Pattern Recognition Letters 33 (2012) 134–143

Contents lists available at SciVerse ScienceDirect

Pattern Recognition Letters

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

Shape matching and classification using height functions

Junwei Wang [a], Xiang Bai [a,*], Xinge You [a], Wenyu Liu [a], Longin Jan Latecki [b]

[a] Department of Electronics and Information Engineering, Huazhong University of Science and Technology, 1037 Luoyu Road, Wuhan, Hubei Province 430074, PR China
[b] Department of Computer and Information Sciences, Temple University, Philadelphia, PA, USA

ARTICLE INFO

Article history:
Received 19 April 2011
Available online 20 October 2011
Communicated by M.S. Nixon

Keywords:
Shape matching
Contour
Height function
Shape retrieval

ABSTRACT

We propose a novel shape descriptor for matching and recognizing 2D object silhouettes. The contour of each object is represented by a fixed number of sample points. For each sample point, a *height function* is defined based on the distances of the other sample points to its tangent line. One compact and robust shape descriptor is obtained by smoothing the height functions. The proposed descriptor is not only invariant to geometric transformations such as translation, rotation and scaling but also insensitive to nonlinear deformations due to noise and occlusion. In the matching stage, the Dynamic Programming (DP) algorithm is employed to find out the optimal correspondence between sample points of every two shapes. The height function provides an excellent discriminative power, which is demonstrated by excellent retrieval performances on several popular shape benchmarks, including MPEG-7 data set, Kimia's data set and ETH-80 data set.

1. Introduction

Shape matching is a very critical problem in computer vision, which has been widely used in many applications such as object recognition (Belongie et al., 2002; Siddiqi et al., 1999), character recognition (Tang and You, 2003; You and Tang, 2007), shape evolution (Lewin et al., 2010), medical image and protein analysis (Wang et al., 2011), robot navigation (Wolter and Latecki, 2004), and topology analysis in sensor networks (Jiang et al., 2009), etc. It is a very difficult problem, as shape instances from the same category, which look similar to humans, are often very different when measured with geometric transformations (translation, rotation, scaling, etc.) and nonlinear deformations (noise, articulation and occlusion). Compared to geometric transformations, the nonlinear deformations are much more challenging for shape similarity measures. Therefore, one key problem of shape matching is to define a shape descriptor which is informative, discriminative, and efficient for matching process. A good shape descriptor should tolerate the geometric differences of objects from the same category, but at the same time should allow to discriminate objects from different shape classes.

As stated in a previous study (Alajlan et al., 2007), shape descriptors with only global or local information may probably fail to be robust enough in these situations. Global descriptors are robust to local deformations, but they cannot capture local details of shape boundary. Local descriptors are precise to represent local shape features, while they are too sensitive to noise. In fact, it is always challenging to distinguish between noise and local details of the shape boundary. Naturally, one solution to this problem is to define a "rich" shape descriptor, which consists of both global and local shape characteristics. By combining local and global shape features, many recent works (Ling and Jacobs, 2007; McNeill and Vijayakumar, 2006; Felzenszwalb and Schwartz, 2007; Alajlan et al., 2007; Xu et al., 2009) achieved excellent performances on the most popular benchmark: MPEG-7 data set (Latecki et al., 2000).

Some other important requirements for a promising shape descriptor include: computational efficiency, compactness, and generality of applications. It is difficult to satisfy all of these requirements. In this paper, we propose a novel shape descriptor that captures both global and local shape features similar to recent works. For each sample point, a height function is defined as a vector of distances of the other sample points to its tangent line. Then, the whole shape contour is represented as a sequence of the height functions. A further process called smoothing is performed on these height functions to make the descriptor more compact and insensitive to local deformations. After the proposed descriptor is calculated, the Dynamic Programming algorithm is employed to accomplish the shape matching task. Experiments in Section 4 demonstrate the excellent discriminative power of this novel shape descriptor.

Using height functions to represent a shape is partly inspired by a recent work (Liu et al., 2008), which provided a novel definition of curvature to discover the extreme points along the curves with the height functions in all the directions. This work demonstrates

* Corresponding author. Tel.: +86 27 87543236; fax: +86 27 87542667.
E-mail addresses: wjw.20012001@163.com (J. Wang), xiang.bai@gmail.com, xbai@hust.edu.cn (X. Bai), youxg@hust.edu.cn (X. You), liuwy@hust.edu.cn (W. Liu), latecki@temple.edu (L.J. Latecki).

0167-8655/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2011.09.042

that the height functions of the sample points of a given contour can represent the geometric changes and deformations. However, the height function was not used in (Liu et al., 2008) as a kind of shape descriptor for shape matching.

1.1. Related work

There are mainly two categories of shape descriptors: contour-based and region-based methods. In general, contour-based shape descriptors exploit only boundary information, and they cannot capture shape interior content. Besides, these methods cannot deal with complex shapes consisting of disjoint parts (Zhang and Lu, 2002). In region-based techniques, shape descriptors are derived using all the pixel information within a shape region. In contrast to contour-based approaches, region-based methods are more reliable for complex shapes such as trademarks, logos and characters (Kim and Kim, 2000). However, (especially some early) region-based methods consist of only global shape characteristics without many important shape details. Therefore, their discriminative power is limited for large databases or complex situations when there are a lot of intra-class variations.

Some well known and recent works for the region-based shape description include Zernike moments (Kim and Kim, 2000), generic Fourier descriptor (Zhang and Lu, 2002), multi-scale Fourier-based description (Direkoglu and Nixon, 2008) and so on. In (Kim and Kim, 2000), a region-based shape descriptor is presented utilizing a set of the magnitudes of Zernike moments. This descriptor has many desirable properties such as rotation invariance, robustness to noise, fast computation and multi-level representation for shape description (Kim and Kim, 2000). Zhang and Lu (2002) propose a generic Fourier descriptor (GFD), which is extracted from spectral domain by applying 2-D Fourier transform on polar raster sampled shape image. Compared with Zernike moments, GFD has no redundant features and allows multi-resolution feature analysis in both radial and angular directions (Zhang and Lu, 2002). In (Direkoglu and Nixon, 2008), the authors believe that the boundary and exterior parts create much more contributions to object recognition than the central part. Based on this observation, they produce new multi-scale Fourier-based descriptors in 2-D space, which represent the boundary and exterior parts of an object and also allow the central part to contribute to shape classification slightly (Direkoglu and Nixon, 2008).

As there have been a lot of contour-based works on shape representation and matching, we only review the most recent approaches. In the last decade, several contour-based shape descriptors have been presented and studied, which are ‘‘rich’’ descriptors with both global and local information. These methods include curvature scale space (CSS) (Mokhtarian et al., 1997), multi-scale convexity concavity (MCC) (Adamek and OĆonnor, 2004), triangle area representation (TAR) (Alajlan et al., 2007, 2008), hierarchical procrustes matching (HPM) (McNeill and Vijayakumar, 2006), shape tree (Felzenszwalb and Schwartz, 2007), contour flexibility (Xu et al., 2009), shape context (SC) (Belongie et al., 2002), inner-distance shape context (IDSC) (Ling and Jacobs, 2007) and so on.

One type of ‘‘rich’’ descriptors is defined in a multi-scale space. curvature scale space (CSS) (Mokhtarian et al., 1997) and multi-scale convexity concavity (MCC) (Adamek and OĆonnor, 2004) are two classical descriptors of this type. In both of them, contour convolution is performed using Gaussian kernel smoothing. By changing the sizes of Gaussian kernels, several shape approximations of the shape contour at different scales are obtained. The descriptors are defined based on these shape approximations. For CSS, a shape feature called Curvature Scale Space image is defined, which records the locations of curvature zero crossings on all evolved curves; the maxima of the curvature zero-crossing contours in the Curvature Scale Space image are used to represent shapes (Mokhtarian et al., 1997). For MCC, the displacements of contour sample points between every two consecutive scale levels are calculated, which are used to represent contour convexities and concavities at different scale levels (Adamek and OĆonnor, 2004). The main limitation for this type of descriptors is that it is difficult to determine the optimal parameter of each scale. Another problem of CSS is that CSS is not a good choice for convex shapes, as there is no curvature zero crossing for convex objects.

