Chord Bunch Walks for Recognizing Naturally Self-overlapped and Compound Leaves

Bin Wang, *Member, IEEE*, Yongsheng Gao, *Senior Member, IEEE*, Changming Sun, Michael Blumenstein, John La Salle

Abstract — Effectively describing and recognizing leaf shapes under arbitrary variations, particularly from a large database, remains an unsolved problem. In this research, we attempted a new strategy of describing leaf shapes by walking and measuring along a bunch of chords that pass through the shape. A novel chord bunch walks (CBW) descriptor is developed through the chord walking behaviour that effectively integrates the shape image function over the walked chord to reflect both the contour features and the inner properties of the shape. For each contour point, the chord bunch groups multiple pairs of chords to build a hierarchical framework for a coarse-to-fine description that can effectively characterize not only the subtle differences among leaf margin patterns but also the interior part of the shape contour formed inside a self-overlapped or compound leaf. Instead of using optimal correspondence based matching, a Log-Min distance that encourages one-to-one correspondences is proposed for efficient and effective CBW matching. The proposed CBW shape analysis method is invariant to rotation, scaling, translation, and mirror transforms. Five experiments, including image retrieval of compound leaves, image retrieval of naturally self-overlapped leaves, and retrieval of mixed leaves on three large scale datasets, are conducted. The proposed method achieved large accuracy increases with low computational costs over the state-of-the-art benchmarks, which indicates the research potential along this direction.

Index Terms — Shape description, shape matching, leaf image retrieval, compound leaf, self-overlapped leaf, leaf margin, chord bunch walks, Log-Min distance.

— — — — — — — — — ◆ — — — — — — — — —

1. INTRODUCTION

The human vision system has the strength of recognizing objects by their shapes. Compared with various visual features, shape is regarded as one of the best features for object recognition [18], [19]. Automated shape analysis by computers is a central and challenging problem in image processing and computer vision. Its main task is to capture the discriminative geometrical characteristics that is independent of the transformational effects such as translation, scaling, rotation, mirror, and articulation for the ultimate goal of recognizing objects.

Leaf shape based plant identification is a challenging example application of shape analysis. Identifying plant species is usually impossible for the general public and often a difficult task for farmers, wood exploiters, and even botanists themselves. Describing and matching leaf shape images is considered as one of the solutions for plant identification. Leaf shapes have many morphological variants between different species, such as long and very narrow, short apex angled, heart-shaped, fan-like, sickle-shaped, diamond-shaped, and wave-like shapes. Moreover, leave shapes of different species usually have their distinctive leaf margin patterns. Fig. 1 shows some of them. For example, serrate is a common margin pattern that has small shape teeth pointing upward, while runcinate margin is a

pattern that the lobes are deeply cut and pointing downward (see Fig. 1).

The leaves of different plant species can appear in the form of either simple leaves or compound leaves. A simple leaf only has a single blade, which is undivided (see the top row in Fig. 10). However, a compound leaf [1] consists of a number of leaflets (see Fig. 2), and the leaflets themselves may be compound (see the leaf on the right in Fig. 2). Recognizing shapes of compound leaves presents a significant challenge to the research community.

Another open problem in describing and recognizing leaf shapes is to handle naturally self-overlapped leaves. (see Fig. 3). When a leaf is self-overlapped (the middle one in Fig. 3), a contour based shape matching approach will misidentify it to another species (the right one in Fig. 3) instead of matching it to the sample of its own species (the left one in Fig. 3). Current shape analysis techniques are not yet designed to tackle above problems.
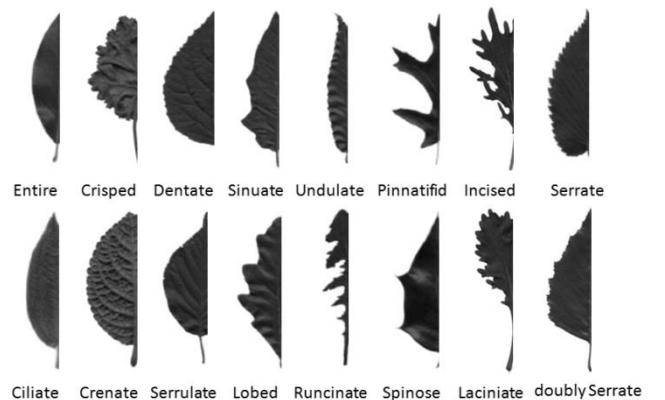
Fig. 1. Examples of leaf margin patterns.

Fig. 2. Examples of compound leaves.

Fig. 3. Examples of self-overlapped leaves (top row) and their contours (bottom row).

In this paper, we present a novel chord bunch walks method for leaf shape description and identification. We propose a new strategy of describing shapes by walking and observing along a group of chords trespassing the shape regions. The observations of the shape when walking along the chords are used to describe not only the outer contour features but also the interior properties of the shape. The small sized chord pair walks tend to capture the leaf margin patterns, while the large sized ones tend to characterize more global and interior information. An early conference version of this research is reported in [26]. In this paper, we present our complete work with new and improved formulations, extended experimental investigation and analysis. To our knowledge, this is believed the first reported attempt that particularly considers compound leaf shapes, self-overlapped leaf shapes, and leaf margin patterns in the design of a shape analysis approach.

The rest of this paper is organized as follows: Related works are briefly reviewed in Section 2. In Section 3, we present the detailed design of the proposed CBW method. Extensive experimental investigations are reported in Section 4. Finally, conclusions are drawn in Section 5.

2. RELATED WORK

Shape analysis and matching methods relevant to this study can be categorized into global approaches and local approaches. The former extracts the shape region information (usually represented as a 2-D function) or the contour information (usually represented as a 1-D function) to form a feature vector, or a sequence of features for describing a shape. Shape dissimilarity can thus be measured by

a metric, for example, using the L_1 or L_2 distance. A classical descriptor for region-based shape analysis is the Zernike moments (ZM) [39] which is generated by projecting the shape image function on a set of orthogonal basis functions defined in the interior of a unit circle. Yap et al. [12] proposed a novel set of 2D transforms, the polar harmonic transforms (PHTs), for rotation invariant feature extraction. The computation of the PHT kernels is significantly faster than that of the Zernike moments. Recently, the Radon transform and its generalized version, Trace transform, have been extensively researched for shape analysis [7],[11],[29], [30],[33]. These methods finely capture the inner structural information via line integrals over the shape region. Charters et al. [45] presented a novel approach for modelling leaf venation structures to improve the accuracy of plant species recognition. More recently, Hong and Soatto [34] proposed a shape descriptor based on integral kernels. They use a series of isotropic kernels to characterize shapes, which results in an invariant and robust shape descriptor.

Another subgroup of global methods are contour-based approaches, which describe a shape only by its contour containing sequential information. A contour based descriptor that combines the features of the centroid-contour distance, eccentricity, and angle code histogram was proposed by Wang et al. [44] and achieved good performance for leaf image retrieval. The spectral methods, such as Fourier descriptor [8],[40] and wavelet descriptor [41], are typical examples. Recently, Hu et al. [3] proposed a novel descriptor, the multiscale distance matrix (MDM), which uses the distances between contour points at multiple scales to build a matrix for reflecting the spatial relationship of the contour points. The distance between the contour points can be Euclidean distance or other metrics such as inner distance [9]. To provide a more efficient solution for shape retrieval, hierarchical string cuts [10] is proposed to extract a group of geometrical features, which reflect the spatial configuration of the curve segments relative to their strings. Backes et al. [46] proposed to model the shape into a small-world complex network and took the measures of the network for shape description. Leaf plant classification is used in their experiment to validate the effectiveness of their complex network based shape descriptors. Inspired by the work in [46], Ribas et al. [47] proposed a distance transform network, which incorporated the network with a Euclidean distance transform, for shape analysis and applied their method for leaf recognition.

Leaf margin (see Fig. 1) is very useful for identifying plant species. Recently, some efforts have been made to extract margin patterns for leaf image analysis. Jin et al. [48] made a morphological analysis on the leaf margin in which the total number of teeth, the ratio between the number of teeth and the length of the margin are measured for plant identification. Cope and Remagnino [49] utilized the leaf's insertion point and apex to generate a margin signature for describing the leaf pattern. Cerutti et al. [50] described the leaf margin by a string representation of symbols, and applied the method for mobile tree identification.

For local approaches, a descriptor is extracted for each

visual primitive such as point, edge of polygon, and curve fragment of the shape. The local descriptors associated with each primitive are collected to describe the shape. Since each primitive of the shape is associated with a local descriptor, these descriptors provide cues for establishing correspondence between the primitives of the shapes to be compared. The matching cost is taken as the shape difference measure. It is a natural choice to take the contour points as primitives and extract their descriptors using the geometric information around them. The shape contexts method [15] builds a histogram for each contour point, which encodes the information of relative spatial distribution (distance and orientation) between the given contour point and all the other contour points. The inner-distance shape contexts approach [9] replaces the Euclidean distance used in the shape context with the inner-distance, which is defined as the length of the shortest path between two landmark points within the shape silhouette, to effectively capture partial structures of the shape. The distance sets method [16] uses a rich local descriptor for points determined by the spatial configuration of their surrounding points. The spatial arrangement is characterized by the set of the distances between the given point and the surrounding points. The height function method [36] uses the height values (perspective distance) from all the contour sample points to the tangent line across a given contour point to generate a local descriptor that captures the geometric relationships of the contour points with respect to the given point.

