

Hierarchical String Cuts: A Translation, Rotation, Scale and Mirror Invariant Descriptor for Fast Shape Retrieval

Author

Wang, Bin, Gao, Yongsheng

Published

2014

Journal Title

IEEE Transactions on Image Processing

DOI

[10.1109/TIP.2014.2343457](https://doi.org/10.1109/TIP.2014.2343457)

Rights statement

© 2014 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Downloaded from

<http://hdl.handle.net/10072/67250>

Griffith Research Online

<https://research-repository.griffith.edu.au>

Hierarchical String Cuts: A Translation, Rotation, Scale and Mirror Invariant Descriptor for Fast Shape Retrieval

Bin Wang and Yongsheng Gao, *Senior Member, IEEE*

Abstract—This paper presents a novel approach for both fast and accurately retrieving similar shapes. A hierarchical string cuts (HSC) method is proposed to partition a shape into multiple level curve segments of different lengths from a point moving around the contour to describe the shape gradually and completely from the global information to the finest details. At each hierarchical level, the curve segments are cut by strings to extract features that characterize the geometric and distribution properties in that particular level of details. The translation, rotation, scale and mirror invariant HSC descriptor enables a fast metric based matching to achieve the desired high accuracy. Encouraging experimental results on four databases demonstrated that the proposed method can consistently achieve higher (or similar) retrieval accuracies than the state-of-the-art benchmarks with a more than 120 times faster speed. This may suggest a new way of developing shape retrieval techniques in which a high accuracy can be achieved by a fast metric matching algorithm without using the time-consuming correspondence optimisation strategy.

Index Terms—Shape description; shape retrieval; hierarchical string cuts

I INTRODUCTION

How to both quickly and effectively retrieve similar shapes from a large image database is an important and challenging problem that continues attracting attentions of many researchers in computer vision and pattern recognition. Its importance lies in that more and more practical applications, such as plant leaf image retrieval [43], fish image retrieval [20], trademark image retrieval [21], medical tumor shape retrieval [22], have encountered the speed and accuracy trade-off barrier. The challenges are twofold. (1) Nowadays, image databases have been growing larger and larger, which make the computational cost of shape retrieval become prohibitively expensive (e.g. there are about 400,000 species of plants all over the world [23] and millions of plant leaf images will be stored in the database). On the other hand, many real-time retrieval tasks (e.g. online shape retrieval) require the retrieval systems respond to the queries very quickly. Therefore,

retrieval efficiency is becoming a more important issue to be addressed in real applications. (2) The shapes in databases usually have large intra-class variations (see Fig. 1) and small inter-class differences (see Fig. 2), which make it very difficult for the retrieval system to achieve a desirable retrieval accuracy.



Fig. 1. Example leaves from the same plant species that have large intra-class shape variations.



Fig. 2. Five example leaves from different plant species that have small inter-class differences.

Shape description and matching are two key parts of shape retrieval. The former extracts effective and perceptually important shape features and organizes them in a data structure such as vector, string, tree and graph. While the later uses the obtained shape descriptors to determine the similarity (or dissimilarity) value of the two shapes to be compared based on a shape distance measure. The performance of any shape retrieval method ultimately depends on the type of shape descriptor used and the matching algorithm applied [31]. According to MPEG-7, a favourable shape descriptor should have a high discriminability so that it can group similar shapes together and separate dissimilar shapes into different groups, and a reliable shape descriptor should be rotation, scaling, and translation invariant [1].

In the past decade, various methods have been proposed for shape retrieval. To evaluate the performance of different methods, MPEG-7 set up a group of challenge datasets (MPEG-7 Core Experiment CE-shape-1) for objective experimental comparison [15],[32]. Many works of recent ten years [18],[13],[12],[5],[4],[3],[33],[34],[35],[37],[38] have reported good retrieval rates (>75%) on the MPEG-7 CE-1 Part B dataset, with some of the most recent ones [4], [5], [13], [38], [37] achieved very high retrieval rates (>85%). However, their high retrieval accuracies are obtained with high computational time cost as these algorithms mainly rely on the primitive correspondence techniques such as dynamic programming (DP). For example, the inner-distance shape context (IDSC) [5]

Manuscript received December 17, 2012. This work was supported in part by the Australian Research Council (ARC) under Discovery Grant DP0877929, and by the National Natural Science Foundation of China under Grant 61372158.

B. Wang is with Key Laboratory of Electronic Business, Nanjing University of Finance and Economics, Nanjing 210096, China (e-mail: wangbin@njue.edu.cn) and School of Engineering, Griffith University, Brisbane, QLD 4111, Australia (e-mail: bin.wang@griffith.edu.au).

Y. Gao is with School of Engineering, Griffith University, Brisbane, QLD 4111, Australia (e-mail: yongsheng.gao@griffith.edu.au).

achieved a very high retrieval rate of 85.40% using DP matching, whereas it only obtained 68.83% when using the shape context distance measure to replace the DP matching part. In theory, the DP based shape matching methods have the time complexity of $O(N^3)$ for matching two shape contours with N sample points [3]. Take $N = 100$ as an example. DP based inner-distance shape context (IDSC+DP) [5] and shape tree [13] spend 8.6 hours (on the reported 2.8GHz PC) and 13.9 hours (on the reported 3GHz computer) respectively to retrieve a shape in a large database with 100,000 images. Actually, almost all the existing methods reported a retrieval rate greater than 80% on the MPEG-7 CE-1 Part B dataset adopted time consuming DP matching. This makes the current methods with high accuracies not suitable for shape retrieval in large image databases and online retrieval, resulting in an accuracy-speed trade-off dilemma to users.

The above limitation of current algorithms motivates us to develop a novel shape description and matching method that considers both speed and accuracy in its design. A Hierarchical String Cuts (HSC) approach is proposed in this paper, which uses a hierarchical coarse to fine cutting and coding framework to describe and match shapes. Each subcurve of the shape is characterized by a set of geometric and distribution features that can capture more discriminative information of the shape than only using a single geometry feature such as curvature [39], angle pattern [12][48][49], pairwise distances [46][47], integral invariants [19][42] and triangle area [4][43][53]. The proposed HSC achieved 87.31% retrieval rate on the MPEG-7 CE-1 Part B dataset, the highest accuracy on leaf retrieval, leaf classification and the second highest retrieval accuracy on marine animal retrieval experiment, which were achieved with a speed of over 120 times faster than the state-of-the-art benchmarks.

The remainder of the paper is organized as follows: A brief review of related work is presented in Section 2. In Section 3, we describe the details of the proposed Hierarchical String Cuts (HSC) method. The computational complexity of HSC is analysed and compared to the state-of-the-art benchmark methods in Section 4. Four groups of experiments are presented and analysed in Section 5. Finally, we draw conclusions in Section 6.

II RELATED WORK

Various shape description and matching methods have been proposed in the past decades [10], [31]. They can be broadly classified into two categories, feature metric measurement method and primitive correspondence method. A feature metric measurement method extracts discriminative features of the two shapes to be matched to form two numeric vectors of a specified length, whose similarity or dissimilarity is measured by a metric distance such as L_1 or L_2 . In the contrary, a primitive correspondence method regards the shape as an aggregation of primitives, where primitives can be points, curve segments, line segments, etc. At the shape description stage, a set of features/attributes is extracted for each primitive, and at the shape matching stage, all pairs of primitives of the two shapes are first compared to generate a cost matrix (or

distance matrix). The least matching cost obtained by the optimal correspondence of the primitives between the two shapes is considered as their dissimilarity.