Another type of multi-scale descriptors is defined directly on the original shape contours without any preprocessing, including triangle area representation, hierarchical procrustes matching and shape tree. triangle area representation (TAR) (Alajlan et al., 2007, 2008) presents a measure of convexity/concavity of each contour point using the signed areas of triangles formed by boundary points at different scales. The area value of every triangle is a measure for the curvature of corresponding contour point, and the sign of the area is positive, negative or zero when the contour point is convex, concave or on a straight line, respectively. This representation is effective in capturing both local and global characteristics of a shape (Alajlan et al., 2007).

Hierarchical procrustes matching (HPM) (McNeill and Vijayakumar, 2006) and shape tree (Felzenszwalb and Schwartz, 2007) are two classical segment-based shape matching algorithms. In both of them, closed shape contours are divided into curve segments hierarchically, and the matching process is performed by comparing these segments explicitly. In HPM (McNeill and Vijayakumar, 2006), the contour of each shape is divided into overlapped segments with relative arc length percentages 50%, 25% and 12.5% of the whole length. HPM achieves segment matching in a global to local direction, i.e., longer segments that have already been matched together provide an initial match for shorter segments (McNeill and Vijayakumar, 2006). For shape tree, one curve can be broken into two halves by the middle point on it, and each of the two sub-curves can be broken into its halves. This hierarchical description is represented by a binary tree called the *shape tree*. When matching two curves $\boldsymbol{A}$ and $\boldsymbol{B}$, they build a Shape Tree for curve $\boldsymbol{A}$ and search a mapping from points in $\boldsymbol{A}$ to points in $\boldsymbol{B}$ such that the shape tree of $\boldsymbol{A}$ is deformed as little as possible (Felzenszwalb and Schwartz, 2007). Shape tree achieves high retrieval rates on both MPEG-7 data set (87.70%) and Kimia's data set (see in Section 4). However, it suffers from an expensive computational complexity.

Another interesting hierarchical approach is proposed by Payet and Todorovic (2009), which convents the contour matching problem into a graph matching framework for the first time. The shape representation in (Payet and Todorovic, 2009) based on salient contour parts is similar to Felzenszwalb and Schwartz (2007).

Contour flexibility (Xu et al., 2009) is a kind of rich descriptor for planar contours, which depicts the deformable potential at each point along a curve. Contour flexibility provides the information about how extensively the neighborhood of a contour point is connected to the main body and about the deformation tolerance of an object at this point. This method achieves an excellent retrieval result of 89.31% on MPEG-7 data set.

Besides the above descriptors, there is another type of rich descriptors, for which the geometric relationship between contour sample points is utilized. This type of descriptors includes Shape Context and Inner-Distance Shape Context. For every sample point, Shape Context (SC) (Belongie et al., 2002) captures the spatial distribution of all the other sample points relative to it. The spatial distribution is represented by a coarse histogram, and the bins in the histogram are uniform in log-polar space, which makes the descriptor more sensitive to nearby sample points than to points farther away (Belongie et al., 2002). To make contour-based shape descriptors articulation insensitive, Ling and Jacobs proposed one novel distance definition called *inner distance* as a replacement for the Euclidean distance. The inner distance is defined as the length of the shortest path between landmark points within the shape silhouette. Using the inner distance, Shape Context can be extended to a novel descriptor called Inner-Distance Shape Context (IDSC) (Ling and Jacobs, 2007). One limitation for the inner distance is that the shape boundary is assumed to be known prior to the computation.

Besides improving shape descriptors, there is now growing interest in learning context-sensitive or contextual similarity (Yang et al., 2008, 2009; Bai et al., 2010,; Kontschieder et al., 2009) from a collection of pairwise similarities among the database (context) and using this context to infer semantic similarities between query shapes/images against the database and to perform concept-based information retrieval. As pointed out by Yang et al. (2008), a "good" similarity between a query shape q and a known shape p should describe the relationship between q and p in the context of the database. The solution of this intuition in (Yang et al., 2008) is to compute a new similarity s for a given similarity measure s_0 by a graph transduction method named Label Propagation (LP). With the similar manner to Yang et al. (2008), Locally Constrained Diffusion Process (LCDP) (Yang et al., 2009) was developed to explore the contextual information, which achieves better performances on several shape benchmarks than LP (Yang et al., 2008). Another contribution of Yang et al. (2009) is that a few ghost objects are constructed and added into the similarities space to enhance the contextual information. Then, Egozi et al. (2010) proposed a contextual similarity function named Meta Similarity (MS) which is to characterize a given shape by its similarity to its K-NN database shapes. One advantage of MS is the low time complexity as it does not require propagating the similarities. Finally, an interesting context learning method called Contextual Dissimilarity Measure (CDM) (Jegou et al., 2007) is motivated by an observation that a good ranking is usually not symmetrical in image search, which is mainly designed for image search problem. To sum up, a new similarity is usually learned through these context-sensitive approaches, which significantly improves the performance of shape retrieval and classification. We combine the shape similarities obtained by the proposed method and a contextual similarity method (Yang et al., 2009), then achieve the best ever score 96.45% on MPEG-7 data set.

The main contribution of this paper is a novel shape descriptor for matching and recognizing planar shape contours that is:

(1) simple and easy to compute (see the definition of the proposed descriptor in Section 2).
(2) invariant to geometric transformations such as translation, rotation and scaling (see the analysis in Section 2).
(3) insensitive to nonlinear deformations (see the analysis in Section 2 and experimental results in Section 4).
(4) compact and computationally efficient (see the analysis in Sections 4.1 and 4.6).

The remainder of this paper is organized as follows. Section 2 presents the novel descriptor based on height functions in details. Section 3 discusses the shape matching algorithm we use, along with the similarity measure. Section 4 presents the experiments and discusses the computational complexity. Finally, Section 5 draws some conclusions.

2. Shape descriptor with height functions

Let $\boldsymbol{X} = \{\boldsymbol{x}_i\}$ $(i = 1,\ldots,N)$ denotes the sequence of equidistant sample points on the outer contour of a given shape, where the index i is according to the order of the sample points along the contour in counter-clockwise direction. In our implementation, we set $N = 100$, which is consistent with the setting in (Adamek and OĆonnor, 2004; Ling and Jacobs, 2007).

To compute the height functions, the most important step is to determine the axes. The method (Liu et al., 2008) which computes the height values in all the directions in some discrete angle steps is not favorable for efficient shape matching due to the high time complexity and limited angular resolution. Therefore, we do not consider any fixed angular directions, but instead adjust the angular direction for each sample point. Here the key observation is that the tangent line provides an excellent reference line for height functions. For each sample point $\boldsymbol{x}_i$, we denote as its reference axis the tangent line l_i. The tangent line l_i inherits its orientation from the contour orientation, i.e., its direction is always starting from $\boldsymbol{x}_{i-1}$ to $\boldsymbol{x}_{i+1}$, where the point indices are considered modulo N.

Then, the distance between the jth $(j = 1,\ldots,N)$ sample point $\boldsymbol{x}_j$ and the tangent line l_i is defined as a height value $h_{i,j}$. Moreover, we give the height value $h_{i,j}$ a symbol. Specifically, the height value $h_{i,j}$ is positive, negative or zero when the jth sample point $\boldsymbol{x}_j$ is to the left of, to the right of or just on the axis l_i. Obviously, the positive/negative symbol of the height value makes a more precise representation for the relative location of the point $\boldsymbol{x}_j$ to the axis l_i: we know not only the distance, but also which side it stands on. Note that some height values may be negative especially when the point $\boldsymbol{x}_i$ is in a concave part of the shape $\boldsymbol{X}$, and the contour of $\boldsymbol{X}$ is divided into two or more parts by the axis l_i consequently. Fig. 1(a) shows an example of the proposed height values. In Fig. 1(a), the height function is defined for the sample point $\boldsymbol{x}_i$. The height value $h_{i,v}$ is positive as the sample point $\boldsymbol{x}_v$ is to the left of the axis l_i, and the height values $h_{i,u}$, $h_{i,w}$ are negative as the corresponding sample points $\boldsymbol{x}_u$, $\boldsymbol{x}_w$ are both to the right of the axis l_i.