Some local descriptors are developed by characterizing the curvature property or bend potential on the shape contour. Alajlan et al. [17] use the areas of triangles formed by the contour points to measure the convexity/concavity of each point at different scales. The integral invariants method [13] introduces a class of functionals obtained by performing integral operations against the shape due to their robustness to high-frequency noise and small deformations. The area integral invariant [13], a particularly useful functional that can reconstruct the curvature of a contour, has been effectively applied to leaf identification [42]. Contour flexibility [23] makes effort to represent the deformable potential at each point along the contour and shows that both local and global features can be extracted by this descriptor. There are also other similar methods in [24],[25],[35],[43].

3. THE PROPOSED METHOD

In this section, we introduce the concept of CBW and its detailed design for a coarse-to-fine shape analysis, which not only can effectively harvest the internal shape information in self-overlapped leaves and compound leaves, but also can differentiate the subtle differences among leaf margin patterns. The invariances derived from the CBW are presented and discussed. Finally, a novel Log-Min distance is defined for efficiently and effectively matching CBWs.

3.1 Chord Pair Walks

Given a binary shape image $f(x, y)$, the shape regions (denoted as D) are formed by a subset of pixels in the image plane $\mathbb{R}^2$. Let Ω be the outer contour of the shape enclosing all pixels in D, which can be represented in an arc-length parameterization form [23]: $z(t) = (\bar{x}(t), \bar{y}(t)), t \in [0,1)$, where $(\bar{x}, \bar{y}) \in \Omega$. Since Ω is a closed contour, we have $z(t + 1) = z(t)$ and $z(t - 1) = z(t)$.

For a contour point $p = z(t)$, we walk along a pair of chords $\overrightarrow{pp'}$ and $\overrightarrow{pp''}$ whose end points are $p' = z(t + s)$ and $p'' = z(t - s)$ respectively, where $s \in (0, \frac{1}{2}]$ is the normalized length of a section of the contour generated by moving point p counterclockwise along the contour to point p'. The paths of walks sometimes fall inside the shape region D, and sometimes fall outside the shape region D (see Figs. 4 and 5). The lengths of walking inside D ($\hat{l}_t^s$ and $\hat{l}_t^{-s}$) and the lengths of walking outside D ($\check{l}_t^s$ and $\check{l}_t^{-s}$) can be mathematically expressed as

$$
\begin{cases}
\hat{l}_t^s = \int_0^{l^s} f(\bar{x}(t) + \tau \cos \theta_t^s, \bar{y}(t) + \tau \sin \theta_t^s) d\tau \\
\hat{l}_t^{-s} = \int_0^{l^{-s}} f(\bar{x}(t) + \tau \cos \theta_t^{-s}, \bar{y}(t) + \tau \sin \theta_t^{-s}) d\tau
\end{cases}
\tag{1}
$$

and

$$
\begin{cases}
\check{l}_t^s = \int_0^{l^s} \left(1 - f(\bar{x}(t) + \tau \cos \theta_t^s, \bar{y}(t) + \tau \sin \theta_t^s)\right) d\tau \\
\check{l}_t^{-s} = \int_0^{l^{-s}} \left(1 - f(\bar{x}(t) + \tau \cos \theta_t^{-s}, \bar{y}(t) + \tau \sin \theta_t^{-s})\right) d\tau
\end{cases}
\tag{2}
$$

where l^s and l^{-s}, θ_t^s and θ_t^{-s} are the lengths and orientations of the chord pair $\overrightarrow{pp'}$ and $\overrightarrow{pp''}$, respectively.

The $\hat{l}_t^s$ and $\hat{l}_t^{-s}$ defined above are the integrals of the shape image function $f(x, y)$ over the chord pairs $\overrightarrow{pp'}$ and $\overrightarrow{pp''}$ while $\check{l}_t^s$ and $\check{l}_t^{-s}$ are the integrals of the complement shape image over the chords $\overrightarrow{pp'}$ and $\overrightarrow{pp''}$. By walking and observing the shape along the chord pair, a quintuplet of measurement is defined to depict the trespassed regions in the shape as

$$
W_t^s = \left(\hat{l}_t^{-s}, \check{l}_t^{-s}, \sin(\theta_t^{-s} - \theta_t^s), \hat{l}_t^s, \check{l}_t^s \right).
\tag{3}
$$

The chord pair walks measurement W_t^s serves as a primitive descriptor to capture the local geometrical and structural features of the neighborhood around the contour point $p = z(t)$. The first two elements reflect the property of its left chord neighborhood with size s, while the last two elements express the property of its right chord neighborhood with the same size s. The middle element $\sin(\theta_t^{-s} - \theta_t^s)$ depicts its curvature property, in which the value of $\sin(\theta_t^{-s} - \theta_t^s)$ being > 0, < 0 and $= 0$ indicates the convex, concave and flat properties respectively. Thus, this chord pair walk measurement W_t^s can capture not only the geometrical information of the outer shape contour but also the structural information of the inner regions within the neighborhood of size s around point $z(t)$.
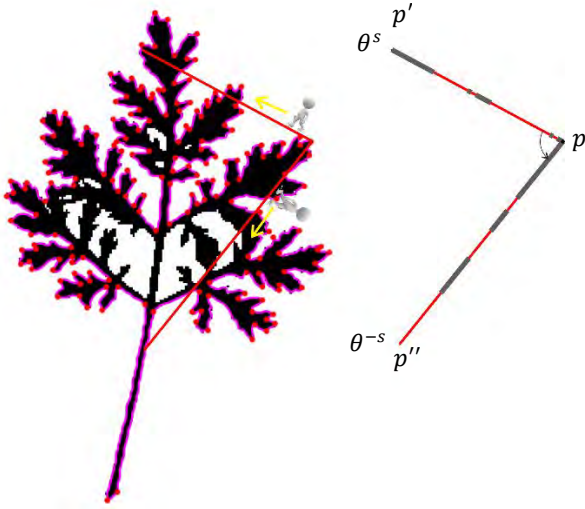
Fig. 4. An example of chord pair walks trespassing the holes formed by self-overlapped leaf regions. The outer contour is shown in magenta color and its salient points (detected by the DCE algorithm [21]) are shown as red dots.
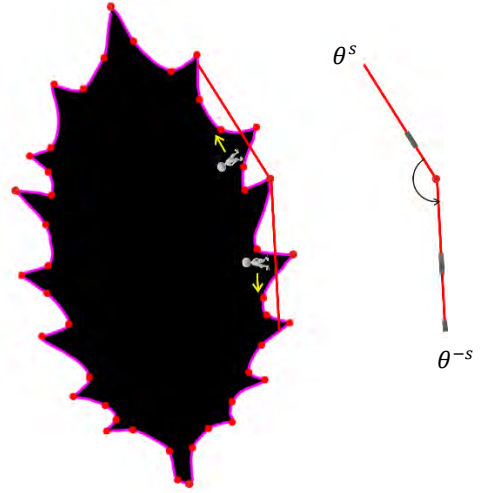
Fig. 5. An example of chord pair walks passing through the leaf margin. The outer contour is shown in magenta color and its salient points (detected by the DCE algorithm [21]) are shown as red dots.
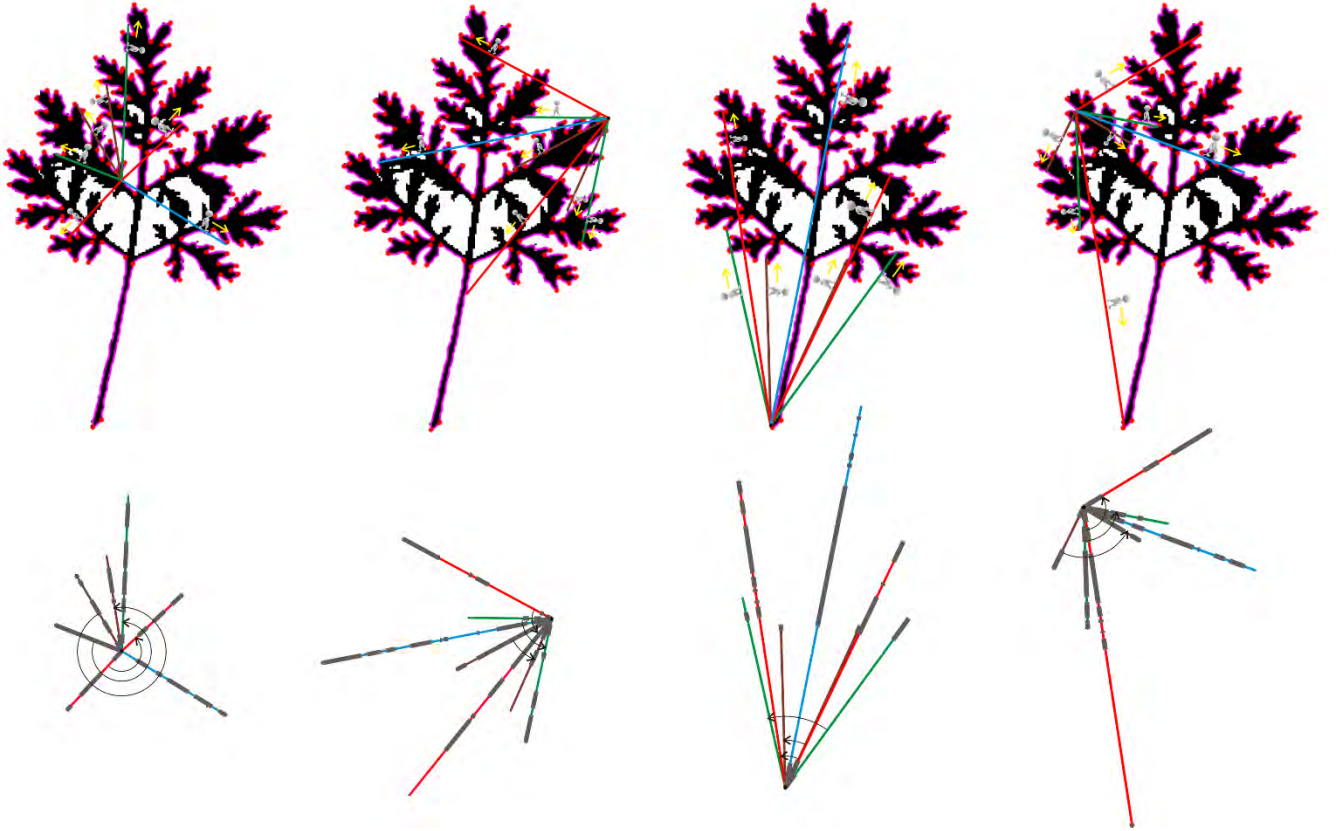