A. Feature metric measurement method

Fourier descriptors (FDs) [40] is a classical feature metric measurement method, in which a 2D shape contour is first represented as an 1D signature, such as complex coordinates [5], centroid distance [41], farthest point distance [9]. A discrete Fourier transform is then applied to the signature to generate a feature vector using Fourier coefficients. Zhang et al. [8] conducted a large amount of shape retrieval experiments to evaluate the performance of FDs derived from various shape signatures and reported that the FDs derived from the centroid distance signature had a higher retrieval performance and is more suitable for shape retrieval than using other shape signatures. Most recently, EI-ghaza et al. [26] treats the curvature-scale-space representation of a shape contour as a binary image and apply a 2D Fourier transform to it. The obtained descriptor captures the detailed dynamics of the shape curvature and the Euclidean distance metric is used for shape matching.

Other spectral transform based shape descriptors have also been proposed. Chuang et al. [17] used one-dimensional continuous wavelet transform to create a multi-scale shape representation. Yang et al. [44] proposed a wavelet descriptor independent of the starting-point location of a contour. As the wavelet coefficients are not invariant to the rotation of a shape, Kunttu et al. [7] applied Fourier transform to the wavelet coefficients and used the obtained Fourier coefficients to describe the shape. Wang et al. [30] used sequency-ordered complex Hadamard transform [27] to build a new shape descriptor.

The spectral transform based descriptors are very compact. This is because the energy of the signal concentrates on the low frequency components and only dozens of low frequency coefficients can describe the shape effectively. In addition, they are also robust to noise because the noise predominantly distributes in the high frequency coefficients which have been discarded in building the shape descriptor. However, these methods only use a single geometric feature such as curvature, centroid distance, and farthest point distance to construct 1D shape signature and lose detailed information that is discarded in the transform domain. This often results in weak discrimination ability of this type of shape recognition methods. According to the reported results on the MPEG-7 CE-1 Part B dataset [17], [30] which are also consistent with the results from our implementations of these algorithms (see Table 2), their retrieval rates are not better than 70%.

Efforts have also been made to yield shape descriptor directly from the spatial domain. These methods can be classified into curvature based and boundary point relationship based techniques. Curvature is a fundamental property of shape. The curvature scale space method (CSS) [52] is a well-known curvature based method, which has been recommended as one of the standard shape descriptor by MPEG-7 [11]. In CSS, Gaussian kernel with increasing standard deviation is used to

smooth the contour at different scales. The inflection points are located at each scale which results in a binary image, termed CSS image. The maxima of the CSS images' contour are used for shape matching. The typical method of computing curvatures is differential techniques. However it amplifies noise and is not stable. To address this issue, Manay et al. [19] used integral invariants instead of differential invariants to compute the curvature features of a shape contour. Most recently, Kumar et al. [42] used the area integral invariants combined with the arc-length integral invariants for plant leaf identification.

The features of the triangles [4][53] formed by shape boundary points are also used as an alternative of curvature. Rubé et al. [53] used the signed triangle area (TAR) to reflect the concave/convex properties of the shape boundary. The wavelet transform is used for smoothing and decomposing the shape boundaries into multiscale levels. The TAR image at each scale that is similar to the CSS image is used for shape matching. Mouine et al. [43] recently extended the triangle area representation into two new descriptors, termed triangle oriented angles (TOA) and triangle side lengths and angle representation (TSLA) which can provide more discriminative information than those of only using the area of triangles. The locality sensitive hashing [29] technique is used for matching.

Foteini et al. [49] used the multiscale angle code sequence incorporated with the Mutual Nearest Point Distance (MNPD) [28] for shape matching. The limitation of this method is its expensive computation cost. Most recently, Hu et al. [48] used the angles formed by the boundary point and its two neighbor points of equal distance from it to form an angle pattern (AP). The relationships of the neighbor APs are encoded into a binary string, termed binary angular pattern (BAP). The multiscale AP and BAP are used to build histogram for shape matching.

Boundary point relationship based methods focus on capturing the geometrical properties from the space relationship of the pairs of boundary points. Hu et al. [47] recently proposed a novel shape descriptor, termed multiscale distance matrix (MDM). In MDM, the distances between all possible pairs of shape boundary points are calculated to form a multiscale distance matrix, where scale is the arc length. Each row of the obtained matrix captures certain range of geometric properties. Through shifting and sorting operations, an invariant matrix is created for shape description. The L_1 metric is used for shape similarity measurement. Biswas et al. [46] used, for each pair of landmark points, the inner distance between two points, relative angle, contour distance and articulation-invariant center of mass associated with the pair of points to form a feature vector. Each of them is quantized and is mapped on the appropriate bins in the hash table. The numbers of feature vectors hashed to various bins are formed into a sequence. The χ^2 distance metric is used for shape similarity measure.

B. Primitive correspondence method

In recent years, many researchers dedicate to develop effective shape description and matching methods using correspondence based techniques to improve the recognition

accuracy. The well-known shape context (SC) [18] builds a histogram primitive for each contour point to describe the relative distribution between the point and the remaining points, which provides rich shape information for finding the best point correspondence between two shapes in a point-by-point manner. It reported 76.51% retrieval rate on the MPEG-7 CE-1 Part B dataset, and achieved 86.8% when dynamic programming (DP) is used in the matching process [14]. To make the shape context descriptor robust to articulation, Ling et al. [5] replace Euclidean distance with inner-distance (ID), which is defined as the shortest path between two contour points within the shape boundary, to extend the SC method to a novel shape description method termed inner-distance shape context (IDSC). IDSC achieved a promising retrieval rate of 85.4% on the MPEG-7 CE-1 Part B dataset.

Inspired by the observation that smoothing a closed contour makes the convex and concave points move inside and outside the contour, respectively, Adamek et al. [3] proposed a novel multi-scale convexity/concavity (MCC) method for shape matching. MCC used the relative displacement of a contour point with respect to its position at the preceding scale level to measure the convexity and concavity properties at a particular scale. The relative displacements of all the scale levels for each contour point are captured to calculate the distance between each pair of contour points. The resulting distance matrix is used to find the best one-to-one dense point correspondence. A retrieval rate of 84.93% is reported on the MPEG-7 CE-1 Part B dataset. Alajlan et al. [4] proposed another multi-scale shape descriptor, termed triangle area representation (TAR). It utilizes the areas of the triangles formed by the boundary points to measure the convexity/concavity of each contour point at different scales, where scale denotes the length of the arc which associates with the triangle formed by the contour point and its two neighbour points. The multi-scale triangle areas associated with each contour point are used to construct the distance matrix for finding the best one-to-one point correspondence. Tested on the MPEG-7 CE-1 Part B dataset, it reported a 85.03% of retrieval rate and achieved 87.23% if three global shape features aspect ratio, eccentricity and solidity are included in the shape matching.

To explicitly capture both global and local geometric properties of the shape of an object, Felzenszwalb et al. [13] proposed a shape tree method, in which a shape contour is modelled to a full binary tree by recursively dividing it into two subcurves of the same length. The dividing points are taken as the node of the tree. Each node of the tree stores the midpoint location of the subcurve relative to their start and end points. Those nodes at the bottom of the tree capture local geometric properties while the nodes near the root capture more global information. When matching curves A and B, a shape tree is built for A to look for a mapping from points in A to points in B such that the shape tree of A is deformed as little as possible. Shape tree is an interesting method and achieved a very high retrieval rate (87.70%) on MPEG-7 CE-1 Part B dataset.

Primitive correspondence methods, including [12], [33], [34], [35], [36], [37], [51], [38], and [50] are reported higher

discriminative capability and often achieved higher than 80% retrieval accuracy on the MPEG-7 CE-1 Part B dataset. However the primitive correspondence methods are computationally slow making them less practical in real world applications particular for large database retrieval, as they use optimisation algorithm such as dynamic programming to look for the best primitive correspondence between two shapes. Their computational complexities range from $O(N^2)$ to $O(N^4)$.

III THE PROPOSED METHOD

In this section, we first describe the proposed string cut process for characterizing a curve segment, and then introduce the hierarchical string cuts that can encode a shape completely from global and coarse information to fine local details. Next, the invariances of the proposed hierarchical string cuts descriptor are discussed. Finally, the dissimilarity measure is presented for matching shapes.