We calculate the height values of every sample point to the axis l_i. Then the shape descriptor of the point $\boldsymbol{x}_i$ with respect to the shape $\boldsymbol{X}$ is the ordered sequence of the height values

$$H_i = (h_i^1, h_i^2, \ldots, h_i^N)^T = (h_{i,i}, h_{i,i+1}, \ldots, h_{i,N}, h_{i,1}, h_{i,2}, \ldots, h_{i,i-1})^T, \quad (1)$$

where $h_{i,j}$ $(j = 1,\ldots,N)$ denotes the height value of the jth sample point $\boldsymbol{x}_j$ according to the reference axis l_i of the point $\boldsymbol{x}_i$. Note that the height values in this sequence is in the order of the sequence of sample points $\{\boldsymbol{x}_i\}$ $(i = 1,\ldots,N)$, always starting from $h_{i,i}$, the height value of the sample point $\boldsymbol{x}_i$ itself to its reference axis l_i. We observe that $h_i^1 = h_{i,i} = 0$ for every $i = 1,\ldots,N$. We treat H_i as a column vector.

The proposed descriptor H_i depends not only on the direction of the reference axis l_i, but also on the location of the sample point $\boldsymbol{x}_i$ on the shape contour $\boldsymbol{X}$. In Fig. 1(b), the height descriptors for three sample points $\boldsymbol{x}_i$, $\boldsymbol{x}_w$ and $\boldsymbol{x}_u$ are given. It is obvious that our descriptor may be totally different even if the reference axis has nearly the same direction. This leads to a strong discriminative power to find out the correct correspondence between the sample points from two shapes.

Since the reference axis and the height values are defined directly on the sample points, the proposed descriptor explicitly contains the information of geometric relationship of the sample points. This makes the height function representation invariant to translations and rotations, as all the sample points will translate and rotate synchronously and the geometric relationship between them will remain unchanged.

However, the descriptor defined above consists of the relative location for EVERY SINGLE sample point to the reference axis. Such a precise description may be too sensitive to local boundary deformations. One solution to this problem is to smooth the predefined descriptor H_i in Eq. (1) without losing too much geometric information. A simple strategy is adopted here: for a given integer k $(1 < k < N)$, we divide the sequence of integers $1, 2, \ldots, N$ into disjoint intervals $[1,k], [k+1, 2k], \ldots$, and compute the mean value of the height values in each interval:

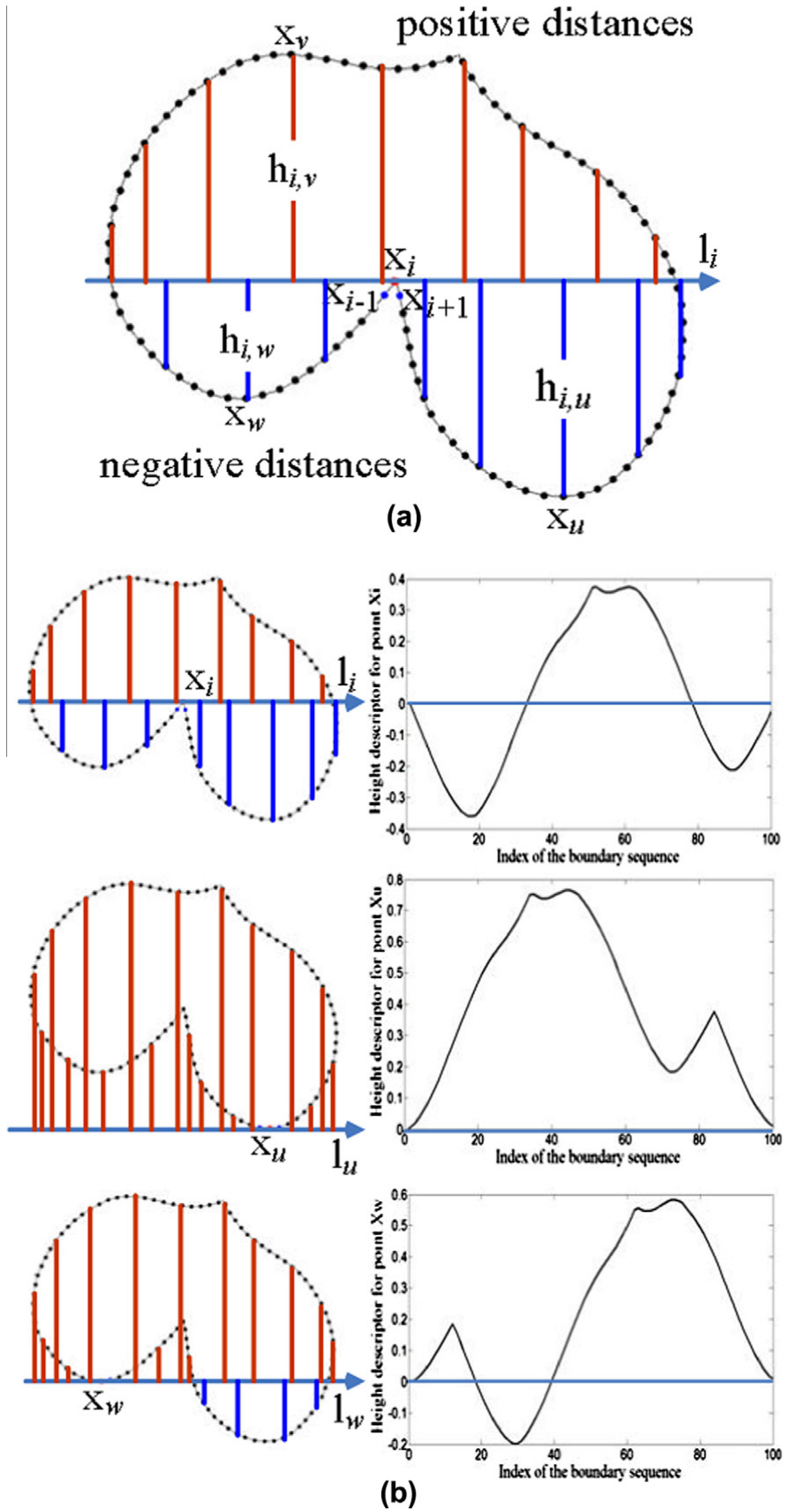

Fig. 1. Height functions. The "heart" shape is selected from MPEG-7 data set (Latecki et al., 2000). (a) The height function for the sample point $\boldsymbol{x}_i$. (b) Height descriptors for different sample points with similar axis direction.

$$f_i^j = \frac{1}{k}\sum_{t=(j-1)k+1}^{jk} h_i^t, \qquad (2)$$

where $j = 1,\ldots,M$ with $M = \lfloor N/k \rfloor$ (the integer part of N/k), and the arithmetic is modulo N. In this way, we obtain a new feature vector representing the smoothed height values of the sample point $\boldsymbol{x}_i$ as

$$F_i = (f_i^1, f_i^2, \ldots, f_i^M)^T. \qquad (3)$$

F_i can be regarded as a smooth version of H_i. The smoothing process not only makes the descriptor more stable to local deformations but also reduces its dimensionality from N to M by the ratio k, since $k > 1$ and $M < N$ consequently. For example, if $N = 100$ and $k = 5$, then the dimensionality of the descriptor is reduced to $M = 20$.

For every sample point $\boldsymbol{x}_i$ ($i = 1,\ldots,N$), we get a descriptor F_i ($i = 1,\ldots,N$). As the proposed descriptor of the shape $\boldsymbol{X}$, we define a matrix

$$\mathcal{F} = \mathcal{F}(\boldsymbol{X}) = (F_1, F_2, \ldots, F_N). \qquad (4)$$

We observe that $\mathcal{F}$ is an $M \times N$ matrix with column i being the shape descriptor F_i of the sample point $\boldsymbol{x}_i$. In our implementation, the values of N and M are 100 and 20, respectively.

