s acting on the sequence and $\mathbf{N}$ is the set of integers. Let $\hat{s}$ be the offset of the cyclic permutation determined by (13). Then, $\hat{s}$ is also used to determine the leading points of two sequences of Ω_1 and Ω_2 in Section 3.1.

Given two sequences Φ_1 and Φ_2 with the same length n, obtained by uniform sampling on curves C_1 and C_2, our scheme for matching two shapes with the local and global features is given as follows:

1. Obtain nonuniform sampling $\bar{\Phi}_1$ and $\bar{\Phi}_2$ as described in Section 3.2.

4. For example, the cyclic permutation of a sequence $\{a\ b\ d\ e\ f\}$ with offset 2 is $\{d\ e\ f\ a\ b\}$.

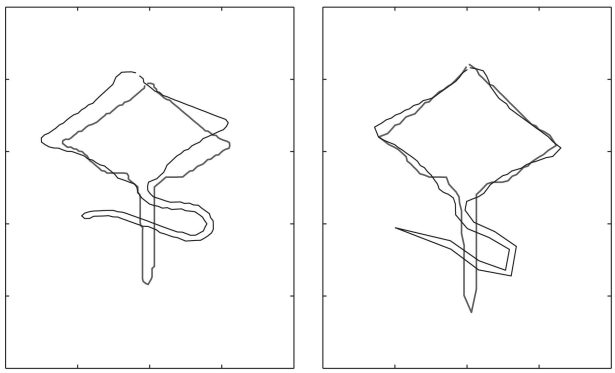


Fig. 7. (a) Global matching of two uniformly sampled sequences. (b) Global matching of two nonuniformly sampled sequences.

2. Determine $\hat{s}$ for finding leading landmarks of $\bar{\Phi}_1$ and $\bar{\Phi}_2$ and then adjust the sequence $\bar{\Phi}_2$ by $\bar{\Phi}_2 \leftarrow \sigma^{\hat{s}}(\bar{\Phi}_2)$.
3. Starting from the leading landmarks $\bar{z}_1^1 \in C_1$ and $\bar{z}_2^1 \in C_2$, uniformly sample C_1 and C_2 to obtain two sequences Ψ_1 and Ψ_2, from which the sequences of the contour flexibility Ω_1 and Ω_2 are obtained by (8).
4. The matching score between two shapes represented by Φ_1 and Φ_2 is determined by

$$M_d = \frac{\alpha}{\pi}\hat{d}(\bar{\Phi}_1, \bar{\Phi}_2) + D(\Omega_1, \Omega_2), \tag{14}$$

where $\hat{d}(\cdot,\cdot)$ and $D(\cdot,\cdot)$ are defined in (13) and (10), respectively, and α is a weighting factor. In (14), the π is used for the normalization of $\hat{d}(\bar{\Phi}_1, \bar{\Phi}_2)$.

4 Implementation

This section discusses how to find the local minima of $\epsilon(t_1, t_2)$ so that a proper bendable size of an object can be obtained and how to carry out nonuniform sampling on a contour to obtain the landmarks.

Let $z(t)$, $0 \le t < 1$, be the arc-length parameterization of a contour. Since $\epsilon(t_1, t_2) = 1$ when $t_1 \ge t_2$, the local minima can be found only in the region

$$\Delta = \{(t_1, t_2) \in [0,1) \times [0,1) | t_1 < t_2\}. \tag{15}$$

We first uniformly select a number of initial points on Δ (say, 100). Then, beginning from each point, a sequence of gradient descent is made to reach a local minimum. An example is given in Fig. 8. After removing the nonprime pairs (see Definition 3) from the obtained local minima, the bendable size r of the contour can be computed with the prime pairs according to (7).

From $z(t)$, a uniformly sampled sequence

$$\Phi = \left\{ z^k = z\left(\frac{k}{n}\right) \mid k = 1, 2, \ldots, n \right\} \tag{16}$$

can be obtained directly. Then, the piecewise polynomial interpolation [26] can be used to generate a nonuniform sequence from Φ with sample density proportional to $\omega(p, r)$.