A. String Cut

A shape can be effectively represented by a sequence of points sampled from the contour on the object with uniform spacing [3], [4], [5], [13], [14]. The benefits of this representation is that it is not required to seek key-points such as points of maximum curvature and we can obtain as good an approximation to the underlying continuous shapes as desired by picking the number of sample points to be sufficiently large (see Fig. 3(a) as an example). Therefore, a shape φ can be described in a form of sample point sequence $\varphi = \{p_i(x_i, y_i), i = 1, \dots, N\}$, where i is the index of the sample point according to its order along the contour in counter-clockwise direction, (x_i, y_i) is the coordinates of the sample points p_i , and N is the total number of sample points on the contour. In the design of the proposed HSC, N is required to be in the power of two ($N=256$ in our experiments). Since the contour is a closed curve, we have $p_{N+i} = p_i$ and $p_{-i} = p_{N-i}$, for $i = 1, \dots, N$, i.e. the contour is treated as a periodic sequence with a period of N .

Curve partition by string: Let sequence $A_{ij} = \{p_i, p_{i+1}, \dots, p_j\}$ denote a piece of curve segment of the contour of the shape φ , which starts from the point p_i and ends at the point p_j . We define the straight line passing through the points p_i and p_j as the string ξ_{ij} to cut the curve segment. A_{ij} can be possibly

cut into three groups S_r, S_l and S_o , where S_r is the set of the points falling to the right side of the string ξ_{ij} , S_l is the set of the points falling to the left side of the string ξ_{ij} , and S_o is the set of the points falling on the string ξ_{ij} . Obviously we have $S_r \cup S_l \cup S_o = A_{ij}$, $S_r \cap S_l = \emptyset$, $S_r \cap S_o = \emptyset$ and $S_l \cap S_o = \emptyset$, i.e. $\{S_r, S_l, S_o\}$ is a string cut partition of curve segment A_{ij} .

Fig. 3 (b) illustrates an example of the string cut process on a curve segment of horse shape from the MPEG-7 database.

String Cut Features: Given a piece of curve segment A_{ij} , let $\{S_r, S_l, S_o\}$ be its string cuts. The geometric properties of the curve segment A_{ij} can be depicted by the distributions of these cuts including deviations to string ξ_{ij} on both sides (f_1 and f_2), imbalance of cut (f_3), and degree of bending (f_4) as defined below.

$$f_1 = \max \left(\frac{1}{T_r} \sum_{p_k \in S_r} h(p_k, \xi_{ij}), \frac{1}{T_l} \sum_{p_k \in S_l} h(p_k, \xi_{ij}) \right), \quad (1)$$

$$f_2 = \min \left(\frac{1}{T_r} \sum_{p_k \in S_r} h(p_k, \xi_{ij}), \frac{1}{T_l} \sum_{p_k \in S_l} h(p_k, \xi_{ij}) \right), \quad (2)$$

$$f_3 = T_r - T_l, \quad (3)$$

$$f_4 = \frac{L_{ij}}{d(p_i, p_j)}, \quad (4)$$

where T_r , T_l and T_o are the number of sample point in S_r , S_l and S_o respectively, L_{ij} is the length of the curve segment, $h(p_k, \xi_{ij})$ is the perpendicular distance from point p_k to the string ξ_{ij} which can be calculated by

$$h(p_k, \xi_{ij}) = \frac{|(x_k - x_i)(y_j - y_i) - (y_k - y_i)(x_j - x_i)|}{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}}, \quad (5)$$

and $d(p_i, p_j)$ denotes the Euclidean distance between the points p_i and p_j .

These string cut features as a whole express the configuration and the behavior of the entire curve segment relative to the reference string and also ensure their invariance to the possible swapping of start and end points of the curve segment (see invariance analysis in Section C).

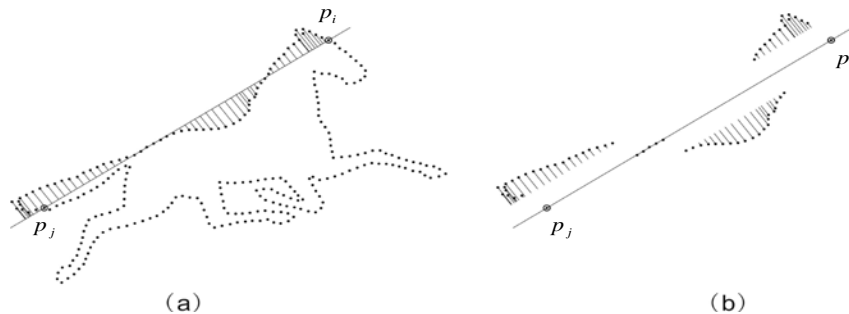


Fig. 3. String and string cut. (a) a curve segment of horse contour and its string. (b) The string cuts the curve segment into three groups; one falls on the string, and the other two fall to the two sides of the string, respectively.

B. Hierarchical String Cuts and Their Signatures

The string cuts on longer curve segments capture the more global shape information, while those on shorter ones are associated with the finer details of the curves. Here we propose a Hierarchical String Cuts (HSC) method that divides a closed contour into curve segments of different lengths, to provide a multiple level coarse-to-fine description.

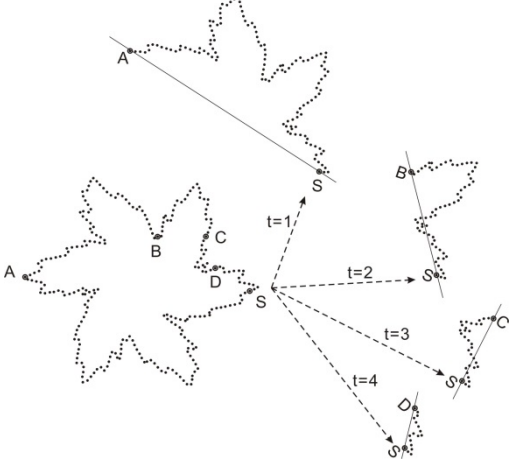


Fig.4. An example result of hierarchical string cuts for a point S at hierarchical levels $t = 1, \dots, 4$.

At hierarchical level t , for each sample point p_i , $i = 1, \dots, N$, we take a piece of curve segment $A_{ij} = \{p_i, p_{i+1}, \dots, p_j\}$ starting from p_i and ending at p_j . The location of the end point p_j and the length of the curve segment are determined by the hierarchical level t using $j = i + \frac{N}{2^t}$.

The string cut features for point p_i at level t , $\{f_1^t(i), f_2^t(i), f_3^t(i), f_4^t(i)\}$, can be calculated using Eqs (1-4). Thus, for each hierarchical level t , we obtain four N -dimension Sting Cut Signatures $f_1^t(i), f_2^t(i), f_3^t(i), f_4^t(i)$, $i = 1, \dots, N$ to describe the shape in a coarse/fine level controlled by t . Each string cut signature is a sequence (with a length of N) of string cut features. As the number of sample points on the contour φ is N , there are $\log N - 1$ levels for $t = 1, \dots, \log N - 1$ that divide φ into curve segments from the coarsest $(\frac{N}{2} + 1)$ -point length to the finest 3-point length. Fig. 4 gives an example of the hierarchical string cut process for a single contour point S that divide the shape into curve segments of different lengths at the coarsest hierarchical levels $t = 1, \dots, 4$.

C. Invariance of Shape Descriptor

A good shape descriptor should be translation, scale, rotation and mirror invariant. Since the string cut features of a curve segment are calculated solely with respect to its string, they have the intrinsic invariance to translation; and so do the string cut signatures that are derived from these features. To make the string cut signatures invariant to scaling, each signature $f_g^t(i)$,

$g = 1, \dots, 4$ is locally normalised by dividing its maximum value $\max_{i=1, \dots, N} f_g^t(i)$. It should be pointed out that since $f_3^t(i)$ can take negative values, it is normalized by $\max_{i=1, \dots, N} |f_3^t(i)|$.