In order to make our shape representation scale invariant, we row wise normalize $\mathcal{F}$ by dividing by the maximal absolute value of each row:

$$f_i^j = \frac{f_i^j}{\max_{t=1,\ldots,N}\{|f_t^j|\}}. \qquad (5)$$

Separated from the usual normalization process by some global measure for the shape contour, the normalization according to Eq. (5) can be regarded as *local normalization*. A detailed analysis for the advantages of this kind of normalization is given in (Alajlan et al., 2007). Consequently, the value of each entry in the matrix $\mathcal{F}$ after normalization is in the interval $[-1, 1]$.

3. Similarity measure using the height descriptor

For the task of shape recognition, usually a shape similarity or dissimilarity (distance) is computed by finding the optimal correspondence of contour points, which is used to rank the database shapes for shape retrieval. In this paper, we use Dynamic Programming (DP) algorithm to find the correspondence. Then the shape dissimilarity is the sum of the distances of the corresponding points.

To match two shapes $\boldsymbol{X}$, $\boldsymbol{Y}$ (represented by two sequences of sample points), the dissimilarity between any pair of points should be computed. Let p, q denote contour points of $\boldsymbol{X}$, $\boldsymbol{Y}$, respectively, and F_p, F_q denote their height features. The cost (distance) of matching p and q is computed by comparing their height features:

$$c(p,q) = \sum_{t=1}^{M} \omega_t |f_p^t - f_q^t|, \qquad (6)$$

where ω_t is the weight coefficient for every component of the height feature. In order to be able to tolerate boundary deformations, the differences of the height values of points further away from points p and q are treated as less important than the differences of points closer to p and q. To achieve this, the weight coefficients are set as follows:

$$\omega_t = \frac{1}{\min\{t, M-t\}}. \qquad (7)$$

Given two shapes $\boldsymbol{X} = \{\boldsymbol{x}_i\}$ and $\boldsymbol{Y} = \{\boldsymbol{y}_i\}$ for $i = 1,\ldots,N$, we compute $c(\boldsymbol{x}_i, \boldsymbol{y}_j)$ for $i, j = 1,\ldots,N$. Then with DP we compute an optimal correspondence $g^*: \boldsymbol{X} \to \boldsymbol{Y}$ such that $\sum_{i=1}^{N} c(\boldsymbol{x}_i, g^*(\boldsymbol{x}_i))$ is minimal. Finally the dissimilarity between the two shapes is given by

$$DP_{min}(\boldsymbol{X}, \boldsymbol{Y}) = \sum_{i=1}^{N} c(\boldsymbol{x}_i, g^*(\boldsymbol{x}_i)). \qquad (8)$$

One additional factor in measuring shape dissimilarity is *shape complexity*, which has been used by MCC (Adamek and OĆonnor, 2004) and TAR (Alajlan et al., 2007, 2008) to improve the shape similarity measure. The idea comes from the observation that humans are generally more sensitive to contour deformations when the complexity of the contour is lower (Adamek and OĆonnor, 2004). The novel descriptor defined in Section 2 can be easily utilized to define the shape complexity:

$$C(\boldsymbol{X}) = \frac{1}{M}\sum_{t=1}^{M} std(f_1^t, f_2^t, \ldots, f_N^t), \qquad (9)$$

where *std* denotes the standard deviation. Note that the shape complexity defined here is different from those defined in (Adamek and OĆonnor, 2004; Alajlan et al., 2007). In (Adamek and OĆonnor, 2004; Alajlan et al., 2007), the difference between the maximum and minimum value of each row of the descriptor matrix is used and averaged instead of the standard deviation used here. We make this change to make the shape complexity values more stable.

The dissimilarity or distance between two shapes $\boldsymbol{X}$, $\boldsymbol{Y}$ normalized by their shape complexity values is given by

$$DCN(\boldsymbol{X}, \boldsymbol{Y}) = \frac{DP_{min}(\boldsymbol{X}, \boldsymbol{Y})}{\beta + C(\boldsymbol{X}) + C(\boldsymbol{Y})}, \qquad (10)$$

where the factor β, which is used to avoid divide-by-zero, is set empirically.

4. Experiments and analysis

In this section, we show that the proposed method achieves encouraging results on three popular benchmarks: MPEG-7 data set (Latecki et al., 2000), Kimia's 99 data set (Sebastian et al., 2004) and ETH-80 data set (Leibe and Schiele, 2003). We also give an analysis for the computational complexity of the proposed method.

4.1. MPEG-7 data set

The MPEG-7 data set (Latecki et al., 2000) is widely used for testing the performances of shape descriptors in the last decade. This database consists of 1400 silhouette images which are grouped into 70 classes with 20 objects per class. Some shapes from different categories are similar, and there are always some complex deformations for the shapes within the same category. Some examples for this data set are given in Fig. 2(a). The retrieval rate on MPEG-7 data set is measured by the so-called *bull's eyes* score which counts how many objects within the 40 most similar objects belong to the class of the query object. Every shape in the data set is used as a query, and the retrieval result for the whole data set is obtained by averaging among all shapes.

In Table 1, the overall retrieval rates on MPEG-7 database for the novel height based descriptor and other recent important methods are put together. We observe that the proposed method DP_{min} outperforms almost all the other shape descriptors even without using the global information of shape complexity presented in Section 3. When shape complexity is further used to normalize the shape distance based on DP_{min}, the overall retrieval rate of the proposed method (with *DCN*) is **90.35%**, the first result over 90.0% for methods only with shape descriptors (without context information among the database). Respectively, the best performance ever reported is 89.62% by the Locally Affine Invariant Descriptors (Wang and Liang, 2010).

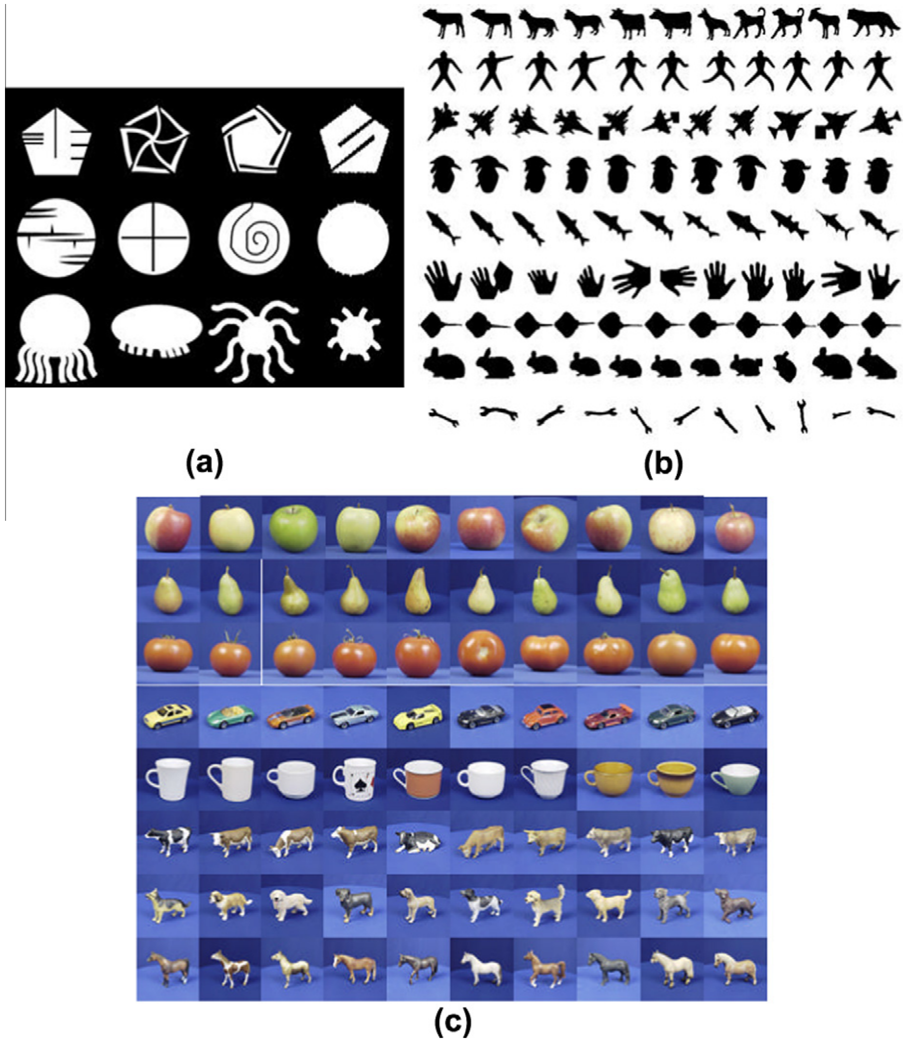

Fig. 2. Databases used in the experiments, (a) Some examples from MPEG-7 data set (Latecki et al., 2000). (b) Kimia's 99 data set (Sebastian et al., 2004). (c) ETH-80 data set (Leibe and Schiele, 2003).