Fig. 6. An example illustrating the concept of chord bunch walks. From left to right: chord bunch walks ($K = 4$) for four points $p = z(t)$ at different t along the contour in a counterclockwise direction. Chord pairs for scales $s = 2^{-1}, \cdots, 2^{-K}$ are shown in blue, red, green, and brown colors, respectively. The figures in the second row visualize the five elements of each quintuplet W_t^s in the CBW. The outer contour is shown in magenta color and its salient points (detected by the DCE algorithm [21]) are shown as red dots.

3.2 Chord Bunch Walks

The parameter s in W_t^s is a scale which controls the size of the neighborhood of point $p = z(t)$. By varying s to take values $2^{-1}, \cdots, 2^{-K}$, we can obtain K chord pair walks originating from the common point p and group them to form a CBW defined as

$$\widetilde{W}_t = \bigcup_{k=1}^{K} \{W_t^s, s = 2^{-k}\}, \qquad (4)$$

where K is the total number of neighborhood scales for the CBWs.

For the first scale $k = 1$, i.e., $s = 2^{-1}$, the whole contour is taken as the neighborhood of point p, and its corresponding CBW capture the coarsest global features of the shape. With the increase of scale index k, the neighborhood becomes smaller and the corresponding CBW capture finer local details. Hence, the CBW descriptor has a desirable ability of capturing the shape characteristics spanning the full coarse-to-fine range of information. An example illustrating the concept of CBW is given in Fig. 6.

To further enhance the discriminability of the CBW descriptor, the spatial relationship between chord pair walks is also considered in the design. For two adjacent chord pair walks, $\{W_t^s, s = 2^{-k}\}$ and $\{W_t^s, s = 2^{-(k+1)}\}$, in the CBWs, their spatial relationship can be represented by the angles $\vartheta_{t,k}^{(L)}$ and $\vartheta_{t,k}^{(R)}$ between the left and right chords, respectively:

$$\vartheta_{t,k}^{(L)} = \theta_t^{-2^{-(k+1)}} - \theta_t^{-2^{-k}}, \qquad \vartheta_{t,k}^{(R)} = \theta_t^{2^{-k}} - \theta_t^{2^{-(k+1)}}. \quad (5)$$

Since there are a total of K chord pair walks in CBW, we have $K - 1$ angle pairs. Fig. 7 gives an example illustrating the computation of these angle pairs. The spatial relationship of the CBWs is characterized by a combination of the angle pairs for all chord pairs as

$$\delta_t = \bigcup_{k=1}^{K-1} \{\sin(\vartheta_{t,k}^{(R)} - \vartheta_{t,k}^{(L)})\}. \qquad (6)$$

Combining the geometrical and structural information (Eq. (4)) and the spatial information (Eq. (6)), a CBW descriptor for a contour point $p = z(t)$ is defined as a vector of $5K - 3 + K - 1 = 6K - 4$ dimentions:

$$\widetilde{W}_t = \left(\bigcup_{k=1}^{K} \{W_t^s, s = 2^{-k}\} \right) \bigcup \left(\bigcup_{k=1}^{K-1} \{\sin(\vartheta_{t,k}^{(R)} - \vartheta_{t,k}^{(L)})\} \right). \qquad (7)$$

A good descriptor is expected to be translation, scale, rotation, and mirror invariant, which allows recognition of shapes disregarding their position, orientation, and distance. Eq. (7) has the intrinsic invariance to translation and rotation. To ensure that our shape descriptor is also invariant to scale variations, $\hat{l}_t^s$ is normalized as

$$\hat{l}_t^s = \frac{\hat{l}_t^s - \min_{0 \le t < 1}\{\hat{l}_t^s\}}{\max_{0 \le t < 1}\{\hat{l}_t^s\} - \min_{0 \le t < 1}\{\hat{l}_t^s\}}. \qquad (8)$$

Similarly, we perform the same normalization for $\check{l}_t^s, \hat{l}_t^{-s}, \check{l}_t^{-s}$ before computing the chord pair walks quintuplets.
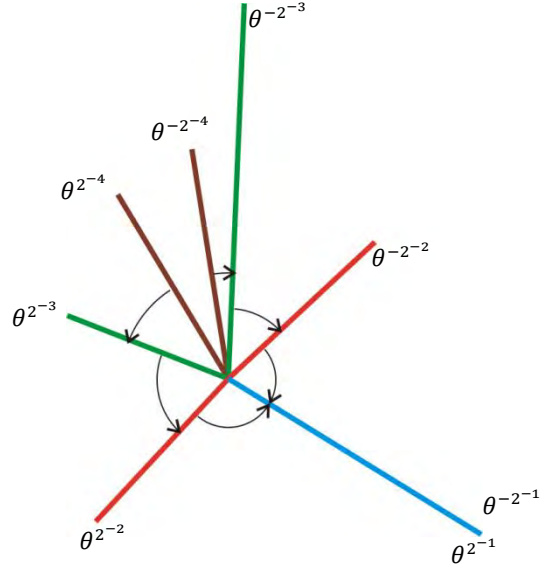

Fig. 7. An example illustrating the computation of angle pairs for the leftmost CBW in Fig. 6.

To make the CBW descriptor invariant to the mirror transform, Eq. (3) and Eq. (7) are reorganized as follows (see Appendix A).

$$W_t^s = (\max(\hat{l}_t^{-s}, \hat{l}_t^s), \max(\check{l}_t^{-s}, \check{l}_t^s), \sin(\theta_t^{-s} - \theta_t^s),$$
$$\min(\hat{l}_t^{-s}, \hat{l}_t^s), \min(\check{l}_t^{-s}, \check{l}_t^s)) \qquad (9)$$

$$\widetilde{W}_t = \left(\bigcup_{k=1}^{K} \{W_t^s, s = 2^{-k}\} \right)$$
$$\bigcup \left(\bigcup_{k=1}^{K-1} \{|\sin(\vartheta_{t,k}^{(R)} - \vartheta_{t,k}^{(L)})|\} \right). \qquad (10)$$

Hence, the proposed CBW $\{\widetilde{W}_t, 0 \le t < 1\}$ is completely translation, rotation, scale, and mirror invariant.

3.3 Log-Min Distance Matching between CBWs of Salient Points

Many shape matching methods [9],[15],[36] employ dynamic programing to find an optimal correspondence [17] between the contour points. A limitation of this scheme is its high computational cost. Secondly, it does not always work well for matching self-overlapped shapes or compound leaf shapes whose local details are hard to be matched in pairs. Thirdly, it is not mirror invariant. On the other hand, the Hausdorff distance [14],[20] has the advantage of being computationally efficient by employing a many-to-one matching scheme. However, a CBW in one image is desired to be matched to a single CBW in another image. In this study, we proposed a *Log-Min* distance, that not only has the advantage of enabling fast calculation but also encourages one-to-one matching, for efficient CBW set-to-set dissimilarity measurement.

We first sample the leaf contour into $T = 2^{K+1}$ uniformly spaced points, where T is sufficiently large for showing details of the leaf contour. K is the number of the scales in the

CBW calculation. Then the salient points on the contour that contribute more discriminative information for shape recognition than the other ones are extracted using the approach in [21]. A leaf image is thus represented by a set of CBWs on salient points.

Given two CBW sets $P = \{\widetilde{W}_1^{(P)}, \widetilde{W}_2^{(P)}, \cdots, \widetilde{W}_M^{(P)}\}$ and $Q = \{\widetilde{W}_1^{(Q)}, \widetilde{W}_2^{(Q)}, \cdots, \widetilde{W}_N^{(Q)}\}$ representing the probe and gallery shapes, respectively. M and N are the numbers of salient points on the contours of the probe and gallery shapes respectively. Note that we do not require $M = N$ which is different from those dynamic programing based methods [9],[15],[17],[36].