When $\max_{i=1, \dots, N} f_g^t(i) = 0$, the above normalisation division cannot be executed. In this case, the normalization division is omitted in implementing the algorithm as the signature with all 0's is already invariant to scaling.

At hierarchical level $t = 1, \dots, \log N - 1$, for shape signature $f_g^t(i)$, $g = 1, \dots, 4$, the magnitudes of their Fourier transform coefficients are calculated by

$$F_g^t(k) = \frac{1}{N} \left| \sum_{i=0}^{N-1} f_g^t(i) \exp\left(\frac{-j2\pi ik}{N}\right) \right|, \quad k = 0, \dots, N-1. \quad (6)$$

From theory, the above obtained $F_g^t(k)$, $k = 0, \dots, N-1$ are invariant to the location of start point p_i of string cut curve segments, and thus invariant to rotation of the whole shape. To make the generated shape descriptor robust to noise and compact, the lowest M order coefficients $F_g^t(k)$, $t = 0, \dots, M-1$ where $M \ll N$ are used to describe the shape.

Finally the hierarchical string cuts (HSC) shape descriptor $\psi = \{\psi_1, \dots, \psi_{\log N - 1}\}$ is a combination of descriptors ψ_t from all hierarchical levels of t defined as

$$\psi_t = \{F_g^t(k), \sigma_g^t \mid g = 1, \dots, 4; k = 0, \dots, M-1\}, \quad (7)$$

where σ_g^t is the standard deviation of the string cut features. It contains supplementary information to $F_g^t(0)$ (which is the mean value) for each signature f_g^t .

Mirror invariant matching is one of the requirements for planar shape recognition methods. Most of the existing techniques [3], [4], [5] address this invariance problem by performing the same shape matching algorithm twice, one is between shapes A and B and the other is between shape A and the mirrored shape of B. To avoid the computational cost of applying the matching algorithm on the mirrored shape, the proposed HSC shape descriptor is designed intrinsically as mirror invariant, that is, $\psi = \{\psi_1, \dots, \psi_{\log N - 1}\}$ remains the same if the start point p_i and the end point p_j are interchanged. In Eqs (1) and (2), f_1 and f_2 are designed as the deviations to its string on both sides according to their values instead of locations, ensuring their invariance to the interchange of p_i and p_j . Imbalance of cut f_3 changes sign when p_i and p_j are interchanged to make its shape signature $f_3^t(i)$ function only have a 180° phase difference, resulting a same $F_3^t(k)$, $k = 0, \dots, N-1$. As f_4 is also independent to the direction of shape contour, the HSC shape descriptor is mirror invariant.

D. Shape Dissimilarity Measure

Given two HSC descriptors $\psi(A) = \{F_g^{(A)t}(k), \sigma_g^{(A)t}\}$ and $\psi(B) = \{F_g^{(B)t}(k), \sigma_g^{(B)t}\}$, extracted from shapes A and shape B , respectively, where $g = 1, \dots, 4$, $k = 0, \dots, M-1$ and $t = 1, \dots, \log N - 1$, we first compare them in each hierarchical level $t = 1, \dots, \log N - 1$ by calculate the sub-level dissimilarity at hierarchical level t using Eq. (8):

$$D_t(A, B) = \sum_{g=1}^4 w_g \left(\sum_{k=0}^{M-1} \left| F_g^{(A)t}(k) - F_g^{(B)t}(k) \right| + \left| \sigma_g^{(A)t} - \sigma_g^{(B)t} \right| \right), \quad (8)$$

where w_g are the weights to allow the adjustment of contribution from each string cut feature. The dissimilarity between two shapes A and B can be easily considered as an aggregation of $\log N - 1$ sub-level dissimilarities as

$$D(A, B) = \sum_{t=1}^{\log N - 1} D_t(A, B). \quad (9)$$

However, the case of $t = 0$ is not considered yet in the design, in which the length of subcurve equals to N , i.e. the curve segment is the whole shape contour. Many contour global features such as circularity, eccentricity, convexity, ratio of principle axis, elliptic variance, circular variance, have already been developed for shape description. To complete the design of the proposed approach, here, we take the two existing global contour features, eccentricity (E) and rectangularity (R), as the features of hierarchical level $t = 0$. Without losing generality, other global contour features can be used as well. The dissimilarity of hierarchical level 0 is:

$$D_0 = \left| E^{(A)} - E^{(B)} \right| + \left| R^{(A)} - R^{(B)} \right|. \quad (10)$$

And the overall dissimilarity between two shapes A and B becomes

$$D(A, B) = \sum_{t=0}^{\log N - 1} D_t(A, B). \quad (11)$$

IV COMPUTATIONAL COMPLEXITY

The computational cost of shape retrieval consists of two parts. One is the time for computing the shape descriptor, and the other is that used for performing shape dissimilarity measurement. The later part is more important and plays a dominant role in determining the retrieval speed than the former one, particularly when the size of database becomes larger. This is because for a shape retrieval task, only the query shape is required to create its descriptor online and all the descriptors of gallery shapes can be calculated offline to be stored in the database beforehand; whereas shape dissimilarity

measurements are conducted online between the query descriptor and all gallery descriptors in the database. In this section, we provide a theoretical computational complexity analysis for the proposed HSC in comparison with the state-of-the-art methods (Experimental comparisons in computational time are also conducted in Section V).

In creating the HSC descriptor, for each hierarchical level $t = 1, \dots, \log N - 1$, the complexity of calculating the string cut

signatures for the whole contour is $O\left(\frac{N^2}{2^t}\right)$ because the time

to compute the string cut features for each curve segment

$A_{ij} = \{p_i, p_{i+1}, \dots, p_j\}$ is $O\left(\frac{N}{2^t}\right)$, where N is the number of

the sample points of the contour and $j = i + \frac{N}{2^t}$. Then the

complexity of calculating the signatures for all the hierarchical levels is

$$O\left(N^2 \left(\frac{1}{2^1} + \dots + \frac{1}{2^K} \right)\right) = O\left(N^2 \left(1 - \frac{1}{2^K}\right)\right) = O(N^2). \quad (12)$$

For each obtained signature ($4(\log N - 1)$ signatures in total), the complexity of calculating its Fourier transform coefficients and standard deviation are $O(N \log N)$ and $O(N)$ respectively. Therefore the time of computing the Fourier coefficients and standard deviations for all the signatures is

$$O(4(\log N - 1)(N \log N + N)) = O(N \log^2 N + N \log N) = O(N \log^2 N). \quad (13)$$

The overall complexity of creating the proposed HSC shape descriptor is $O(N^2 + N \log^2 N)$.

At the dissimilarity measurement part, the complexity of calculating Eq. (8) is $O(4(M+1)) = O(M)$. Therefore, the complexity of calculating Eq. (9) is

$$O(M(\log N - 1)) = O(M \log N). \quad (14)$$

Since calculating Eq. (10) only requires time $O(1)$, the time of computing Eq. (11), i.e. the overall computational complexity of dissimilarity measurement is $O(M \log N)$. In Table I, we compare our method with several recent representative methods which are reported high retrieval accuracies (>80%) on the MPEG-7 Part B retrieval test, where N is the number of sample points of the contour, and K (for IDSC+DP [5] and SC+DP [14]) denotes the number of possible starting points for alignment used in their dynamic programming part. From this table, we can see that the proposed HSC has the lowest order of computational complexity among the competing methods with high recognition accuracies.

TABLE I. Comparison of the computational complexity of different methods at the stage of shape dissimilarity measurement.