Given two curves $z_1(t)$ and $z_2(t)$, suppose that their nonuniform sample sequences are $\bar{\Phi}_1 = \{\bar{z}_1^k\}$ and $\bar{\Phi}_2 = \{\bar{z}_2^k\}$, respectively. An algorithm presented in [21] is available to find the best offset $\hat{s}$ of cyclic permutation such that

$$\hat{s} = \arg\max_{s\in\mathbf{N}} \left|\left\langle \bar{\Phi}_1, \sigma^s(\bar{\Phi}_2)\right\rangle\right|. \tag{17}$$

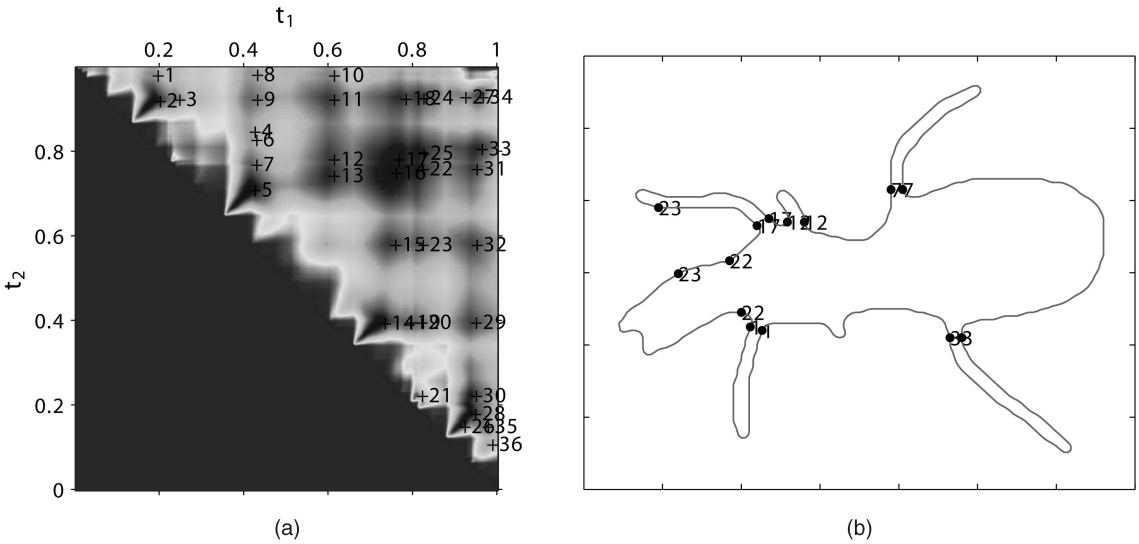


Fig. 8. (a) The terrain map of $\epsilon(t_1, t_2)$ for the contour shown in the right image. There are 36 local minima labeled by "+" on the map. (b) The prime pairs superimposed on the contour. The local minima corresponding to the prime pairs can be found in the terrain map.

5 Experiments

To demonstrate the feasibility of the proposed approach, this section gives the experimental results obtained by applying it to 2D object retrieval based on the contours (silhouettes) of the objects. We test our algorithm on three benchmark databases with comparisons to several recent methods. The weighting factor α in (14) is set to 1/3.

5.1 MPEG-7 Database

Part B of the MPEG-7 Core Experiment CE-Shape-1 data set [35] is often used to test shape matching methods. There are 70 groups of objects in the database and 20 binary images in each group. In a benchmarking test (Bullseye test), each shape is taken as a query and the 40 shapes with the highest scores (or smallest distances in our case) are retrieved from the database (including the query). The percentage of matched images out of $20 \times 70 \times 20 = 28{,}000$ is the retrieval rate of the test.

Our comparison includes nine most recent methods, which are all shown in Table 1.[5] This table indicates that our algorithm outperforms the others.

In our experiments, we also test our algorithm in which either local features or global features are used. We find that the former results in a better score, with 82.65 percent versus 71.43 percent in the Bullseye test. However, the global features are quite useful in matching heavily noisy samples, such as some shapes from the "Devices" class in MPEG-7.

To demonstrate that the flexibility can improve the matching performance, another experiment is designed to carry out the Bullseye test on MPEG-7 by only using the Procrustean distance with uniformly sampled or flexibly sampled landmarks. The experiment shows that the flexibility can improve the Bullseye test score from 63.01 percent to 71.43 percent.