For each CBW $\widetilde{W}_i^{(P)}$ in P, we find its nearest correspondence, $\pi(\widetilde{W}_i^{(P)})$, among all CBWs in Q using $\pi(\widetilde{W}_i^{(P)}) = \arg \min_{\widetilde{W}_j^{(Q)} \in Q} (d(\widetilde{W}_i^{(P)}, \widetilde{W}_j^{(Q)}))$, where $d(\cdot)$ denotes an L_1 distance. Let $u(\widetilde{W}_j^{(Q)}) = \{\widetilde{W}_i^{(P)} | \pi(\widetilde{W}_i^{(P)}) = \widetilde{W}_j^{(Q)}, i = 1, \dots, M\}$ be the subset of P whose elements are matched to the same nearest correspondence $\widetilde{W}_j^{(Q)}$ in Q. Assume that there are C subsets $u_1, u_2, \cdots, u_C$ in P, we have

$$P = \bigcup_{p=1}^{C} u_p, \qquad u_m \cap u_n = \emptyset \text{ for } m \neq n,$$

where $\emptyset$ is a null set, and $m, n \in \{1, 2, \dots, C\}$. Because all the elements in a subset u_p are matched to the same CBW in Q, C is not greater than the number of CBWs in Q (i.e., $C \leq N$). The Log-Min distance between the two CBW sets P and Q is defined as

$$H(P, Q) = \max(h(P, Q), h(Q, P)), \qquad (11)$$

$$h(P, Q) = \frac{1}{|P|} \sum_{p=1}^{C} \log_2(|u_p| + 1) D(u_p, Q), \qquad (12)$$

where $|P|$ is the cardinality of the set P, $|u_p|$ is the cardinality of subset u_p. $D(u_p, Q)$ is the sum of L_1 distances between the CBWs in the subset u_p and the single matched CBW in Q which is defined as

$$D(u_p, Q) = \sum_{\widetilde{W}_i^{(P)} \in u_p} \left(\min_{\widetilde{W}_j^{(Q)} \in Q} d(\widetilde{W}_i^{(P)}, \widetilde{W}_j^{(Q)}) \right). \qquad (13)$$

When $|u_p| = 1$, the subset $u_p \subseteq P$ only contains one element which indicates a one-to-one correspondence. When $|u_p| = 2$, the subset $u_p \subseteq P$ contains two elements that indicate a two-to-one correspondence. Similarly, $|u_p| = k$ indicates a k-to-one correspondence. Because one-to-one correspondences are desirable, $\log_2(|u_p| + 1)$ in Eq. (12) plays a role as a penalty factor to discourage matching multiple CBWs in P to a single CBW in Q. Thus, the proposed Log-Min distance can enforce the desirable one-to-one correspondence that the Hausdorff distance cannot, while preserving the advantage of fast computation.

4. EXPERIMENTAL RESULTS

To evaluate the effectiveness of the proposed CBW approach, an extensive experimental investigation has been conducted on compound leaves, self-occluded leaves, and mixed leaves on three publicly available datasets: ICL, MEW2012, and CVIP100 leaf datasets. The performance (both accuracy and speed) of the proposed CBW is compared against six state-of-the-art methods. Among them, the Shape Context [15] and Inner Distance Shape Context [9] are the most well-known shape analysis approaches, which are widely used as benchmarks for performance comparison. The Height Function [36], MDM [3], Complex Network [46], and HSC [10] are the recent contour shape analysis methods with state-of-the-art performances. In all the experiments, the number of scales K for the proposed method is set to 7 and the threshold τ of the algorithm DCE [21] used for salient point extraction is set to 2.5×10^{-2}.

4.1 ICL Leaf Dataset for Compound Leaves, Self-overlapped Leaves and Mixture of Leaves

Automatic plant leaf identification is an important application of computer vision and has received much attention [2],[3],[5],[6],[9],[10],[38],[37] in recent years. The ICL leaf dataset [3] is a publicly available leaf database, which contains 16,846 samples from 220 species with 26 to 1,078 samples per species (see Fig. 8). This dataset is the largest dataset currently available to the public, which presents significant challenges to the research community due to compound leaves, naturally self-overlapped leaves, and their very large intra-class variations (see examples in Fig. 9).

In the ICL dataset, there are 654 compound leave images from 11 species (see examples in the first row of Table I). The numbers of images from each species are 63, 97, 49, 69, 90, 41, 48, 54, 26, 62, and 55 respectively. Also, there are 8 species whose leaves are naturally self-overlapped (see examples in the first row of Table II). The numbers of images from each species are 96, 53, 37, 46, 41, 36, 58, 52, respectively, which form a self-overlapped leave subset with 419 images.

Some experiments [3],[24],[38] have been conducted on the ICL dataset. However, they only chose part of the leaf samples (less than 6,000) for testing and some of them preprocess the images by cutting off the footstalks from all the leaf samples before applying their methods. To have a rigorous and honest evaluation that reflects the real world conditions, we use the original leave images as is without any preprocessing and use all the 16,846 images without selection in our experiment. Three datasets are created for our experiments: the compound leaf subset (654 images), the self-overlapped leaf subset (419 images), and the mixture of leaves of the complete ICL dataset (16,846 images).

Fig. 8. Two hundred and twenty sample images (one sample per species) from the ICL leaf dataset [3],[37].

Fig. 9. Example leaf shapes from the ICL leaf dataset that show large intra-class variations including natural deformations, naturally self-overlapped leaves, and compound leaves.

Table I. The MAP scores (%) obtained using images in the compound leaf subset (654 images) as queries to retrieve images from the complete ICL dataset (16,846 images).

Method												AVG.
IDSC+DP [9]	62.1	66.9	96.4	51.8	39.2	95.2	71.2	75.2	57.5	40.5	59.7	65.0
SC+DP [9][15]	50.2	87.2	100.0	65.1	33.6	96.1	75.5	82.0	75.7	43.7	69.2	70.8
MDM-CD-RM [3]	48.6	46.0	88.8	30.3	29.9	98.5	33.2	65.2	40.8	84.2	39.4	55.0
MDM-ID-RA [3]	46.4	37.5	66.2	23.2	29.9	72.9	26.7	52.9	43.3	49.1	28.1	43.3
Height Function [36]	67.0	81.6	100.0	57.1	35.4	92.9	76.2	79.4	65.2	42.0	61.6	68.9
HSC [10]	55.2	75.4	100.0	53.4	34.6	96.3	62.7	77.4	69.3	73.9	66.6	69.5
Complex Network [46]	38.5	43.7	55.4	28.6	25.5	63.5	46.9	66.7	45.8	35.3	33.1	43.9
Proposed CBW+Hausdorff	**74.6**	**73.4**	**100.0**	**72.7**	**40.6**	**99.4**	**70.6**	**81.0**	**87.4**	**92.2**	**69.0**	**78.3**
Proposed CBW+DP	**58.2**	**73.9**	**100.0**	**58.5**	**34.0**	**98.6**	**69.8**	**83.5**	**66.5**	**79.2**	**64.9**	**71.6**
Proposed CBW+LogMin	**76.5**	**78.4**	**100.0**	**79.0**	**41.8**	**98.5**	**68.5**	**82.2**	**94.3**	**90.6**	**67.0**	**79.7**

Table II. The MAP scores (%) obtained using images in the self-overlapped leaf subset (419 images) as queries to retrieve images from the complete ICL dataset (16,846 images).

Method									AVG.
IDSC+DP [9]	90.1	89.9	91.8	44.3	75.3	63.5	85.7	89.4	78.8
SC+DP [9][15]	81.9	97.9	95.5	63.3	85.8	87.9	90.4	97.9	87.6
MDM-CD-RM [3]	58.7	67.2	51.1	36.7	42.5	60.6	70.6	91.2	59.8
MDM-ID-RA [3]	57.3	65.2	36.5	30.7	50.7	56.3	59.7	83.6	55.0
Height Function [36]	70.2	93.6	88.5	60.7	75.0	86.5	89.7	97.0	82.6
HSC [10]	66.8	92.4	74.7	55.0	82.2	92.9	81.2	97.3	80.3
Complex Network [46]	52.9	53.8	26.8	19.3	36.7	41.0	37.4	72.0	42.5
Proposed CBW+Hausdorff	**82.7**	**98.4**	**99.2**	**62.5**	**83.3**	**97.1**	**90.4**	**99.7**	**89.2**
Proposed CBW+DP	**83.6**	**98.0**	**97.1**	**48.6**	**76.0**	**93.0**	**90.0**	**99.6**	**85.8**
Proposed CBW+LogMin	**90.1**	**98.5**	**99.3**	**59.0**	**85.6**	**96.7**	**90.8**	**99.8**	**90.0**

Mean average precision (MAP) [1] [4] is a standard measure for evaluating the performance of information retrieval systems, which is widely used in evaluating systems for image retrieval [5],[6],[31],[32], speech index [28], and video retrieval [27],[22]. Each sample in the dataset is taken as a query to retrieve the similar ones from all the samples in the dataset. The MAP scores obtained by the proposed method and the seven state-of-the-art approaches are summarized in Table I for compound leaf retrieval, in Table II for self-overlapped leaf retrieval, and in Table III for mixed leaf retrieval.

From Table I, it is observed that the average MAP score obtained by the proposed CBW+LogMin method for the compound leaf species are higher than those of the benchmark methods with large margins. On average, the CBW+LogMin achieves a MAP score of 79.7%, which is 8.9% higher than the second best method of Shape Context using dynamic programming (SC+DP). The performance increases (by 8.9% up to 36.4%) over the seven state-of-the-art benchmarks (whose MAP scores ranges from 43.3% to 70.8%) are very encouraging in handling the challenging compound leaves.