Proposed HSC	MCC [3]	TAR+DP [4]	IDSC+DP [5]	SC+DP [14]	Shape Tree [13]
$O(M \log N)$	$O(N^3)$	$O(N^3)$	$O(KN^2)$	$O(KN^2)$	$O(N^4)$

V EXPERIMENTS

To evaluate the effectiveness and efficiency of the proposed

approach, an extensive experimental investigation is conducted on MPEG-7 shape, plant leaf shape, and marine animal shape databases. The performances of the proposed method are

compared with the state-of-the-art approaches in both accuracy and speed. In all the experiments, the same parameters ($M = 7$, $w_1 = 1.4$, $w_2 = 0.5$, $w_3 = 0.5$, $w_4 = 0.4$) that is heuristically determined, are used without tuning the system setting to best suit individual dataset.

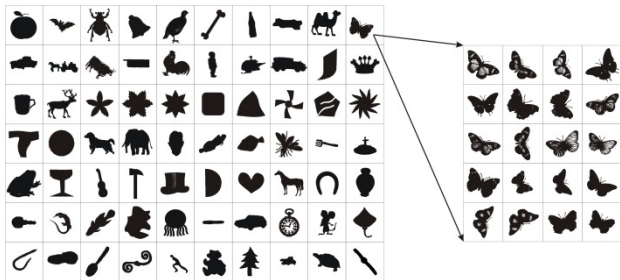


Fig. 5. MPEG-7 Part B shape database, which contains 70 classes with 20 images in each class.

A. MPEG-7 Shape Database

The MPEG-7 CE-1 Part B shape database [15] is a widely used dataset for evaluating performances of similarity-based shape retrieval. It contains 1400 shape images, consisting of 70 classes of various shapes with 20 images in each class (see Fig. 5). The retrieval accuracy is measured using the well-known “bull’s-eye test” [15], [3], [4], [5], [14]. In this measurement, each shape is used in turn as a query and matched with all the shapes in the database. The number of correct matches (that is the retrieved shape and the query shape belong to the same class) in the top $2 \times 20 = 40$ retrieved shapes that have the smallest dissimilarity values are counted. Since the maximum number of correct matches for a single shape is 20, the total number of correct matches is $1400 \times 20 = 28000$. The percentage of matched shapes out of 28000 is the retrieval rate of the bull’s-eye test. The computational time of each shape retrieval is the time of matching the query with all the 1400 shapes including the feature extraction time of the query shape. To finely compare the behavior of the proposed HSC against benchmarks, their Precision-Recall (PR) curves are provided.

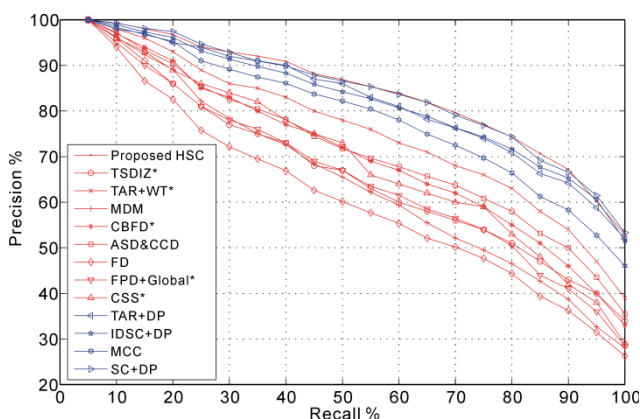


Fig. 6. Precision-Recall (PR) curves of the proposed HSC and the state-of-the-art approaches, including TSDIZ [2], TAR+WT[53], MDM [47], CBFDP [26], ASD&CCD [49], FD [6][8], FPD+Global [9], CSS [52], TAR+DP [4], IDSC+DP [5], MCC [3] and SC+DP [4][18], obtained on the MPEG-7 Part B shape database. The non-DP methods are plotted in red, while the DP methods are plotted in blue. * indicates that the results are from the published papers.

The retrieval rate and speed of the proposed Hierarchical String Cuts (HSC) approach together with those of the

state-of-the-art approaches, including the well-known inner distance (IDSC) [5], shape contexts (SC) [18] and their variants are tabulated in Table II. Their Precision-Recall (PR) curves are plotted in Fig. 6. Fig. 7 provides 4 detailed examples that list the top 10 retrieved shapes and their scores obtained by the proposed HSC, IDSC [5], SC [18].

Table II. Retrieval results on the MPEG-7 database. The values with * are from the published papers.

Algorithm		Retrieval rate (%)	Average retrieval time (ms)
DP	MCC [3]	84.93*	2.30×10^4
	TAR+DP [4]	87.13*	2.77×10^4
	IDSC+DP [5]	85.40*	1.81×10^4
	SC+DP [14][18] [†]	86.80*	2.17×10^4
	Shape Tree [13]	87.70*	NA
Non-DP	Proposed HSC	87.31	1.40×10^2
	MDM [47]	69.19	1.61×10^2
	ASD & CCD [49]	76.20*	1.17×10^4
	FD[6][8]	67.94	2.64×10^1
	FPD [9]	64.29	3.43×10^1
	IDSC+SC distance measure [5]	68.83*	NA
	SC+SC distance measure [5]	64.59*	NA
	WTD [17]	67.76*	NA

It can be seen that the proposed approach achieved similar retrieval accuracy to inner distance with dynamic programming (IDSC+DP), shape contexts with dynamic programming (SC+DP), and other dynamic programming based algorithms with the highest reported accuracies. The 87.31% of retrieval rate of the proposed HSC is slightly higher than the renowned IDSC+DP (85.40%), SC+DP (86.80%), but lower than that of Shape Tree method [13] by 0.30%, placing it among the most accurate shape retrieval algorithms. On the other hand, the above accuracy is obtained in a speed over 120 times faster than those methods with comparable accuracies (only 0.61%, 0.51%, 0.77% and 0.65% of the retrieval time used by MCC [3], TAR+DP [4], IDSC+DP [5] and SC+DP [14], [18]). In our experiments, all the algorithms are written in Matlab and are run on a PC with Intel Core-2 Duo 2.8 GHz CPU and 2 GB DDR2 RAM under Windows XP. Due to the very large computation cost in dynamic programming (DP) part of the benchmark algorithms, the DP part in MCC [3], TAR+DP [4], IDSC+DP [5] and SC+DP [14], [18] are implemented in C allowing the comparative experiments to be completed in an acceptable time. Note that the retrieval speed of Shape Tree [13] was not reported. But, as its computational complexity is $O(N^4)$ (see Table I), it is computationally more expensive than the other four benchmarks (whose complexities are $O(N^3)$ and $O(KN^2)$ respectively).

The performances of some fast non-DP shape matching algorithms are also listed in Table II for comparison purpose, whose accuracies are much lower than the proposed method and the state-of-the-art benchmarks by around 20%.

[†] [14] replaced the thin plate spline matching process used in [18] with Dynamic Programming, which increased the retrieval rate from 76.51% [18] to 86.8%.

HSC				SC+DP				IDSC+DP			
2.96	4.66	2.05	1.65	42.5	37.9	55.9	24.8	47.9	65.7	67.6	45.4
3.14	4.72	2.42	2.07	44.1	44.2	56.5	26.1	53.4	68.9	70.0	50.6
3.37	4.78	2.90	2.41	54.3	54.3	56.9	26.6	68.5	69.9	70.1	56.9
3.55	5.24	3.28	3.83	67.8	55.0	57.1	63.8	69.5	72.3	70.5	66.5
3.68	5.31	3.38	3.93	67.9	55.0	57.1	66.1	70.3	73.9	70.5	67.8
3.73	5.32	3.58	4.11	68.1	58.4	57.2	67.1	72.4	74.1	70.7	68.2
3.78	5.35	3.65	4.83	68.3	58.6	57.4	67.5	73.5	74.6	70.8	69.2
3.89	5.50	3.71	5.01	68.3	61.9	58.5	67.9	74.0	75.1	70.9	69.6
3.94	5.52	3.80	5.17	68.4	62.5	58.8	68.0	74.2	76.2	71.0	70.2
4.00	5.53	3.81	5.18	68.5	65.2	59.2	68.1	74.2	76.7	71.4	70.3

Fig. 7. The comparative retrieval results of 4 shapes obtained by the proposed HSC, SC+DP [18], and IDSC+DP [5]. Each column shows the query shape (top image) and the 10 most similar shapes retrieved from the MPEG-7 database (the 2nd-11th images). The number below each retrieved shape is its dissimilarity score. The incorrect hits are circled in red colour.