Note that a shape recognition scheme can be easily performed here based on the pairwise shape matching scores by one kind of very simple classifier: *one nearest neighbor* (1NN). With the leave-one-out procedure, every one of the 1400 shapes in the database is used as a test input, and the recognition result is considered correct if the closest (non-identical) match in the database is in the same category as the query shape (Super, 2006). Experiment shows that height functions achieves an excellent result comparative with some recent approaches. With *DCN*, the recognition rate is **98.86%**, which equals the state-of-art result reported by Variational Shape Matching (Nasreddine et al., 2010). This result is followed by that of Symbolic Representation (Daliri and Torre, 2008) (98.57%), Class Segments Set (Sun and Super, 2006) (97.93%), and Polygonal Multiresolution (Attalla and Siy, 2005) (97.79%). The value of our result means that only 16 out of 1400 shapes are assigned to the wrong class, which is a rather low error rate.

For the experiments on MPEG-7 data set, we set $k = 5$ so that the height descriptor for each sample point is a vector with the size **20** (i.e., $M = 20$), which is significantly smaller than that of SC (5 distance scales and 12 angle scales for building the histogram, thus **60** bins in total (Belongie et al., 2002)), that of IDSC (8 distance scales and 12 angle scales, thus **96** bins in total (Ling and Jacobs, 2007)) and that of TAR (the number of scales is $\lfloor (N-1)/2 \rfloor$, i.e., the integer part of $(N-1)/2$, which is **63** when $N = 128$ (Alajlan et al., 2007, 2008)). Although the proposed descriptor leads to a more compact representation, we achieved better retrieval performance than these descriptors. This is because our descriptor includes not only the height value of each sample point but also the order information of the sample points along the contour (see Eq. (1)).

Besides the retrieval and recognition experiments on MPEG-7 data set, we show that the similarities obtained by the proposed method can be improved by combining them with a context-sensitive learning method: Locally Constrained Diffusion Process (LCDP) (Yang et al., 2009). Table 2 shows the result and compares it to other context-sensitive methods. With LCDP, our method achieves the best ever score **96.45%** on MPEG-7 data set.

For LCDP, we set the number of nearest neighbors $K_1 = 8$ and the number of iterations $r = 18$. When converting distance matrices to affinity matrices, the parameters for Gaussian kernel are set as $K_2 = 31$ and $\alpha = 0.34$. Details of these parameters can be checked in (Yang et al., 2009).

4.2. Kimia's data set

The Kimia's data set (Sebastian et al., 2004) is also widely used in shape matching and classification. It contains 99 images from nine categories (as shown in Fig. 2(b)). There are some occlusions,

Table 1
Bull's eyes scores on MPEG-7 data set (Latecki et al., 2000).

Algorithm	Score (%)
Height functions + shape complexity	**90.35**
Height functions	**89.66**
Locally affine invariant descriptors (Wang and Liang, 2010)	89.62
Contour flexibility (Xu et al., 2009)	89.31
Variational shape matching (Nasreddine et al., 2010)	89.05
Layered graph (Lin et al., 2009)	88.75
Two strategies (Temlyakov et al., 2010)	88.39
Aspect shape context (Ling et al., 2010)	88.30
Hierarchical parts (Payet and Todorovic, 2009)	88.30
Hilbert curve (Ebrahim et al., 2009)	88.30
Shape tree (Felzenszwalb and Schwartz, 2007)	87.70
TAR + shape complexity + global (Alajlan et al., 2007)	87.23
TAR + shape complexity (Alajlan et al., 2008)	87.13
SC + DP (Bai et al., 2010)	86.80
HPM (McNeill and Vijayakumar, 2006)	86.35
Symbolic representation (Daliri and Torre, 2008)	85.92
IDSC + DP (Ling and Jacobs, 2007)	85.40
Planar graph cuts (Schmidt et al., 2009)	85.00
MCC + shape complexity (Adamek and OĆonnor, 2004)	84.93
Polygonal multiresolution (Attalla and Siy, 2005)	84.33
Fixed correspondence (Super, 2006)	84.05
Optimized CSS (Mokhtarian et al., 1997)	80.54
Generative model (Tu and Yuille, 2004)	80.03
Skeletal contexts (Xie et al., 2008)	79.92
Distance set (Grigorescu and Petkov, 2003)	78.38
SC + TPS (Belongie et al., 2002)	76.51

Table 2
Bull's eyes scores of descriptors combined with context-sensitive methods on MPEG-7 data set (Latecki et al., 2000).

Algorithm	Score
Height functions + LCDP	**96.45**
Aspect shape context + LCDP (Ling et al., 2010)	95.96
Two strategies + LCDP (Temlyakov et al., 2010)	95.60
IDSC + mutual graph (Kontschieder et al., 2009)	93.40
IDSC + SSP (Wang et al., 2011)	93.35
IDSC + LCDP (Yang et al., 2009)	93.32
SC + LP (Bai et al., 2010)	92.91
ZGM + SC + meta descriptor (Egozi et al., 2010)	92.51
IDSC + LP (Bai et al., 2010)	91.61
GM + IDSC + meta descriptor (Egozi et al., 2010)	91.46

missing parts, and articulations in this data set. Thus, it is popular for both contour-based approaches and skeleton-based ones. In the experiment, every shape in the database is considered as a query, and the retrieval result is summarized as the number of top 1 to top 10 closest matches in the same class (excluding the query itself). The best possible result for each of the ranking is 99. Table 3 lists the results of height functions and some other recent methods. Our method performs comparably with some recent approaches such as perceptual strategies (Temlyakov et al., 2010), shape tree (Felzenszwalb and Schwartz, 2007) and geometric shape matching (Egozi et al., 2010). The best performance on this data set is achieved by Symbolic Representation (Daliri and Torre, 2008).

4.3. ETH-80 data set

The ETH-80 data set (Leibe and Schiele, 2003) contains eight categories of objects with 10 objects per category. For each object, 41 color images from different viewpoints are provided. Thus there are 3,280 ($8 \times 10 \times 41$) images in total in this database. 80 instances (one image/viewpoint for every object) of this data set are given in Fig. 2(c).

The intended test mode is *leave-one-object-out cross-validation*. That is, we train with 79 objects and test with the one unknown object. Recognition is considered successful if the object is assigned to the correct category label. The final result is achieved by averaging over all 80 objects (Leibe and Schiele, 2003).

Table 4 lists the results of height functions and some other single-cue methods. Again, the proposed method achieves an excellent recognition rate, which outperforms all methods except for only one result (comparable) reported by Symbolic Representation (Daliri and Torre, 2008).

4.4. Matching and retrieval under noisy conditions

The experiments on the above three commonly used data sets have demonstrated the effectiveness of the proposed method. However, the shapes are usually quite smooth in these data sets. To evaluate the performance of our descriptor under noisy conditions, we add Gaussian noise to shape boundaries and perform matching between the noise-deformed shapes.

We use Kimia's 99 data set (Sebastian et al., 2004) as the original shape boundaries. Noise is added by perturbing all pixels on each shape contour in both x- and y-coordinates by values drawn from a Gaussian random variable with zero mean and standard derivation σ. As the parameter σ increases, we add increasing Gaussian noise to the shape boundaries. Fig. 3 shows an example of shape boundaries with increasing Gaussian noise.