Table III. The MAP scores (%) obtained using images in the complete ICL leaf dataset as queries to retrieve images from the complete ICL dataset (16,846 images). "~" indicates that the program did not finish in 20 days.

Algorithm	MAP score (%)
IDSC+DP [9]	~
SC+DP [9][15]	~
MDM-CD-RM [3]	37.5
MDM-ID-RA [3]	34.2
Height Function [36]	~
HSC [10]	51.7
Complex Network [46]	28.5
Proposed CBW+Hausdorff	**55.2**
Proposed CBW+DP	~
Proposed CBW+LogMin	**56.7**

Table II reports the results of the second experiment on the self-overlapped leaves. It can be seen that the proposed method achieves an average MAP score of 90.0%, which is higher than those of benchmarks by 2.4% up to 47.5%. Again, the CBW+LogMin obtained the highest score on every species compared to the seven benchmarks.

The third experiment is to evaluate the performance of the proposed method on mixed leaves using the whole ICL database. To the best of our knowledge, this is believed to be the first report on performance comparison using the complete ICL dataset (i.e. all 16,846 leaves). The experimental results are summarized in Table III. We also recorded the average number of the salient points on leaves extracted by [21], which is 132 in this experiment. Note that three benchmark methods are not able to finish the computation, because the computational costs of IDSC [9], SC+DP [9][15] and Height Function [36] are too high (see average retrieval time reported in Table IV). In this challenging situation, the proposed CBW+LogMin method achieves a MAP score of 56.7% which is 5.0%, 19.2%, 22.5%, and 28.2% higher than the scores of HSC [10], MDM-CD-RM [3], MDM-CD-RA [3], and Complex Network [46] methods respectively.

Note that the performance improvement of the proposed CBW+LogMin over HSC [10], MDM-CD-RA [3], MDM-CD-RM [3], and Complex Network [46] methods on the compound leaves are 10.2%, 36.4%, 24.7%, 35.8% respectively. The performance improvement of the proposed CBW+LogMin over HSC [10], MDM-CD-RA [3], MDM-CD-RM [3], Complex Network [46] methods on the self-overlapped leaves are 9.7%, 35.0%, 30.2%, 47.5% respectively. However, the performance improvement of the proposed CBW+LogMin over HSC [10], MDM-CD-RA [3], and MDM-CD-RM [3], Complex Network [46] methods on mixed leaves of the whole ICL database are 5.0%, 22.5%, 19.2%, and 28.2% respectively. The larger accuracy increases of the proposed CBW+LogMin (e.g., over HSC) on the compound leaves (10.2% vs 5.0%) and self-overlapped leaves (9.7% vs 5.0%) demonstrate the effectiveness of our

CBW design as presented in Section 3.

We also conducted another group of experiments to evaluate the proposed CBW descriptor and the proposed Log-Min distance matching separately. In the experiments, the proposed CBW descriptor is integrated with the DP matching as used by IDSC+DP [9], SC+DP [15][9], and Height Function [36], as well as the Hausdorff distance matching, respectively. The results are reported in Tables I-III. It can been seen that when using the DP distance, the proposed CBW descriptor still achieves the highest MAP score of 71.6% on the compound leaves compared to all the benchmarks including the three DP based methods, IDSC+DP [9], SC+DP [9][15] and Height Function [36]. While for the self-overlapped leaf test, the proposed CBW with DP achieves a MAP score of 85.8% which is better than almost all the benchmarks except for SC+DP [9][15]. These results further demonstrate the stronger discriminative ability of the proposed CBW descriptor. It is also observed that when using the proposed Log-Min distance matching, the proposed CBW consistently achieved higher scores than using Hausforff and DP matching strategies, which validates the effectiveness and superiority of the proposed Log-Min distance.

4.2 MEW2012 Leaf Dataset

The Middle European Woody plants (MEW2012) [37] is a publicly available leaf dataset, which contains native or frequently cultivated trees and shrubs of the central Europe region. There are a total of 9,745 leaf images which belong to 153 species with at least 50 samples in each one. One sample image per species is shown in Fig. 10. It can be seen that among the many species in the dataset, the very minor differences of leaf margin patterns (as explained in Fig. 1) between them make retrieving similar leaf shapes become very challenging.

In our fourth experiment, each sample in the MEW2012 dataset is taken as a query to retrieve the similar ones from all the samples in the dataset. The MAP scores obtained by the proposed method and the seven state-of-the-art approaches are summarized in Table IV. We also record the average computation time of matching one query with all the 9,745 leaf shapes including the feature extraction time of the query shape and the matching time for all the comparative methods. The average number of the salient points on leaves extracted by [21] in this experiment is 127.

Fig.10. One hundred and fifty three sample images (one sample per species) from the MEW2012 leaf dataset [37].

Fig. 11. One hundred sample images (one sample per species) from the CVIP100 dataset [10].

It can be seen that on this challenging leaf margin pattern dataset, the proposed method achieves the best MAP score of 68.25% which is 13.27% higher than the second best method HSC and is more than 18% over the other benchmark methods. These attractive results demonstrate that the proposed CBW+LogMin method also has superior ability to distinguish the subtle differences of leaf margin patterns.

The retrieval time listed in Table IV is obtained by running all the competing algorithms on a workstation with Intel Core(TM) i7-4910MQ 2.9 GHz CPU and 32 GB RAM under MATLAB r2012a programming environment. It can be seen that the proposed CBW+LogMin method achieves an average 4.66 seconds for a single retrieval which is 18.3, 19.6, and 21.0 times faster than IDSC [9], SC+DP [15], and Height Function [36] respectively. The proposed method has a slightly slower retrieval speed than the MDM-CD-RM, MDM-ID-RA, and Complex Network methods. The HSC method is 27.4 times faster than the proposed method, with a tradeoff of 13.27% drop in accuracy.

Table IV. MAP score (%) on the MEW2012 leaf dataset.

Algorithm	MAP score (%)	Average retrieval time (s)
IDSC+DP [9]	45.36	85.23
SC+DP [9][15]	47.94	91.47
MDM-CD-RM [3]	39.14	1.56
MDM-ID-RA [3]	33.71	1.89
Height Function [36]	49.76	97.81
HSC [10]	54.98	0.17
Complex Network [46]	28.59	1.45
Proposed CBW+Hausdorff	**64.79**	**4.57**
Proposed CBW+DP	**62.05**	**75.64**
Proposed CBW+LogMin	**68.25**	**4.66**

In Table IV, the MAP scores for the proposed CBW with Hausdorff distance and the proposed CBW with DP matching are also reported. It can be seen that using the DP matching, the proposed CBW achieves a MAP score of 62.05% which is more than 7.07% higher than all the benchmarks and is more than 12.29% higher than the three DP based methods. It is also noticed that when using the proposed Log-Min distance, the MAP score of the proposed CBW is 3.46% and 6.20% higher than that of CBW+Hausdorff and CBW+DP, which indicates again superiority of the proposed Log-Min distance over the Hausdorff distance and DP matching for leaf image retrieval.

4.3 CVIP100 Leaf Dataset

The CVIP100 leaf database [10] contains 1,200 leaf images from 100 plant species with 12 samples in each class. Fig. 11 shows example leaves of the 100 species with one sample for each species. The MAP scores of the proposed method compared with the other six benchmark methods are tabulated in Table V. It is observed that the proposed CBW+LogMin method achieves a 90.72% MAP score which is more than 4.64% higher than the other compared methods. Consistent with the previous four experiments,

the proposed CBW+LogMin again demonstrates its superior performance over all the benchmarks. In this experiment, the average number of the salient points extracted by the algorithm is 127. Comparing the MAP score of the proposed CBW+DP with that of all the benchmarks, we can see that the proposed CBW still consistently performs better than all the benchmark methods including the three DP based methods. While comparing the MAP score of the proposed CBW+Log-Min with that of the proposed CBW+Hausdorff and CBW+DP, the results again indicate the effectiveness and superiority of the proposed Log-Min distance.

Table V. MAP score (%) on the CVIP100 leaf dataset.

Algorithm	MAP score (%)
IDSC+DP [9]	80.53
SC+DP [9][15]	81.05
MDM-CD-RM [3]	71.50
MDM-ID-RA [3]	68.03
Height Function [36]	84.25
HSC [10]	86.08
Complex Network [46]	60.84
Proposed CBW+Hausdorff	**89.87**
Proposed CBW+DP	**86.86**
Proposed CBW+LogMin	**90.72**

4.4 Effect of Parameters

Two parameters are involved in the proposed algorithm. One is the number of scales K which also determines $T = 2^{K+1}$. The other is the threshold τ used in the DCE algorithm [21] for salient point detection. The smaller the parameter τ is, the larger the number of the salient points will be extracted. When τ=0, all the sample points are taken as the salient points. To analyze the effects of the parameters K and τ on the performance of the proposed CBW+Log-Min, we conducted a group of experiments on the CVIP100 leaf dataset.

Firstly, we analyze the effect of varying τ on the performance of the proposed method by setting K=7. In our experiment, the MAP scores of the proposed CBW+LogMin are computed by varying the parameter τ from 0 to 0.1 with step size of 5×10^{-3} and the results are plotted as a curve in Fig. 12. Since the parameter τ controls the number of the salient points used in the proposed CBW+LogMin, the curve for the number of salient points versus the parameter τ is also plotted in Fig. 12.