Fig. 7 gives the retrieval results of four typical shapes obtained using the SC+DP, IDSC+DP and the proposed approach. The results are listed and sorted in ascending order of dissimilarity (the first 10 ranked matches are shown). For the first query shape with small occlusions, the proposed method obtains 7 correct hits which is more than 5 for SC+DP and 2 for IDSC+DP, indicating the more robustness of the proposed approach to small occlusions than SC and IDSC. For the second shape which have larger occlusions, the performance of the proposed method decreases drastically (only obtains 2 correct hits), while the SC and IDSC are stable which indicates that SC and IDSC are more robust to larger occlusions than the proposed method. For the third and the fourth query shapes with large intra-class variations, the proposed method obtains 9 and 7 correct hits respectively which are much better than 0 and 4 for SC, and 3 and 6 for IDSC. These comparative results shows that the proposed approach, without using dynamic programming, can achieve a similar/better accuracy as SC and IDSC with an over 120 times faster speed for large database on-line retrieval.

B. Plant Leaf Databases

In this section, we apply the proposed HSC on a real and challenging application, plant leaf image retrieval and classification, and compare its performance with the

benchmark approaches. The challenge of this application lies in the high inter-class similarity between some of the leaf shapes and the large intra-class variations of plant leaves from the same species.

Leaf retrieval

The same performance evaluation method as used in MPEG-7 shape dataset is applied on a plant leaf image database collected by ourselves. It contains 1200 leaf images collected from 100 plant species, with 12 different leaves from each class (see Fig. 8).

Table III shows the comparative results of the proposed Hierarchical String Cuts (HSC) approach and the state-of-the-art benchmarks including the well-known inner distance (IDSC+DP) [5] and shape contexts (SC+DP) [14], [18]. The Precision-Recall (PR) curves of these methods are plotted in Fig. 9. It can be seen that the proposed approach achieves the highest retrieval rate and obtains the best Precision-Recall (PR) curve among all the competing methods. The 89.40% of retrieval rate of the proposed HSC is 3.76% and 2.58% higher than the inner distance (IDSC) [5] and the shape contexts [14][18] respectively, and is 12.30% and 11.74% higher than the other dynamic programming based algorithms MCC [3] and TAR+DP [4] respectively. It is worth noting that this high accuracy is achieved with a speed of over 125 times faster than those dynamic programming based methods (only 0.61%, 0.54%, 0.80% and 0.70% of the retrieval time used by MCC [3], TAR+DP [4], IDSC+DP [5] and SC+DP [14], [18], respectively). The performances of some fast shape matching algorithms are also listed in Table III for comparison purpose, whose accuracies are much lower than the proposed method by around 25%.

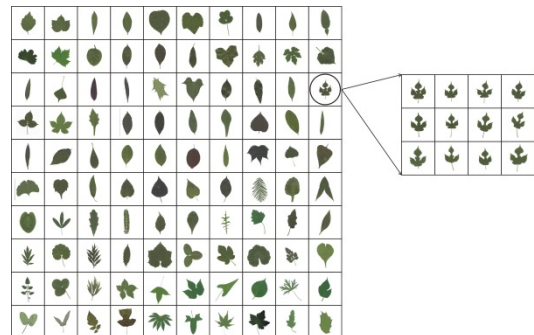


Fig. 8. Plant leaf image database collected by ourselves, which contains 100 species with 12 images in each class.

Table III. Retrieval results on the plant leaf image database.

Algorithm		Retrieval rate (%)	Average retrieval time (ms)
DP	MCC [3]	77.10	1.97×10^4
	TAR+DP [4]	77.66	2.25×10^4
	IDSC+DP[5]	85.64	1.52×10^4
	SC+DP [14] [18]	86.82	1.72×10^4
Non-DP	ASD&CCD [49]	74.97	9.83×10^3
	FD [6][8]	59.30	2.32×10^1
	MDM [47]	64.84	1.37×10^2
	Proposed HSC	89.40	1.21×10^2

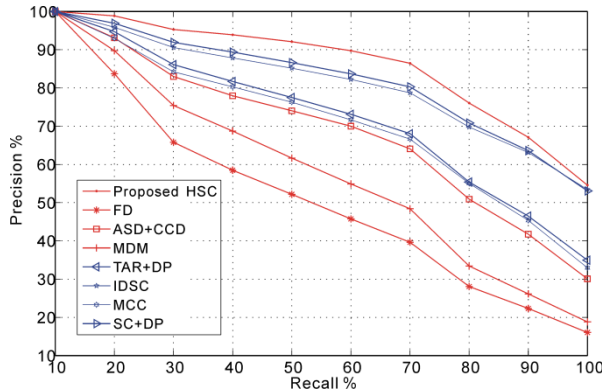


Fig. 9. The Precision-Recall (PR) curves of the proposed HSC and the state-of-the-art approaches on the plant leaf dataset collected by ourselves.

Leaf classification

The publicly available Swedish leaf database, which comes from a leaf classification project at Linköping University and the Swedish Museum of Natural History [45] is also used in our experiments. The database contains 15 different Swedish tree species (see Fig. 10) with 75 samples in each species. Due to the nature of this database (small number of classes and large number of samples in each class), the same leaf classification protocol and accuracy measurement as used in [5], [13], [45] are adopted in our experiment. 25 images randomly selected from each species are used as models and the remaining images are used as testing images. The classification rate is calculated using the nearest-neighbor classification rule. The comparative results of the proposed Hierarchical String Cuts (HSC) approach with the state-of-the-art approaches are listed in Table IV. From Table IV, it can be seen that the proposed approach maintained the highest classification accuracy (96.91%), consistent with the previous results in retrieval setting.



Fig. 10. Fifteen samples (one per species) from the Swedish leaf database.

TABLE IV. Classification rates on the Swedish leaf database. The values with * are from the published papers.

Algorithm	Classification rate (%)
MCC [3]	94.75
TAR+DP [4]	95.97
IDSC+DP [5]	94.13*
SC+DP [5]	88.12*
Shape Tree [13]	96.28*
TSLA+LSH [43]	96.53*
MDM+ Dimensionally reduction [47]	93.60*
TAR+LSH [43]	90.40*
ASD+CCD[49]	93.08
Proposed HSC	96.91

C. Marine Database

The marine database was originally built by Mokhtarian *et al.* [24]. It consists of 1100 marine animals (samples are shown in Fig. 11 (b)) which are unclassified. In MPEG-7 (Part C), this

dataset is combined with 200 breem fish images which are frames extracted from a video clip of a swimming breem fish (examples are shown in Fig. 11 (a)) to form a database consisting of $1100+200=1300$ images for testing the robustness to shape changes caused by non-rigid motion and different viewing angle [15]. In our experiment, each of the 200 breem fish images is used as a query in turn to match against the 1300 images in the database. The number of breem fish images in the top 200 most similar shapes is counted, and the percentage of the matched shapes out of 200 is calculated as retrieval rate (see also [12], [16], [17], [25], [26]).