The test mode is the same as in Section 4.2. We evaluate our method along with two other kinds of classical descriptors, SC and IDSC. In this experiment, we use 8 distance scales and 12 angle scales for building the histograms of both SC and IDSC. In the matching stage, we use dynamic programming algorithm for all of the three methods. Table 5 lists the results of height functions, SC and IDSC with different Gaussian noise levels. Under low level noisy conditions (when $\sigma \leqslant 0.4$), our method outperforms SC and IDSC; when noise becomes serious, i.e., when $\sigma \geqslant 0.6$, the three methods perform comparably.

To make it easier to compare the performances of our method with other descriptors, Fig. 4 shows the results of the 2nd, 5th, 8th,

Table 3
Retrieval results on Kimia's 99 data set (Sebastian et al., 2004).

Algorithm	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
SC (Belongie et al., 2002)	97	91	88	85	84	77	75	66	56	37
Generative model (Tu and Yuille, 2004)	99	97	99	98	96	96	94	83	75	48
Path similarity (Bai and Latecki, 2008)	99	99	99	99	96	97	95	93	89	73
Hierarchical parts (Payet and Todorovic, 2009)	99	99	98	98	98	97	96	94	93	82
Shock Graph (Sebastian et al., 2004)	99	99	99	98	98	97	96	95	93	82
IDSC (Ling and Jacobs, 2007)	99	99	99	98	98	97	97	98	94	79
TAR (Alajlan et al., 2007)	99	99	99	98	98	97	98	95	93	80
Shape tree (Felzenszwalb and Schwartz, 2007)	99	99	99	99	99	99	99	97	93	86
GM + SC (Egozi et al., 2010)	99	99	99	99	99	99	99	97	93	86
Two strategies (Temlyakov et al., 2010)	99	99	99	98	99	99	99	97	96	84
Height functions	**99**	**99**	**99**	**99**	**98**	**99**	**99**	**96**	**95**	**88**
Symbolic representation (Daliri and Torre, 2008)	99	99	99	98	99	98	98	95	96	94

Table 4
Recognition rates on ETH-80 data set (Leibe and Schiele, 2003).

Algorithm	Score (%)
Symbolic representation (Daliri and Torre, 2008)	90.28
Height functions + shape complexity	**89.73**
Height functions	**88.72**
IDSC + DP (Ling and Jacobs, 2007)	88.11
SC + DP (Leibe and Schiele, 2003)	86.40
SC greedy (Leibe and Schiele, 2003)	86.40
PCA masks (Leibe and Schiele, 2003)	83.41
PCA gray (Leibe and Schiele, 2003)	82.99
Mag-lap (Leibe and Schiele, 2003)	82.23
D_xD_y (Leibe and Schiele, 2003)	79.79
Color histogram (Leibe and Schiele, 2003)	64.85

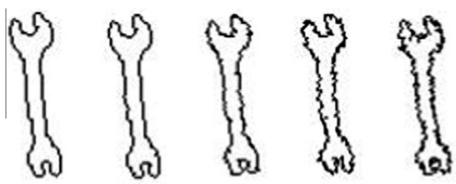

Fig. 3. An example of shape boundaries with increasing Gaussian noise. (a) The original shape boundary. The parameter σ from (b) to (e) increases from 0.2 to 0.8.

and 10th top matches separately. The performances of the three methods with respect to increasing noise levels are put together. Note that all data in Fig. 4 is selected from Table 5. From the results we find that, for the 2nd and 5th top matches, our method outperforms SC and IDSC; for the 8th and 10th top matches, the three methods perform comparably.

4.5. The settings of parameters

There are three parameters in the proposed method: k, the number of consecutive sample points when we smooth the original height functions; λ, the penalty factor in the DP algorithm; and β, the factor when we improve the shape distances with shape complexity.

For different data sets, the three parameters are adjusted accordingly to achieve the optimal retrieval results. For MPEG-7 data set, the best result is achieved when $k = 5$, $\lambda = 0$ and $\beta = 0.19$. For Kimia's 99 data set, these parameters are set as $k = 5$, $\lambda = 0.1$ and $\beta = 0.46$. For ETH-80 data set, the optimal values for the parameters are $k = 5$, $\lambda = 0$ and $\beta = 0.023$. We find that the optimal values for the parameters k and λ are quite stable over different data sets. On the contrary, the optimal value for the parameter β is not that stable. One possible explanation to the instability of β is given here. As defined by Eq. (10), the effect of the parameter β is not only to avoid divide-by-zero, but also to help to achieve a more reasonable distance (DCN) between two shapes. As Eq. (10) makes use of the summation of β and the shape complexity values as the normalization factor, the optimal value of β is closely dependent on the shape complexity values. For example, if the shape complexity values for one database are all between 0 and 1, then we cannot set $\beta = 10{,}000$; if the value of β is too large, the parameter β will be dominant, and the distance between shapes will not get any improvement through Eq. (10). Moreover, from Eq. (9) we know that the shape complexity is directly defined on the descriptor-height functions, i.e., dependent on characteristics of every shape. Different shapes (from different databases) have different shape complexity values. As a result, the optimal value of the parameter β varies for different data sets.

In noisy shape matching experiments, parameter value adjustment is performed in a similar manner. For different noise levels, the parameters are adjusted accordingly to achieve the optimal retrieval results. As we test three descriptors (height functions, SC, and IDSC) under noisy conditions, we introduce the parameter settings for each method one by one. For the proposed method, there are three parameters as stated in previous paragraphs. In all noise levels, the optimal value of the parameter k is set to 5. The value of the parameter λ is set to 0.1 under low level noisy conditions (when $\sigma \leqslant 0.4$), and is set to 0.2 when noise becomes serious (i.e., when $\sigma \geqslant 0.6$). The optimal value of the parameter β varies under different noise levels: the optimal value is 0.46, 0.68, 0.27, 0.37, and 0.49 when σ = 0, 0.2, 0.4, 0.6, and 0.8, respectively. Again, we find that the optimal values for the parameters k and λ are quite stable when the proposed method is used for different noise levels, and the optimal value for the parameter β varies.

For the experiments of SC and IDSC under noisy conditions, there is only one parameter-λ, the penalty factor in the DP algorithm. The value of this parameter is set manually depending on the noise level in the experiment. For SC, its value is 0.0, 0.9, 0.9, 0.5, and 0.5 when σ = 0, 0.2, 0.4, 0.6, and 0.8, respectively. For IDSC, the corresponding value is 0.3 (this value is given in (Ling and Jacobs, 2007)), 0.8, 0.7, 1.0, and 0.6.

The parameter k has no significant influence to the final retrieval result when the value is in some proper range. Fig. 5 illustrates the changes of the retrieval result of the proposed method (DP_{min}) on MPEG-7 data set as a function of the parameter k. The retrieval rate remains stable above 89.3% when $2 \leqslant k \leqslant 9$.

Table 5
Retrieval results on Kimia's 99 data set (Sebastian et al., 2004) with contours perturbed by adding Gaussian noise with the parameter σ = 0, 0.2, 0.4, 0.6 and 0.8.

σ	Algorithm	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
0	SC	99	97	97	97	97	97	96	98	95	87
	IDSC (Ling and Jacobs, 2007)	99	99	99	98	98	97	97	98	94	79
	Height functions	**99**	**99**	**99**	**99**	**98**	**99**	**99**	**96**	**95**	**88**
0.2	SC	99	97	96	97	97	97	96	95	95	85
	IDSC	99	97	97	97	97	96	98	95	95	82
	Height functions	**99**	**99**	**99**	**99**	**99**	**98**	**98**	**95**	**94**	**88**
0.4	SC	99	97	96	97	96	98	96	96	94	83
	IDSC	99	97	97	97	97	97	96	96	94	82
	Height functions	**99**	**99**	**99**	**99**	**98**	**98**	**97**	**95**	**94**	**84**
0.6	SC	99	96	97	97	97	98	94	91	93	86
	IDSC	99	97	97	97	96	97	97	93	94	83
	Height functions	**99**	**99**	**99**	**99**	**98**	**97**	**97**	**94**	**91**	**80**
0.8	SC	99	97	97	97	96	97	97	95	91	80
	IDSC	99	97	97	97	97	97	98	94	92	80
	Height functions	**99**	**99**	**98**	**98**	**98**	**98**	**95**	**92**	**92**	**78**