From Fig. 12, it can be seen that when $\tau = 0$, i.e., all 256 sample points on the contour are used as salient points, the MAP score of the proposed CBW+LogMin is 91.01%. When τ increases to 0.005, 0.010 and 0.015 (the number of salient points decreases to 201, 179 and 159), the MAP score increases to be greater than 91.01%. It means that by only choosing 159 salient points for chord bunch walks, we can achieve a better MAP score than that of using all the 256 sample points for chord bunch walks. Even when τ increases to 0.1 which largely reduces the number of salient points to around 50 (less than 20% of the total sample points), we still achieved a high MAP score of 86.91% which is only 4.1% lower than the score obtained by using

all the sample points. These experimental results indicate that in the proposed CBW+LogMin, using the salient points instead of using all the sample points can make the algorithm more effective and efficient.
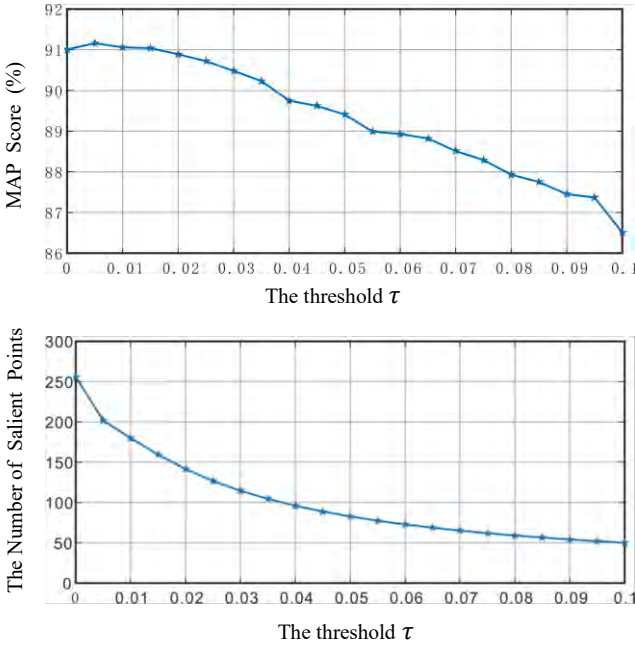

Fig. 12. The curves of the MAP score versus the threshold τ (above) and the number of salient points versus the threshold τ (below).

To study the effect of the parameter K on the performance of the proposed CBW+LogMin, we fix the parameter τ=0.025 and vary K from 2 to 11 with a step size of 1. The experimental results are plotted as a curve of MAP score versus the parameter K (see Fig. 13). As can be seen, the increase of the parameter K from 2 to 6 greatly improves the ability of the proposed CBW+LogMin to capture finer details of the shape. Then, the MAP score reaches its peak and remains stable at around 90% when K ranges from 6 to 8. Further increase of K from 8 to 11 will cause the performance drop of the proposed method because of the interference of noise it captures. Using a large number of scales will increase the computational cost for the proposed algorithm and degrade its efficiency.
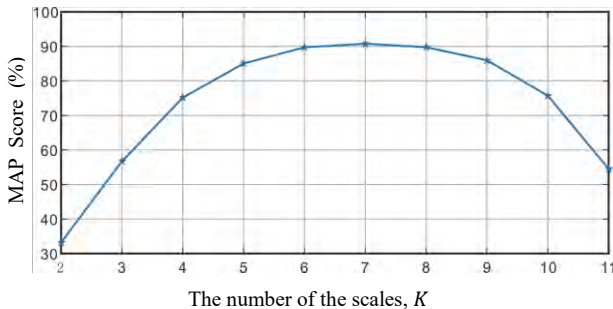

Fig. 13. The curve of the MAP score versus the number of the scales, K.

4.5 General Shape Recognition

Although the proposed CBW+LogMin is designed for recognizing naturally self-overlapped and compound leaves, we also tested its performance for general shape recognition using the MPEG-7 shape databases. The MPEG-7 CE-1 database [9][10][36] is a contour shape dataset in which each shape is enclosed by an outer contour without interior contents. This database has a total of 70 classes of images with each one containing 20 samples. However, each shape in this CE-1 dataset does not contain any interior content inside its contour that the proposed CBW+LogMin can use. The MPEG-7 CE-2 shape database is a region shape dataset suitable for evaluating CBW+LogMin, in which each shape contains interior content. However, some images contain multiple shape regions that are not connected each other, which cannot be used by the proposed method. We select all the images from the groups (i.e. classes) that contains 21 images and the shape regions are enclosed in one outer contour as a testing dataset for our experiment. It contains 336 images of 16 groups (see Fig. 14) and each group has 21 similar shapes including one original shape, five scaled versions, five rotated versions, and ten perspective transformed versions (see Fig. 15).

We conducted shape retrieval experiments on these two datasets respectively using the same evaluation protocol as used before. The MAP scores of the proposed CBW+LogMin together with the seven benchmark methods are displayed in Table VI. The proposed method achieves the highest score (88.20%) on the region shape dataset. On the contour shape dataset, the proposed method is ranked the fourth in accuracy (82.78%).

Fig.14. Example region shape images of sixteen groups selected from the MPEG-7 CE-2 database. One sample is shown for each group.
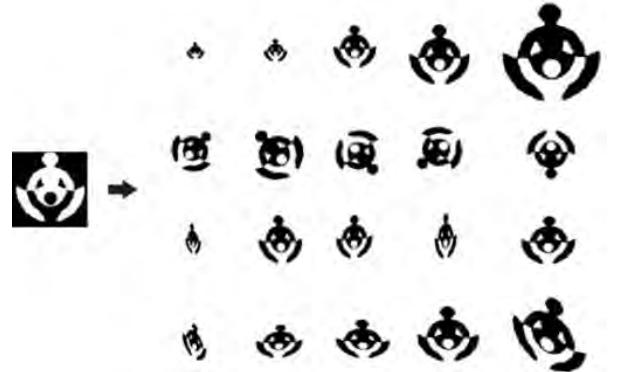

Fig.15. An example illustrating all 21 samples from a group shown in Fig. 14. They are: one original shape (the leftmost one), five scaled versions (the first row), five rotated versions (the second row), and ten perspective transformed versions (the last two rows).

Table VI. MAP score (%) on the MPEG-7 CE-1 dataset and the subset of the MPEG-7 CE-2 dataset.

Algorithm	MPEG-7 CE-1	The subset of MPEG-7 CE-2
IDSC+DP [9]	81.48%	84.31%
SC+DP [9][15]	83.80%	84.29%
MDM-CD-RM [3]	71.50%	77.34%
MDM-ID-RA [3]	68.03%	78.85%
Height Function [36]	86.52%	82.96%
HSC [10]	83.35%	85.94%
Complex Network [46]	63.21%	74.82%
Proposed CBW+LogMin	**82.78%**	**88.20%**

5. CONCLUSION

This paper presented a novel chord bunch walks approach for describing and matching shapes with self-overlaps. By walking and measuring along the chord bunch originated from each contour point, both the contour features and the inner properties are characterized. Each chord walk integrates the shape image function over the walked chord to capture the inner and outer distributions of the shape region. This can particularly handle the problem of describing compound leaves and self-overlapped leaves.

The chord bunch walks employ a hierarchical framework to provide a coarse-to-fine shape description for effectively differentiating subtle differences among leaf margin patterns of plant species. It is invariant to translation, rotation, scaling, and mirror transforms. The proposed Log-Min matching scheme provides significantly faster CBW matching than the widely used dynamic programming technique, making it suitable for large database retrieval tasks. Five experiments are conducted on three challenging leaf image datasets containing 16,846, 9,745, and 1,200 images respectively. The proposed CBW+Log-Min approach consistently outperforms the state-of-the-art methods in all the experiments, in particular, with larger accuracy increases on compound leaves and self-overlapped leaves. The idea presented in this paper may also be used in other similar applications.

APPENDIX A

Lemma 1: Let $W_t^s = \left(\hat{l}_t^{-s}, \breve{l}_t^{-s}, \sin(\theta_t^{-s} - \theta_t^s), \hat{l}_t^s, \breve{l}_t^s \right)$ be a quintuplet measurement for a chord pair walks. By redefining the measurement as

$$\bar{W}_t^s = \left(\max(\hat{l}_t^{-s}, \hat{l}_t^s), \max(\breve{l}_t^{-s}, \breve{l}_t^s), \sin(\theta_t^{-s} - \theta_t^s), \right.$$
$$\left. \min(\hat{l}_t^{-s}, \hat{l}_t^s), \min(\breve{l}_t^{-s}, \breve{l}_t^s) \right),$$

the set $\{\bar{W}_t^s, 0 \leq t < 1\}$ becomes invariant to mirror transform.