The comparative results of the proposed Hierarchical String Cuts (HSC) approach with MCC [3], TAR+DP [4], IDSC+DP [5], and SC+DP [14],[18] are given in Table V. Considering the large shape variations caused by non-rigid motion and different viewing direction (see Fig. 11 (a)), it is very encouraging to see that the proposed HSC possess the similar level of tolerance to shape distortions as the best performing approaches, while its computational speed is more than 133 times faster than the competing algorithms (188.3, 184.7, 133.8 and 165.1 times faster than MCC, TAR+DP, IDSC+DP and SC+DP respectively). The average retrieval rate (over 200 tests by using every breem fish shape as the query in each test) for HSC remains higher than MCC, TAR+DP and Shape Context. It also worth comment that our result shows that the Inner Distance approach is the most robust to distortions (which support the claim in [5]) as we can see that IDSC+DP becomes the most accurate one in this experiment, while its accuracy is ranked the second and the third in the leaf databases experiments.

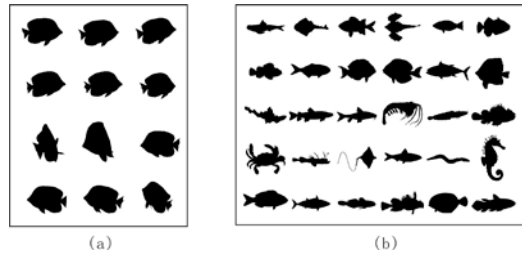


Fig. 11. Samples from 200 breem fishes (a) and 1100 marine animals (b).

TABLE V. Retrieval results on the marine database.

Algorithm	Average retrieval rate (%)	Average retrieval time (ms)
MCC [3]	88.64	2.12×10^4
TAR+DP [4]	92.15	2.08×10^4
IDSC+DP [5]	95.84	1.51×10^4
SC+DP [14],[18]	91.93	1.86×10^4
Proposed HSC	94.02	1.12×10^2

D. Sensitivity to Noise

To evaluate the sensitivity to noise, a group of experiments with different levels of additive noise are conducted. We adopt the same noise perturbation scheme as used in [26][9][19][50]. For each shape image in MPEG-7 Part B dataset, its coordinates (x, y) of the boundary points are randomly varied by a Gaussian noise with different variance. The noise corruption of shape is measured by the signal-to-noise ratio (SNR) as

$$SNR = 10 \log \frac{\sigma_s^2}{\sigma_n^2}, \quad (16)$$

where σ_s^2 and σ_n^2 are the variances of the signal and noise sequences, respectively. In our experiments, various levels of Gaussian noise (SNR=20dB-50dB) are added to the query shapes in MPEG-7 Part B dataset, respectively. Fig. 12 summarizes the accuracies of retrieving shapes corrupted by various levels of noise. It can be seen that the accuracy is not affected when the additive noise is at the level of SNR=50dB, and then drops gracefully when noise increases to SNR=30dB. When the shapes are heavily corrupted to the level of SNR=20dB (see the leftmost corrupted shape), the proposed HSC still maintained an accuracy over 70%.

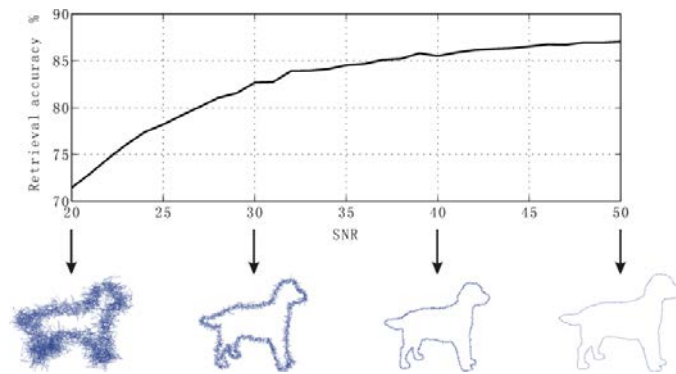


Fig. 12. Retrieval accuracy of the proposed method on MPEG-7 Part B shapes corrupted by different levels of Gaussian noise.

VI CONCLUSION

In this paper, we have presented a new hierarchical string cuts (HSC) approach to both fast and effectively retrieve similar shapes. From a start point moving along the shape contour, the shape is partitioned into multiple level curve segments of different lengths, which are cut by their corresponding strings for extracting their geometric and distribution properties. At the coarse hierarchical levels, the string cut features of longer curve segments describe the global properties of the shape; while at the fine levels, the features extracted from the shorter ones depict the detailed information of the shape. The proposed HSC descriptor is a translation, rotation, scaling and mirror invariant shape descriptor. The dissimilarity of two shapes can be efficiently measured by comparing their string cut signatures of all hierarchical levels using the L_1 distance without the need of using the time costly dynamic programming algorithm. The performance of the proposed HSC has been evaluated extensively on three well-known shape databases and a leaf database of 100 plant species collected by ourselves. The experimental results show that the proposed HSC method can consistently achieve the same level of highest retrieval accuracy as the state-of-the-art methods with an over 120 times faster speed. This indicates the potential of metric based shape matching strategy in developing both fast and accurate algorithms for large database retrieval and/or online retrieval.

REFERENCES

[1] H. Kim, J. Kim, Region-based shape descriptor invariant to rotation, scale and translation, *Signal Process: Image Commun.*, 16 (2000) 87-93.

[2] F.A. Andaló, P.A.V. Miranda, R. da S. Torres, A.X. Falcão, Shape feature extraction and description based on tensor scale, *Pattern recognition* 43 (2010) 26-36.

[3] T. Adamek, N.E. O'Connor, A multiscale representation method for non-rigid shape with a single closed contour, *IEEE Trans. on Circuits and systems for video technology* 14 (2004) 742-743.

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

[5] H. Ling, D.W. Jacobs, Shape classification using the inner-distance, *IEEE Trans. Pattern Anal. Machine Intell.* 29 (2007) 286-299.

[6] T.P. Wallace, P.A. Wintz, An efficient three-dimensional aircraft recognition algorithm using normalized Fourier descriptor, *Comput. Graph. Image Process.* 13 (1980) 99-126.

[7] I. Kunttu, L. Lepistö, J. Rauhamaa, A. Visa, Multiscale Fourier descriptors for defect image retrieval, *Pattern Recognition Letters* 27 (2006) 123-132.

[8] D. Zhang, G. Lu, Study and evaluation of different Fourier methods for image retrieval, *Image and Vision Comput.* 23 (2005) 33-49.

[9] A. El-ghazal, O. Basir, S. Belkasim, Farthest point distance: A new shape signature for Fourier descriptors, *Signal Processing: Image Communication* 24 (2009) 572-586.

[10] D. Zhang, G. Lu, Review of shape representation and description techniques, *Pattern Recognition* 37 (2004) 1-19.

[11] The MPEG Home Page, <http://www.chiariglione.org/mpeg>, 2008.

[12] N. Arica, F.T.Y. Vural, BAS: a perceptual shape descriptor based on the beam angle statistics, *Pattern Recognition Letters* 24 (2003) 1627-1639.

[13] P.F. Felzenszwalb, J.D. Schwartz, Hierarchical matching of deformable shapes, *IEEE International Conference on Computer Vision and Pattern Recognition* Vol. 1, pp. 1-8, 2007.

[14] X. Bai, B. Wang, C. Yao, W. Liu, Z. Tu, Co-Transduction for shape retrieval, *IEEE Trans. Image Processing* 21 (2012) 2747-2757.

[15] L. J. Latecki, R. Lakämper, U. Eckhardt, Shape descriptor for non-rigid shapes with a single closed contour, in: *IEEE International Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp. 424-429, 2000.

[16] L. J. Latecki, R. Lakämper, Shape similarity measure based on correspondence of visual parts, *IEEE Trans. Pattern Anal. Machine Intell.* 22 (2000) 1185-1190.

[17] G.C-H. Chuang, E.G.M. Petrakis, Wavelet descriptor of planar curves: Theory and applications, *IEEE Trans. on Image Processing* 5 (1996) 56-70.

[18] S. Belongie, J. Malik, J. Puzicha, Shape matching and object recognition using shape contexts, *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (2002) 509-522.