5.2 Gesture Database

Here we test our algorithm with the Gesture database.[6] There are 17 classes and 980 hand gesture samples in it. All of the queries are shown in Fig. 9a. In the benchmarking test proposed in [23], human relevance data are used to evaluate the performance of retrieval. In [23], the authors provided the evaluation on four methods including sequential moments [12], Fourier descriptor [34], geometric moments [15], and the dynamic programming method proposed by themselves. These data are included in our comparison. McNeill and Vijayakumar's evaluation on another two methods, HPM-F and HPM-Fn, with this data set [22] is also included. Fig. 10a shows the precision-recall plots from [23] and [22] with our results added, where the precision is defined as the ratio of relevant gestures retrieved out of all the retrievals, while the recall is the ratio of relevant gestures retrieved out of all of the relevant gestures in the database. The results again show that our algorithm works best.

5.3 Marine Database

There are 20 classes and 1,100 marine species samples in the database. In Fig. 9b, all of the queries are shown. The same tests and evaluations as those in Section 5.2 are performed with this data set. The precision-recall diagram is shown in Fig. 10b, in which the data from [23] and [22] are also included. In this test, our algorithm again shows the best performance.

6 Discussion and Conclusion

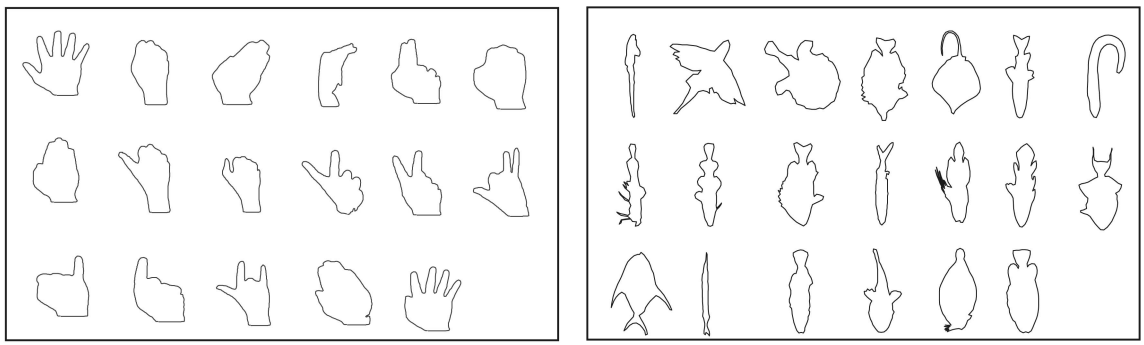
We have presented a novel shape descriptor, called contour flexibility, for 2D shape matching. The bendable potential for each point on the contour can be characterized by this descriptor. The application of the descriptor to extracting the local and global features of a shape is given. With the obtained features, we have

TABLE 1
Retrieval Rates for the MPEG-7 Database

Algorithm	Score(%)
Our method	**89.31**
Shape-tree [11]	87.70
HPM-Fn [22]	86.35
Inner Distance [19]	85.40
Multiscale Representation [1]	84.93
Polygonal Multiresolution [2]	84.33
Chance Probability Function [32]	82.69
Curvature Scale Space [24]	81.12
Generative Model [33]	80.03
Shape Contexts [5]	76.51

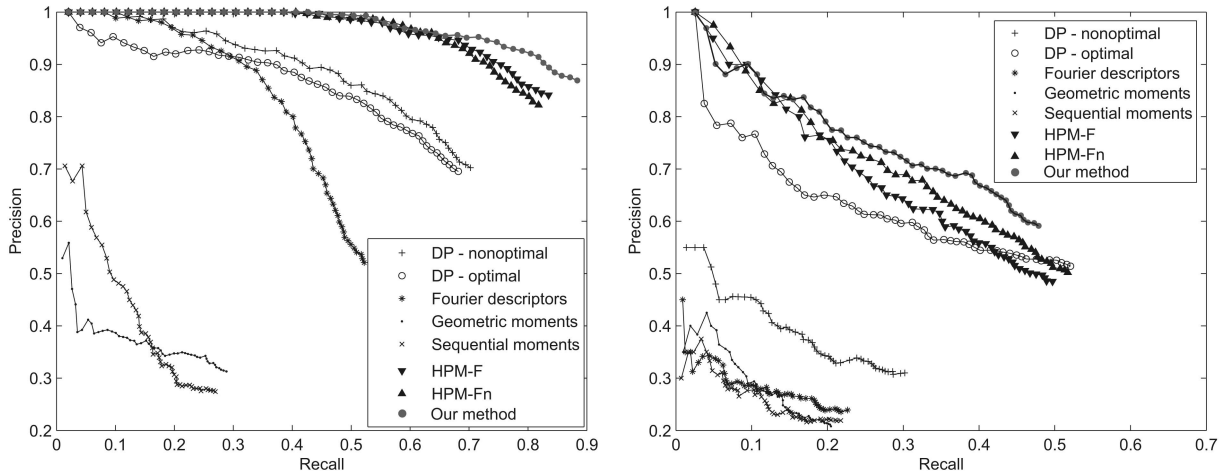
5. These scores, except ours, are from [22], [11], and [5].