Proof:

Without losing generality, here we assume that the shape is reflected over the y-axis. Given a shape image function $f(x, y)$ together with its contour parameter equation $\dot{z}(t) = \left(\bar{x}(t), \bar{y}(t)\right), t \in [0,1)$, the mirrored version of the shape can be expressed as

$$\dot{f}(x, y) = f(-x, y) \tag{A-1}$$

and

$$\dot{z}(t) = \left(-\bar{x}(1 - t), \bar{y}(1 - t)\right), t \in [0,1). \tag{A-2}$$

The redefined chord pair walks measurement for the mirrored version of the shape is

$$\bar{\dot{W}}_t^s = \left(\max(\hat{l}_{1-t}^{-s}, \hat{l}_{1-t}^s), \max(\breve{l}_{1-t}^{-s}, \breve{l}_{1-t}^s), \sin((\pi - \theta_{1-t}^{-s}) - (\pi - \theta_{1-t}^s)), \min(\hat{l}_{1-t}^{-s}, \hat{l}_{1-t}^s), \min(\breve{l}_{1-t}^{-s}, \breve{l}_{1-t}^s) \right). \tag{A-3}$$

Because the redefined chord pair walks measurement for the shape is

$$\bar{W}_t^s = \left(\max(\hat{l}_t^{-s}, \hat{l}_t^s), \max(\breve{l}_t^{-s}, \breve{l}_t^s), \sin(\theta_t^{-s} - \theta_t^s), \min(\hat{l}_t^{-s}, \hat{l}_t^s), \min(\breve{l}_t^{-s}, \breve{l}_t^s) \right), \tag{A-4}$$

we have

$$\bar{\dot{W}}_t^s = \left(\max(\hat{l}_{1-t}^{-s}, \hat{l}_{1-t}^s), \max(\breve{l}_{1-t}^{-s}, \breve{l}_{1-t}^s), \sin(\theta_{1-t}^{-s} - \theta_{1-t}^s), \min(\hat{l}_{1-t}^{-s}, \hat{l}_{1-t}^s), \min(\breve{l}_{1-t}^{-s}, \breve{l}_{1-t}^s) \right) = \bar{W}_{1-t}^s. \tag{A-5}$$

Because a shape is described by a whole set of quintuplet measurements, we have

$$\left\{\bar{\dot{W}}_t^s, 0 \leq t < 1\right\} = \{\bar{W}_{1-t}^s, 0 \leq t < 1\} = \{\bar{W}_t^s, 0 \leq t < 1\}, \tag{A-6}$$

Thus, the set $\{\bar{W}_t^s, 0 \leq t < 1\}$ is invariant to mirror transform.

∎

Lemma 2: Given $\vartheta_{t,k}^{(L)}$ and $\vartheta_{t,k}^{(R)}$ that are defined in Eq. (5), the set $\left\{\left|\sin\left(\vartheta_{t,k}^{(R)} - \vartheta_{t,k}^{(L)}\right)\right|, 0 \leq t < 1\right\}$ is invariant to mirror transform.

Proof:

From Eq. (5), the $\vartheta_{t,k}^{(L)}$ and $\vartheta_{t,k}^{(R)}$ for the mirrored version of the shape can be calculated as

$$\dot{\vartheta}_{t,k}^{(R)} = \left(\pi - \theta_{1-t}^{-2^{-k}}\right) - \left(\pi - \theta_{1-t}^{-2^{-(k+1)}}\right)$$
$$= \theta_{1-t}^{-2^{-(k+1)}} - \theta_{1-t}^{-2^{-k}} = \vartheta_{1-t,k}^{(L)} \tag{A-7}$$

and

$$\dot{\vartheta}_{t,k}^{(L)} = \left(\pi - \theta_{1-t}^{2^{-(k+1)}}\right) - \left(\pi - \theta_{1-t}^{2^{-k}}\right)$$
$$= \theta_{1-t}^{2^{-k}} - \theta_{1-t}^{2^{-(k+1)}} = \vartheta_{1-t,k}^{(R)}. \tag{A-8}$$

Thus we have

$$\left|\sin\left(\dot{\vartheta}_{t,k}^{(R)} - \dot{\vartheta}_{t,k}^{(L)}\right)\right| = \left|\sin\left(\vartheta_{1-t,k}^{(L)} - \vartheta_{1-t,k}^{(R)}\right)\right|. \tag{A-9}$$

Because a shape is described by a whole set of measurements, we have

$$\left\{\left|\sin\left(\dot{\vartheta}_{t,k}^{(R)} - \dot{\vartheta}_{t,k}^{(L)}\right)\right|, 0 \leq t < 1\right\}$$
$$= \left\{\left|\sin\left(\vartheta_{1-t,k}^{(L)} - \vartheta_{1-t,k}^{(R)}\right)\right|, 0 \leq t < 1\right\}$$
$$= \left\{\left|\sin\left(\vartheta_{t,k}^{(R)} - \vartheta_{t,k}^{(L)}\right)\right|, 0 \leq t < 1\right\}. \tag{A-10}$$

Thus, the set $\left\{ \left| \sin\left(\vartheta_{t,k}^{(R)} - \vartheta_{t,k}^{(L)} \right) \right|, 0 \leq t < 1 \right\}$ is invariant to mirror transform. ∎

REFERENCES

[1] G. Bharathan and N.R. Sinha, "The Regulation of Compound Leaf Development," Plant Physiology, vol. 127, no. 4, pp. 1533-1538, 2001.

[2] S.G. Wu, F.S. Bao, E.Y. Xu, Y.-X. Wang, Y.-F. Chang, and Q.-L. Xiang, "A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Nerual Network," in Proc. IEEE Int. Sym. Signal Processing and Information Technology, vol.1, pp. 11-16, 2007.

[3] R. Hu, W. Jia, H. Ling, and D. Huang, "Multiscale Distance Matrix for Fast Plant Leaf Recognition," IEEE Trans. Image Processing, vol. 21, no.11, pp. 4667-4672, 2012.

[4] C.D. Manning, P. Raghavan, and H. Schütze, "Introduction to information retrieval," Cambridge University Press, New York, NY, USA, 2008.

[5] H. Laga, S. Kurtek, A. Srivastava, M. Golzarian, and S.J. Miklavcic, "A Riemannian Elastic Metric for Shape-based Plant Leaf Classification," in Proc. Int. Conf. Digital Image Computing: Techniques and Applications (DICTA), vol. 1, pp. 1-7, 2012.

[6] S. Mouine, I. Yahiaoui, and A. Verroust-Blondet, "A Shape-based Approach for Leaf Classification Using Multiscale Triangular Representation," in: Proc. The 3rd ACM Int. Conf. Multimedia Retrieval (ICMR), vol. 1, pp. 127-134, 2013.

[7] T.V. Hoang and S. Tabbone, "Invariant Pattern Recognition Using the RFM Descriptor," Pattern Recognition, vol. 45, no. 1, pp. 271-284, 2012.

[8] D. Zhang and G. Lu, "Study and Evaluation of Different Fourier Methods for Image Retrieval," Image Vis. Comput., vol. 23, no. 1, pp. 33-49, 2005.

[9] H. Ling, D.W. Jacobs, "Shape Classification Using the Inner-Distance," IEEE Trans. Pattern Anal. Machine Intell., vol. 29, no. 2, pp. 286-299, 2007.

[10] B. Wang and Y. Gao, "Hierarchical String Cuts: A Translation, Rotation, Scale and Mirror Invariant Descriptor for Fast Shape Retrieval," IEEE Trans. Image Processing, vol. 23, no. 9, pp. 4101-4111, 2014.

[11] Y.W. Chen and Y.Q. Chen, "Invariant Description and Retrieval of Planar Shapes Using Radon Composite Features," IEEE Trans. Signal Processing, vol. 56, no. 10, pp. 4762-4771, 2008.

[12] P.-T. Yap, X. Jiang, and A.C. Kot, "Two-Dimensional Polar Harmonic Transforms for Invariant Image Representation," IEEE Trans. Pattern Anal. Machine Intell., vol. 32, no. 7, pp. 1259-1270, 2010.

[13] S. Manay, D. Cremers, B.-W Hong, A.J. Yezzi, and S. Soatto, "Integral Invariants for Shape Matching," IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 10, pp. 1602-1618, 2006.

[14] M.-P. Dubuisson and A.K. Jain, "A Modified Hausdorff Distance for Object Matching," in Proc. The Twelfth Int. Conf. Pattern Recognition (ICPR), pp. 566-568, 1994.

[15] S. Belongie, J. Malik, and J. Puzicha, "Shape Matching and Object Recognition Using Shape Contexts," IEEE Trans. Pattern Anal. Machine Intell., vol. 24, no.4, pp. 509-522, 2002.

[16] C. Grigorescu and N. Petkov, "Distance Sets for Shape Filters and Shape Recognition," IEEE Trans. Image Processing, vol. 12, no. 10, pp. 1274-1286, 2003.

[17] N. Alajlan, I.E. Rube, M.S. Kamel and G. Freeman, "Shape Retrieval Using Triangle-area Representation and Dynamic Space Warping," Pattern Recognition, vol. 40, no. 7, pp. 1911-1920, 2007.

[18] I. Biederman, G. Ju, "Surface versus Edge-based Determinants of Visual Recognition," Cognitive Psychology, vol. 20, no.1, pp. 38-64, 1988.

[19] L. Schomaker, E. de Leau, and L. Vuurpijl, "Using Pen-based Outlines for Object Based Annotation and Image-based Queries," in Proc. Third Int. Conf. Advances in Visual Information Systems, Amsterdam, The Netherlands, June, pp. 585-592, 1999.

[20] Y. Gao and M.K.H. Leung, "Face Recognition Using Line Edge Map," IEEE Trans. Pattern Anal. Machine Intell., vol. 24, no. 6, pp. 764-779, 2002.

[21] L.J. Latecki and R. Lakämper, "Convexity Rule for Shape Decomposition Based on Discrete Contour Evolution," Comput. Vis. Image Understand., vol. 73, no. 3, pp. 441-454, 1999.

[22] C.-L. Chou, H.-T. Chen, S.-Y. Lee. "Pattern-based Near-Duplicate Video Retrieval and Localization on Web-scale Videos," IEEE Trans. Multimedia, vol. 17, no. 3, pp. 382-395, 2015.