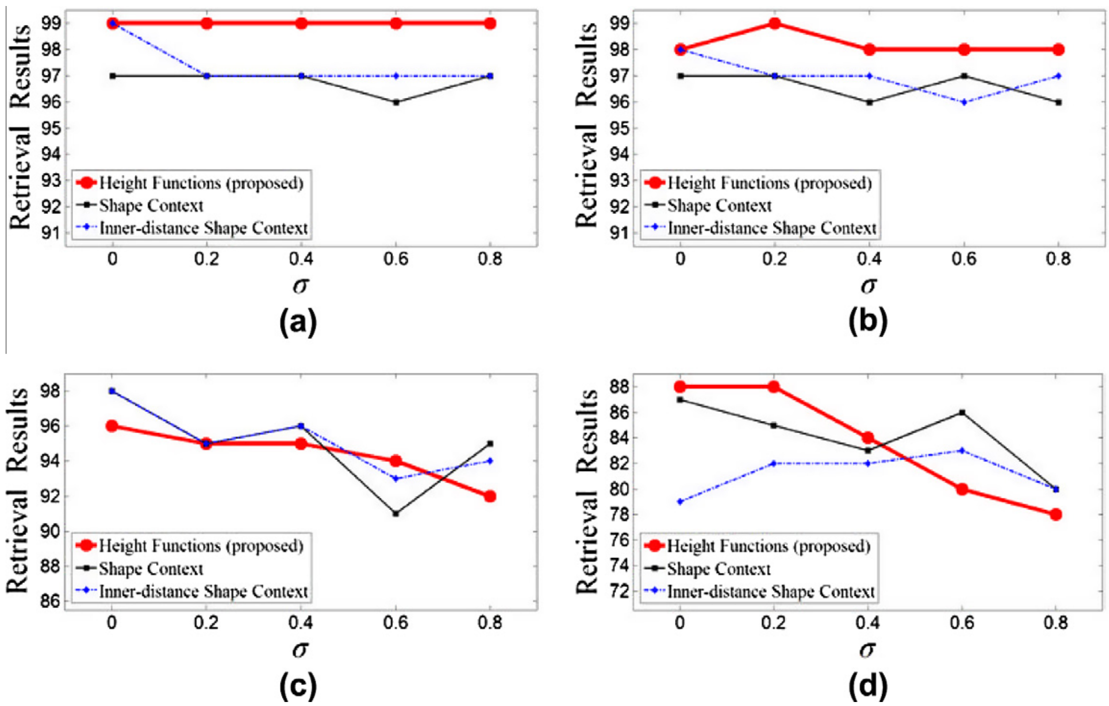

Fig. 4. The (a) 2nd, (b) 5th, (c) 8th, and (d) 10th top matches of the methods with respect to increasing noise levels on Kimia's 99 data set (Sebastian et al., 2004).

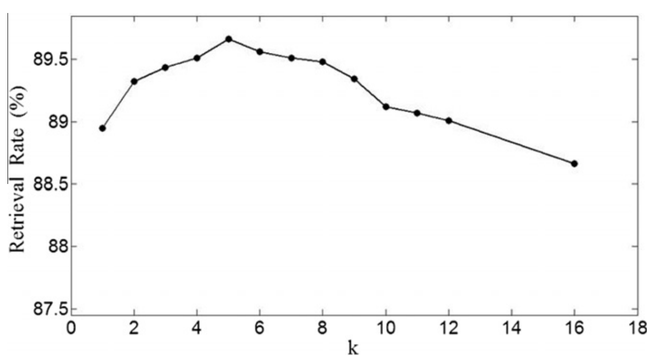

Fig. 5. The changes in retrieval rate of the proposed method (DP_{min}) on MPEG-7 data set when the parameter k changes.

4.6. Analysis of time complexity

The calculation of the shape descriptor and shape matching are two independent stages for the proposed method, and the complexity of height functions and matching are evaluated separately here. Assume there are N uniform sample points for every shape contour, k consecutive sample points when we smooth the original height functions, and M is the dimension number of the height vector.

To compute height functions, the signed distances from all sample points to every axis need to be calculated. According to Eq. (1), for the height vector of every sample point, there are $N-1$ distances to be calculated (the only exception is the distance of the sample point itself, as $h_i^1 = h_{i,i} = 0$ for every $i = 1,\ldots,N$), thus the time complexity of obtaining the original height vectors is $O(N(N-1)) = O(N^2)$. Then, the height vectors are smoothed according to Eq. (2), and there are M values to be calculated for every height vector. This process takes a complexity of $O(MN) = \frac{1}{k}O(N^2)$. Finally, the row wise normalization according to Eq. (5) for the $M \times N$ matrix defined by Eq. (3) takes a complexity of $O(MN) = \frac{1}{k}O(N^2)$. Thus the total computational complexity for calculating the height functions for every shape contour is $O(N^2) + \frac{1}{k}O(N^2) + \frac{1}{k}O(N^2) = (1 + \frac{2}{k})O(N^2)$. This complexity is comparable to that of SC (Belongie et al., 2002) and TAR (Alajlan et al., 2007, 2008), and is significantly less than that of IDSC, which is $O(N^3)$ (Ling and Jacobs, 2007).

In the matching stage, to match a pair of shapes, the time complexity of calculating feature distances according to Eq. (6) is $O(N^2)$, and then the DP algorithm takes a complexity of $O(N^3)$. The total computational complexity for the matching stage is $O(N^2) + O(N^3) = O(N^3)$, which equals the complexity of many approaches (such as MCC (Adamek and OĆonnor, 2004) and IDSC (Ling and Jacobs, 2007)).

5. Conclusions and future works

We presented a new shape representation and matching method based on the height functions of each sample contour point. The height function for one sample point includes the distances of all the other sample points to its tangent line. Height functions consist of not only the height value of every sample point, but also the order of the sample points along the shape contour. The smoothing process for the original height sequences makes the proposed descriptor more compact and insensitive to local deformations. Although the proposed descriptor is very simple, it achieves excellent retrieval results, which makes it attractive for adoption in different applications. The experiments on three popular benchmarks proved that the proposed method is effective under both geometric transformations and nonlinear deformations. In a word, the height functions is easy to implement, computationally efficient, and with an excellent performance in shape matching and retrieval tasks.

Several extensions of the proposed approach are possible. In this paper, height functions is only used for binary images to analyze the outer closed contour of objects. It is possible to include the inner contours in addition to the outer contour, which would require an order definition of the sample points. One simple way would be to order the sample points according to their projections on the tangent line of the reference point. Moreover, it is possible to apply height functions into the hierarchical matching frameworks such as (Felzenszwalb and Schwartz, 2007; Payet and Todorovic, 2009) to obtain a more confident shape similarity measure due to its compact and discriminative description.

Acknowledgements

The authors thank Ling and Jacobs for releasing their source code of IDSC online. Parts of our work (including boundary extraction, contour sampling and the dynamic programming algorithm) are accomplished by directly using corresponding parts of this code or slightly modifying them. The authors also would like to thank the anonymous referees who gave us many helpful comments and suggestions. This work was supported by the Fundamental Research Funds for the Central Universities' HUST 2011TS110, and National Natural Science Foundation of China #60903096 and #60873127.

References

Adamek, T., OĆonnor, N.E., 2004. A multiscale representation method for nonrigid shapes with a single closed contour. IEEE Trans. Circuits Systems Video Technol. 14 (5), 742–753.

Alajlan, N., Rube, I.E., Kamel, M.S., Freeman, G., 2007. Shape retrieval using triangle-area representation and dynamic space warping. Pattern Recognition 40, 1911–1920.

Alajlan, N., Kamel, M., Freeman, G., 2008. Geometry-based image retrieval in binary image databases. IEEE Trans. Pattern Anal. Machine Intell. 30 (6), 1003–1013.

Attalla, E., Siy, P., 2005. Robust shape similarity retrieval based on contour segmentation polygonal multiresolution and elastic matching. Pattern Recognition 38 (12), 2229–2241.

Bai, X., Latecki, L.J., 2008. Path similarity skeleton graph matching. IEEE Trans. Pattern Anal. Machine Intell. 30 (7), 1282–1292.