6. The Gesture database and the Marine database discussed in Section 5.3 can be downloaded from http://www.intelligence.tuc.gr/~petrakis/.



(a) (b)

Fig. 9. (a) The queries in the Gesture database. (b) The queries in the Marine database.



(a) (b)

Fig. 10. The precision-recall plots for the retrieval tests on (a) the Gesture database and (b) the Marine database.

proposed a shape matching scheme that combines the effectiveness of both local and global features. Our experiments on three benchmark data sets show that our method performs better than the state-of-the-art methods for shape matching.

It should be mentioned that our matching scheme does not consider significant occlusions between objects. When we need to design methods to handle occlusions, however, the flexibility can help in two ways: 1) to determine the decompositions of boundaries and 2) to be used for boundary segment matching.

When the flexibility along the boundary of an object varies insignificantly, it does not help very much for object matching. These objects can be classified into two categories: 1) objects without protruding parts (such as a disk) and 2) ribbon-like objects (such as handwriting numerals). For the first case, the global matching in our scheme plays the main role and the sampling from the flexibility is close to a uniform one, making the global matching reduce to a Procrustean distance matching, which can manage the matching of these objects well. For the second case, other features and matching schemes should be employed because the flexibility extracts little information from the contours.

Future work may include the design of a more sophisticated matching scheme, instead of the linear combination of the two distances in (14). As an intrinsic property of contours, the contour flexibility may be useful for other tasks in computer vision, such as finding the skeleton of a binary shape.

ACKNOWLEDGMENTS

The work described in this paper was supported by a grant from the Research Grants Council of Hong Kong (Project CUHK 414306) and a CUHK Direct Grant.

REFERENCES

[1] T. Adamek and N.E. O'Connor, "A Multiscale Representation Method for Nonrigid Shapes with a Single Closed Contour," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 14, no. 5, pp. 742-753, 2004.

[2] E. Attalla and P. Siy, "Robust Shape Similarity Retrieval Based on Contour Segmentation Polygonal Multiresolution and Elastic Matching," *Pattern Recognition*, vol. 38, no. 12, pp. 2229-2241, 2005.

[3] F. Attneave, "Dimensions of Similarity," *Am. J. Psychology*, vol. 63, no. 4, pp. 516-556, 1950.

[4] R. Basri, L. Costa, D. Geiger, and D. Jacobs, "Determining the Similarity of Deformable Shapes," *Vision Research*, vol. 38, nos. 15-16, pp. 2365-2385, 1998.

[5] S. Belongie, J. Malik, and J. Puzicha, "Shape Matching and Object Recognition Using Shape Contexts," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 509-522, Apr. 2002.

[6] I. Biederman, "Recognition-by-Components: A Theory of Human Image Understanding," *Psychological Rev.*, vol. 84, pp. 115-147, 1987.

[7] F.L. Bookstein, "Size and Shape Spaces for Landmark Data in Two Dimensions," *Statistical Science*, vol. 1, no. 2, pp. 181-222, 1986.

[8] C.C. Chen, "Improved Moment Invariants for Shape Discrimination," *Pattern Recognition*, vol. 26, no. 5, pp. 683-686, 1993.

[9] G.C.H. Chuang and C.C.J. Kuo, "Wavelet Descriptor of Planar Curves: Theory and Applications," *IEEE Trans. Image Processing*, vol. 5, no. 1, pp. 56-70, 1996.

[10] S.A. Dudani, K.J. Breeding, and R.B. McGhee, "Aircraft Identification by Moment Invariants," *IEEE Trans. Computers*, vol. 26, no. 1, pp. 39-46, Jan. 1977.