[23] C. Xu, J. Liu, and X. Tang. "2D Shape Matching by Contour Flexibility," IEEE Trans. Pattern Anal. Machine Intell., vol. 31, no. 1, pp. 180-186, 2009.

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

[25] B. Wang, D. Brown, Y. Gao, and J. La Salle, "MARCH: Multiscale-arch-height Description for Mobile Retrieval of Leaf Images," Information Sciences, vol. 302, pp. 132-148, 2015.

[26] B. Wang, Y. Gao, C. Sun, M. Blumentein, and J. La Salle, "Can Walking and Measuring along Chord Bunches Better Describe Leaf Shapes?," in Proc. IEEE Conf. Comput. Vis. Pattern Recognition (CVPR), pp. 2047-2056, 2017.

[27] Y.-H. Yang, W.H. Hsu, and H.H. Chen. "Online Reranking via Ordinal Informative Concepts for Context Fusion in Concept Detection and Video Search," IEEE Trans. Circuits and Systems for Video Technology, vol. 19, no. 12, pp. 1880-1890, 2009.

[28] Y.-C. Pan and L.-S. Lee. "Performance Analysis for Lattice-Based Speech Indexing Approaches Using Words and Subword Units," IEEE Trans. Audio, Speech, and Language Processing, vol. 18, no. 6, pp. 1562-1574, 2010.

[29] S. Tabbone, L. Wendling and J.-P. Salmon, "A New Shape Descriptor Defined on the Radon Transform," Computer Vision and Image Understanding, vol. 102, no. 1, pp. 42-51, 2006.

[30] X. Wang, B. Xiao, J.-F. Ma, and X.-L. Bi, "Scaling and Rotation Invariant Analysis Approach to Object Recognition Based on Radon and Fourier-Mellin Transforms," Pattern Recognition, vol. 40, no. 12, pp. 3503-3508, 2007.

[31] J. Panda, M.S. Brown, and C.V. Jawahar. "Offline Mobile Instance Retrieval with a Small Memory Footprint," in Proc. IEEE Int. Conf. Computer Vision (ICCV), pp. 1257-1264, 2013.

[32] W. Dong, Z. Wang, M. Charikar and K. Li. "Efficiently Matching Sets of Features with Random Histograms," in Proc. The 16th ACM Int. Conf. Multimedia, pp. 179-188, 2008.

[33] A. Kadyrov and M. Petrou, "The Trace Transform and Its Applications," IEEE Trans. Pattern Anal. Machine Intell, vol. 23, no. 8, pp. 811-828, 2001.

[34] B.-W. Hong and S. Soatto, "Shape Matching Using Multiscale Integral Invariants," IEEE Trans. Pattern Anal. Machine Intell, vol. 37, no. 1, pp. 151-160, 2015.

[35] F. Janan and M. Brady, "Shape Descrption and Matching Using Integral Invariants on Eccentricity Transformed Images", Int. J. Comput. Vis., vol. 113, no. 2, pp. 92-112, 2015.

[36] J. Wang, X. Bai, X. You, W. Liu, and L.J. Latecki, "Shape Matching and Classification Using Height Functions," Pattern Recognition Letters, vol. 33, no. 2, pp. 134-143, 2012.

[37] P. Novotný and T. Suk, "Leaf Recognition of Woody Species in Central Europe," Biosystems Engineering, vol. 115, no. 4, pp. 444-452, 2013.

[38] C. Zhao, S.S.F. Chan, W.-K. Cham, and L.M. Chu, "Plant Identification Using Leaf Shapes–A Pattern Counting Approach," Pattern Recognition, vol. 48, no. 10, pp. 3203-3215, 2015.

[39] A. Khotanzad and Y.H. Hong, "Invariant Image Recognition in Zernike Moments," IEEE Trans. Pattern Anal. Machine Intell., vol. 12, no. 5, pp. 489-497-261, 1990.

[40] C.T. Zahn and R.Z. Roskies, "Fourier Descriptors for Plane Closed Curves," IEEE Trans. Comput. vol. 21, no. 3, pp. 269-281, 1972.

[41] Q.M. Tieng and W.W. Boles, "Recognition of 2D Object Contours Using the Wavelet Transform Zero-crossing Representation," IEEE Trans. Pattern Anal. Machine Intell., vol. 19, no. 8, pp. 910-916, 1997.

[42] N. Kumar, P.N. Belhumeur, A. Biswas, D.W. Jacobs, W.J. Kress, I.C. Lopez, and J.V.B. Soares, "Leafsnap: A Computer Vision System for Automatic Plant Species Identification," in Proc, Eur. Conf. Comput. Vis. (ECCV), pp. 502-516, 2012.

[43] N. Arica and F.T.Y. Vural, "BAS: a Perceptual Shape Descriptor Based on the Beam Angle Statistics," Pattern Recognition Letters, vol. 24, no. 9-10, pp.1627-1639, 2003.

[44] Z. Wang, Z. Chi and D. Feng, "Shape Based Leaf Image Retrieval," IEE Proceedings – Vision, Image and Signal Processing, vol. 150, no. 1, pp. 34-43, 2003.

[45] J. Charters, Z. Wang, Z. Chi, A.C. Tsoi, and D. Feng, "EAGLE: a Novel Descriptor for Identifying Plant Species Using Leaf Lamina Vascular Features," in Proc. IEEE Conf. Multimedia & Expo (ICME), 2014.

[46] A.R. Backes, D. Casanova, and O.M. Bruno, "A Complex Network-based Approach for Boundary Shape Analysis," Pattern Recognition, vol. 42, no. 1, pp.54-67, 2009.

[47] L.C. Ribas, M.B. Neiva, and O.M. Bruno, "Distance Transform Network for Shape Analysis," Information Sciences, vol. 470, pp. 28-42, 2019.

[48] T. Jin, X. Hou, P. Li, F. Zhou, "A Novel Method of Automatic Plant Species Identification Using Sparse Representation of Leaf Tooth Features," PLoS ONE, vol 10, no. 10, e0139482. https://doi.org/10.1371/journal.pone.0139482, 2015.

[49] J. Cope, P. Remagnino, "Classifying Plant Leaves from their margins Using Dynamic Time Warping," In: Advanced Concepts for Intelligent Vision Systems, Lecture Notes in Computer Sciences, vol 7517, Springer, Berlin, pp. 258-267, 2012.

[50] G. Cerutti, L. Tougne, D. Coquin, A. Vacacant, "Leaf Margins as Sequences: A Structural Approach to Leaf Identification," Pattern Recognition Letter, vol 49, pp. 177-184, 2014.

Bin Wang received the PhD degree in computer science in 2007 from Fudan University, Shanghai, China. Since 2007, he had been on the faculty of Nanjing University of Finance and Economics, Nanjing, China, where he became a professor in 2015. 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. In October 2015, he joined the School of Engineering at Griffith University, Australia, where he is currently a research fellow. His main research interests include computer vision, image processing and pattern recognition. He is a member of IEEE.

Yongsheng Gao received the B.Sc. and M.Sc. degrees in electronic engineering from Zhejiang University, Hangzhou, China, in 1985 and 1988, respectively, and the Ph.D. degree in computer engineering from Nanyang Technological University, Singapore. He is currently a Professor with the School of Engineering, Griffith University, Brisbane, QLD, Australia. He had been the Leader of Biosecurity Group, Queensland Research Laboratory, National ICT Australia (ARC Centre of Excellence), a consultant of Panasonic Singapore Laboratories, and an Assistant Professor in the School of Computer Engineering, Nanyang Technological University, Singapore. His research interests include smart farming, intelligent agriculture, biosecurity, face recognition, biometrics, image retrieval, computer vision, pattern recognition, environmental informatics, and medical imaging. He is a senior member of IEEE.

Changming Sun received his PhD degree in computer vision from Imperial College London, London, UK in 1992. He then joined CSIRO, Sydney, NSW, Australia, where he is currently a Principal Research Scientist carrying out research and working on applied projects. His current research interests include computer vision, image analysis, and pattern recognition. He has served on the program/organizing committees of various international conferences. He is an Associate Editor of the EURASIP Journal on Image and Video Processing. He is a member of the Australian Pattern Recognition Society.

Michael Blumenstein received his PhD in Computational Intelligence from Griffith University in Queensland, Australia in 2001. He is a Professor and Associate Dean Research in the Faculty of Engineering and IT at the University of Technology Sydney, Australia. Michael is an internationally and nationally renowned expert in the areas of Pattern Recognition and Artificial Intelligence (specifically Machine learning and Neural Networks). He has published over 200 papers in refereed conferences, journals and books in these areas. His research also spans various projects applying Artificial Intelligence to the fields of

Engineering, Environmental Science, Neurobiology and Coastal Management. Michael has secured internal/nationally competitive research grants to undertake these projects with funds exceeding AUD$4.6 Million. Components of his research into the predictive assessment of beach conditions have been commercialized for use by local government agencies, coastal management authorities and in commercial applications.

John La Salle received his PhD in Entomology from the University of California Riverside in 1984. He is an internationally recognised insect taxonomist, who served as Director of the Australian National Insect Collection (ANIC) from 2001-2012 and since then has served as the Director of the Atlas of Living Australia. His current research interests are in adopting emerging technologies to accelerate the processes of taxonomy, species discovery and description, and making biodiversity information accessible and useable online.