[19] S. Manay, D. Cremers, B. Hong, A.J. Yezzi, Jr., S. Soatto, Integral invariants for shape matching, *IEEE Trans. Pattern Anal. Mach. Intell.* 28 (2006) 1602-1618.

[20] E. Miliou, E.G.M. Petrakis, Shape retrieval based on dynamic programming, *IEEE Trans. on Image Processing* 9 (2000) 141-147.

[21] C.H. Wei, Y. Li, W.Y. Chau, C.T. Li, Trademark image retrieval using synthetic features for describing global shape and interior structure, *Pattern Recognition* 42 (2009) 386-394.

[22] P. Korn, N. Sidiropoulos, C. Faloutsos, E. Siegel, Z. Protopapas, Fast and effective retrieval of medical tumor shapes, *IEEE Trans. on Knowledge and Data Engineering* 10 (1998) 889-904.

[23] J.X. Du, X.F. Wang, G.J. Zhang, Leaf shape based plant species recognition, *Applied Mathematics and Computation* 205 (2008) 916-926.

[24] F. Mokhtarian, S. Abbasi, J. Kittler, Efficient and robust retrieval by shape content through curvature scale space, in *Proc. Int. Workshop Image Database and Multimedia Search*, Amsterdam, The Netherlands, pp. 35-42, 1996.

[25] D. Zhang, G. Lu, A comparative study of curvature scale space and Fourier descriptor for shape based image retrieval, *Journal of Visual Communication & Image Representation* 14 (2003) 41-60.

[26] A. El-ghazal, O. Basir, S. Belkasim, Invariant curvature-based Fourier shape descriptors, *Journal of Visual Communication & Image Representation* 23 (2012) 622-633.

[27] A. Aung, B.P. Ng, S. Rahardja, Sequence-ordered complex Hadamard transform: properties, computational complexity and applications, *IEEE Trans. Signal Processing* 56 (2008) 3562-3571.

[28] Shih-Chin Fang, Hsiao-Lung Chan, Human identification by quantifying similarity and dissimilarity in electrocardiogram phase space, *Pattern Recognition*, 42(2009) 1824-1832.

[29] L. Pauleve, H. Jegou, L. Amsaleg, Locality sensitive hashing: A comparison of hash function type and querying mechanisms. *Pattern Recognition Letters*, 31 (2010) 1348-1358..

- [30] B. Wang, J.S. Wu, H.Z. Shu, L.M. Luo, Shape description using sequency-ordered complex Hadamard transform, *Optics communication* 284 (2011) 2726-2729.
- [31] S. Loncaric, A survey of shape analysis techniques, *Pattern Recognition* 31 (1998) 983-1001.
- [32] M. Bober, J.D. Kim, H.K. Kim, Y.S. Kim, W-Y. Kim, K. Muller. Summary of the results in shape descriptor core experiment. MPEG-7, ISO/IEC/JTC1/SC29/WG11/MPEG99/M4869, Vancouver, July 1999.
- [33] C. Grigorescu, N. Petkov, Distance sets for shape filters and shape recognition, *IEEE Trans. Image Processing* 12 (2003) 729-739.
- [34] J. Xie, P.A. Heng, M. Shah, Shape matching and modelling using skeletal context, *Pattern Recognition* 41 (2008) 1756-1767.
- [35] T.B. Sebastian, P.N. Klein, B.B. Kimia, On aligning Curves, *IEEE Trans. Pattern Anal. Machine Intell.* 25 (2003) 116-124.
- [36] E. Attalla, P. Siy, Robust shape similarity retrieval based on contour segmentation polygonal multiresolution and elastic matching, *Pattern Recognition* 38 (2005) 2229-2241.
- [37] C. Xu, J. Liu, X. Tang, 2D shape matching by contour flexibility, *IEEE Trans. Pattern Anal. Machine Intell.* 31 (2009) 180-186.
- [38] G. McNeill, S. Vijayakumar, Hierarchical procrustes matching for shape retrieval, *IEEE International Conference on Computer Vision and Pattern Recognition* Vol. 1, pp. 885-894, 2006.
- [39] F. Mokhtarian, A.K. Mackworth, A theory of multiscale, curvature based shape representation for planar curves, *IEEE Trans. Pattern Anal. Mach. Intell.* 14 (1992) 789-805.
- [40] C.T. Zahn, R.Z. Roskies, Fourier descriptors for plane closed curves, *IEEE Trans. Comput.* 21 (1972) 269-281.
- [41] C.C. Chang, S.M. Hwang, D.J. Buehrer, A shape recognition scheme based on relative distances of features points from the centroid, *Pattern Recognition* 24 (1991) 1053-1063.
- [42] N. Kumar, P. N. Belhumeur, A. Biswas, D.W. Jacobs, W.J. Kress, I. Lopez, J.V.B. Soares, Leafsnap: A Computer Vision System for Automatic Plant Species Identification, *ECCV 2012*, Part II, pp. 502-516.
- [43] S. Mouine, I. Yahiaoui, A. Verroust-Blondet, A shape-based approach for leaf classification using multiscale triangular representation, in: *Proceedings of the 3rd ACM International Conference on Multimedia Retrieval, 2013*, pp. 127-134.
- [44] H.S. Yang, S.U. Lee, K.M. Lee, Recognition of 2-D object contours using starting-point independent wavelet coefficient matching, *J. Visual Commun. Image Representation* 9 (1998) 171-181.
- [45] O. Söderkvist, Computer vision classification of leaves from Swedish trees, Master's thesis, Linköping University, 2001.
- [46] S. Biswas, G. Aggarwal, R. Chellappa, An efficient and robust algorithm for shape indexing and retrieval, *IEEE Trans. Multimedia.* 12 (2010) 372-385.
- [47] R. Hu, W. Jia, H. Ling, D. Huang, Multiscale distance matrix for fast plant leaf recognition, *IEEE Trans. Image Processing.* 21 (2012) 4667-4672.
- [48] R. Hu, W. Jia, H. Ling, Y. Zhao, J. Gui, Angular pattern and binary angular pattern for shape retrieval, *IEEE Trans. Image Processing.* DOI: 10.1109/TIP.2013.2286330.
- [49] F. Foteini, E. George, Multivariate angle scale descriptor for shape retrieval, in *Proc. Signal Processing and Applied Mathematics for Electronics and Communications (SPAMEC)*, 2011, pp. 105-108.
- [50] J. Wang, X. Bai, X. You, W. Liu, L.J. Latecki, Shape matching and classification using height functions, *Pattern Recognition Letters*, 33 (2012) 134-143.
- [51] R. Gopalan, P. K.Turaga, R. Chellappa, Articulation-invariant representation of non-planar shapes, in *Proc. ECCV, 2010*, PP. 286-299.
- [52] F. Mokhtarian and M. Bober, Curvature scale space representation: theory, applications, and MPEG-7 standardization, Kluwer Academic Publishers, 2003.
- [53] I.E. Rubé, N. Alajlan, M.S. Kamel, MATR: A robust 2D shape representation, *International Journal of Image and Graphics*, 6 (2006) 421-443.



Bin Wang received the PhD degree in computer science in 2007 from Fudan University, Shanghai, China. Since 2007, he has been on the faculty of Nanjing University of Finance and Economics, Nanjing, China, where he is currently an associate professor. From 2007 to 2011, he was a Postdoctoral Research Fellow (part time) with Southeast University, Nanjing, China. From December 2011 to December 2012, he was with the School of Engineering, Griffith University, Australia, as a Visiting Scholar. His main research interests include computer vision, image processing and pattern recognition.



Yongsheng Gao received the B.Sc. and M.Sc. degrees in electronic engineering from Zhejiang University, China, in 1985 and 1988, respectively, and the Ph.D. degree in computer engineering from Nanyang Technological University, Singapore. Currently, he is a Professor with the School of Engineering, Griffith University, Australia. His research interests include face recognition, biometrics, biosecurity, image retrieval, computer vision, pattern recognition, environmental informatics and medical imaging.