Bai, X., Wang, B., Wang, X., Liu, W., Tu, Z., 2010. Co-transduction for shape retrieval. In: ECCV: European Conf. on Computer Vision, pp. 328–341.

Bai, X., Yang, X., Latecki, L.J., Liu, W., Tu, Z., 2010. Learning context-sensitive shape similarity by graph transduction. IEEE Trans. Pattern Anal. Machine Intell. 32 (5), 861–874.

Belongie, S., Malik, J., Puzicha, J., 2002. Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Anal. Machine Intell. 24 (4), 522–705.

Daliri, M.R., Torre, V., 2008. Robust symbolic representation for shape recognition and retrieval. Pattern Recognition 41 (5), 1782–1798.

Direkoglu, C., Nixon, M.S., 2008. Shape classification using multiscale Fourier-based description in 2-D space. In: ICSP: IEEE Internat. Conf. on Signal Processing, pp. 820–823.

Ebrahim, Y., Ahmed, M., Abdelsalam, W., Chau, S.C., 2009. Shape representation and description using the Hilbert curve. Pattern Recognition Lett. 30, 348–358.

Egozi, A., Keller, Y., Guterman, H., 2010. Improving shape retrieval by spectral matching and meta similarity. IEEE Trans. Image Process. 19 (5), 1319–1327.

Felzenszwalb, P.F., Schwartz, J., 2007. Hierarchical matching of deformable shapes. In: CVPR: IEEE Internat. Conf. on Computer Vision and Pattern Recognition, pp. 1–8.

Grigorescu, C., Petkov, N., 2003. Distance sets for shape filters and shape recognition. IEEE Trans. Image Process. 12 (7), 729–739.

Jegou, H., Harzallah, H., Schmid, C., 2007. A contextual dissimilarity measure for accurate and efficient image search. In: CVPR: IEEE Internat. Conf. on Computer Vision and Pattern Recognition, pp. 1–8.

Jiang, H., Liu, W., Wang, D., Tian, C., Bai, X., Liu, X., Wu, Y., Liu, W., 2009. CASE: Connectivity-based skeleton extraction in wireless sensor networks. In: INFOCOM: IEEE Internat. Conf. on Computer Communications, pp. 2916–2920.

Kim, W.Y., Kim, Y.S., 2000. A region-based shape descriptor using Zernike moments. Signal Process.: Image Commun. 16 (1), 95–102.

Kontschieder, P., Donoser, M., Bischof, H., 2009. Beyond pairwise shape similarity analysis. In: ACCV: Asain Conf. on Computer Vision, pp. 655–666.

Latecki, L.J., Lakämper, R., Eckhardt, U., 2000. Shape descriptors for non-rigid shape with a single closed contour. In: CVPR: IEEE Internat. Conf. on Computer Vision and Pattern Recognition, pp. 424–429.

Leibe, B., Schiele, B., 2003. Analyzing appreance and contour based methods for object categorization. In: CVPR: IEEE Internat. Conf. on Computer Vision and Pattern Recognition, pp. 409–415.

Lewin, S., Jiang, X., Clausing, A., 2010. Perceptually motivated shape evolution with shape-preserving property. Pattern Recognition Lett. 31 (6), 447–453.

Lin, L., Zeng, K., Liu, X., Zhu, S., 2009. Layered graph matching by composite cluster sampling with collaborative and competitive interactions. In: CVPR: IEEE Internat. Conf. on Computer Vision and Pattern Recognition, pp. 1351–1358.

Ling, H., Jacobs, D., 2007. Shape classification using the inner-distance. IEEE Trans. Pattern Anal. Machine Intell. 29 (2), 286–299.

Ling, H., Yang, X., Latecki, L.J., 2010. Balancing deformability and discriminability for shape matching. In: ECCV: European Conf. on Computer Vision, pp. 411–424.

Liu, H., Latecki, L.J., Liu, W., 2008. A unified curvature definition for regular, polygonal, and digital planar curves. Internat. J. Comput. Vision 80 (1), 104–124.

McNeill, G., Vijayakumar, S., 2006. Hierarchical procrustes matching for shape retrieval. In: CVPR: IEEE Internat. Conf. on Computer Vision and Pattern Recognition, pp. 885–894.

Mokhtarian, F., Abbasi, S., Kittler, J., 1997. Efficient and robust retrieval by shape content through curvature scale space. IN: Smeulders, A.W.M., Jain, R. (Eds.) Image Databases and Multi-Media Search, pp. 51–58.

Nasreddine, K., Benzinou, A., Fablet, R., 2010. Variational shape matching for shape classification and retrieval. Pattern Recognition Lett. 31 (12), 1650–1657.

Payet, N., Todorovic, S., 2009. Matching hierarchies of deformable shapes. In: Proc. IAPR-TC-15 Workshop on Graph-based Representations in Pattern Recognition (GbR).

Schmidt, F.R., Toeppe, E., Cremers, D., 2009. Efficient planar graph cuts with applications in computer vision. In: CVPR: IEEE Internat. Conf. on Computer Vision and Pattern Recognition, pp. 351–356.

Sebastian, T.B., Klein, P.N., Kimia, B.B., 2004. Recognition of shapes by editing their shock graphs. IEEE Trans. Pattern Anal. Machine Intell. 25, 116–125.

Siddiqi, K., Shokoufandeh, A., Dickinson, S.J., Zucker, S.W., 1999. Shock graphs and shape matching. Internat. J. Comput. Vision 35 (1), 13–32.

Sun, K., Super, B., 2006. Classification of contour shapes using class segment sets. In: CVPR: IEEE Internat. Conf. on Computer Vision and Pattern Recognition, pp. 727–733.

Super, B., 2006. Retrieval from shape databases using chance probability functions and fixed correspondence. Internat. J. Pattern Recognition Artif. Intell. 20 (8), 1117–1138.

Tang, Y., You, X., 2003. Skeletonization of ribbon-like shapes based on a new wavelet function. IEEE Trans. Pattern Anal. Machine Intell. 25 (9), 1118–1133.

Temlyakov, A., Munsell, B.C., Waggoner, J.W., Wang, S., 2010. Two perceptually motivated strategies for shape classification. In: CVPR: IEEE Internat. Conf. on Computer Vision and Pattern Recognition, pp. 2289–2296.

Tu, Z., Yuille, A.L., 2004. Shape matching and recognition using generative models and informative features. In: ECCV: European Conf. on Computer Vision, pp. 195–209.

Wang, Z., Liang, M., 2010. Locally affine invariant descriptors for shape matching and retrieval. IEEE Signal Process. Lett. 17 (9), 803–806.

Wang, J., Li, Y., Bai, X., Zhang, Y., Wang, C., Tang, N., 2011. Learning context-sensitive similarity by shortest path propagation. Pattern Recognition 44 (10–11), 2367–2374.

Wolter, D., Latecki, L.J., 2004. Shape matching for robot mapping. In: PRICAI: Pacific Rim Internat. Conf. on Artificial Intelligence.

Xie, J., Heng, P., Shah, M., 2008. Shape matching and modeling using skeletal context. Pattern Recognition 41 (5), 1756–1767.

Xu, C., Liu, J., Tang, X., 2009. 2D shape matching by contour flexibility. IEEE Trans. Pattern Anal. Machine Intell. 31 (1), 180–186.

Yang, X., Bai, X., Latecki, L.J., Tu, Z., 2008. Improving shape retrieval by learning graph transduction. In: ECCV: European Conf. on Computer Vision, pp. 788–801.

Yang, X., Koknar-Tezel, S., Latecki, L.J., 2009. Locally constrained diffusion process on locally densified distance spaces with applications to shape retrieval. In: CVPR: IEEE Internat. Conf. on Computer Vision and Pattern Recognition, pp. 357–364.

You, X., Tang, Y., 2007. Wavelet-based approach to character skeleton. IEEE Trans. Image Process. 16 (5), 1220–1231.

Zhang, D.S., Lu, G., 2002. Generic Fourier descriptor for shape-based image retrieval. In: ICME: IEEE Internat. Conf. on Multimedia and Expo, pp. 425–428.