[11] P.F. Felzenszwalb and J. Schwartz, "Hierarchical Matching of Deformable Shapes," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* vol. 1, pp. 1-8, 2007.
[12] L. Gupta and M.D. Srinath, "Contour Sequence Moments for the Classification of Closed Planar Shapes," *Pattern Recognition,* vol. 20, no. 3, pp. 267-272, 1987.
[13] D.D. Hoffman and W.A. Richards, "Parts of Recognition," *Cognition,* vol. 18, nos. 1-3, pp. 65-96, 1984.
[14] B.W. Hong, E. Prados, S. Soatto, and L. Vese, "Shape Representation Based on Integral Kernels: Application to Image Matching and Segmentation," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* vol. 1, pp. 833-840, 2006.
[15] M.K. Hu, "Visual Pattern Recognition by Moment Invariants," *IEEE Trans. Information Theory,* vol. 8, no. 2, pp. 179-187, 1962.
[16] D.G. Kendall, "Shape Manifolds, Procrustean Metrics and Complex Projective Spaces," *Bull. London Math. Soc.,* vol. 16, pp. 81-121, 1984.
[17] E. Klassen, A. Srivastava, W. Mio, and S.H. Joshi, "Analysis of Planar Shapes Using Geodesic Paths on Shape Spaces," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 26, no. 3, pp. 372-383, Mar. 2004.
[18] I. Kunttu, L. Lepisto, J. Rauhamaa, and A. Visa, "Multiscale Fourier Descriptor for Shape Classification," *Proc. 12th Int'l Conf. Image Analysis and Processing,* vol. 1, pp. 536-541, 2003.
[19] H. Ling and D.W. Jacobs, "Using the Inner-Distance for Classification of Articulated Shapes," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* vol. 2, pp. 719-726, 2005.
[20] K.V. Mardia and I.L. Dryden, "Shape Distributions for Landmark Data," *Advances in Applied Probability,* vol. 21, no. 4, pp. 742-755, 1989.
[21] J.S. Marques and A.J. Abrantes, "Shape Alignment-Optimal Initial Point and Pose Estimation," *Pattern Recognition Letters,* vol. 18, no. 1, pp. 49-53, 1997.
[22] G. McNeill and S. Vijayakumar, "Hierarchical Procrustes Matching for Shape Retrieval," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* vol. 1, pp. 885-894, 2006.
[23] E. Milios and E.G.M. Petrakis, "Shape Retrieval Based on Dynamic Programming," *IEEE Trans. Image Processing,* vol. 9, no. 1, pp. 141-147, 2000.
[24] F. Mokhtarian, *Curvature Scale Space Representation: Theory, Applications, and MPEG-7 Standardization.* Springer, 2003.
[25] S.O.O.C. Pei and C.N.A.N. Lin, "Image Normalization for Pattern Recognition," *Image and Vision Computing,* vol. 13, no. 10, pp. 711-723, 1995.
[26] M.J.D. Powell, *Approximation Theory and Methods.* Cambridge Univ. Press, 1981.
[27] L. Rabiner and B.H. Juang, *Fundamentals of Speech Recognition.* Prentice Hall, 1993.
[28] T.B. Sebastian, P.N. Klein, and B.B. Kimia, "On Aligning Curves," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 25, no. 1, pp. 116-125, Jan. 2003.
[29] T.B. Sebastian, P.N. Klein, and B.B. Kimia, "Recognition of Shapes by Editing Their Shock Graphs," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 26, no. 5, pp. 550-571, May 2004.
[30] K. Siddiqi and B.B. Kimia, "Parts of Visual Form: Computational Aspects," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* vol. 1, pp. 75-81, 1993.
[31] K. Siddiqi, A. Shokoufandeh, S.J. Dickinson, and S.W. Zucker, "Shock Graphs and Shape Matching," *Int'l J. Computer Vision,* vol. 35, no. 1, pp. 13-32, 1999.
[32] B.J. Super, "Learning Chance Probability Functions for Shape Retrieval or Classification," *Proc. IEEE Workshop Learning in Computer Vision and Pattern Recognition,* vol. 6, pp. 93-98, 2004.
[33] Z. Tu and A.L. Yuille, "Shape Matching and Recognition—Using Generative Models and Informative Features," *Proc. Eighth European Conf. Computer Vision,* pp. 195-209, 2004.
[34] T.P. Wallace and P.A. Wintz, "An Efficient Three-Dimensional Aircraft Recognition Algorithm Using Normalized Fourier Descriptors," *Computer Graphics and Image Processing,* vol. 13, no. 2, pp. 99-126, 1980.
[35] L.J. Latecki, R. Lakaemper, and U. Eckhardt, "Shape Descriptors for Non-Rigid Shapes with a Single Close Contour," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 424-429, 2000.