19th International Computer Science and Engineering Conference (ICSEC)
Chiang Mai, Thailand, 23-26 November, 2015

Detection of Leaf Apex and Base by Using Contour and Symmetry Analysis

Ukrit Watchareeruetai[1], Matchima Ditthawibun[2], and Kriangkrai Phanjan[3]
International College
King Mongkut's Institute of Technology Ladkrabang
Chalongkrung Road, Ladkrabang, Bangkok 10520 Thailand
ukrit@ieee.org[1], s6090011@kmitl.ac.th[2], s4090005@kmitl.ac.th[3]

Abstract**—**This paper proposes a contour analysis method for automatic detection of leaf apex and base. The contour is investigated to determine the optimal pair of points that are possibly the apex and base. The results are then affirmed using a leaf symmetry analysis and post-processing. Experimental results show that the proposed method can detect apex and base with an accuracy of 85%, given a 5 mm distance error.

Keywords—leaf apex; leaf base; leaf identification; contour analysis; image processing

I. Introduction

Considering that plants are inevitably crucial to human in various ways, abilities to identify and classify the species of plants are essential. Generally, these tasks are carried out by botanists inspecting the morphological characteristics of plants including leaves, flowers, stems, roots, etc. The scarcity of botanists may prevent other professions from exploiting this beneficial information. Nowadays, plant specimens can be easily digitized by using a digital camera or scanner, leading to an idea to create an automated system for identifying and classifying plants with the use of image processing and computer vision techniques. Presently, a number of plant identification algorithms, which are mostly based on leaf characteristics, have been proposed [3]. Mobile applications for plant identification based on leaves are also available in the market, recently [9, 10].

Apex and base are also one of the key characteristics of leaves often used by botanists. For the reason that the apex and base are landmarks presenting in almost every taxa of leaves, many leaf identification algorithms used apex and base as references in calculation of leaf features. For example, Wu et al. [13] and Iwata and Saitoh [8] defined a feature called physiological length, which was also used in the definition of several other features, as the distance between the apex and base of a leaf. Cerutti et al. proposed a deformable model obtained from given apex and base of leaves [1, 2]. The set of parameters describing the model was used as features in classification. The positions of apex and base were also essential in feature extraction from leaf margin. In particular, they partitioned a leaf contour into areas of apex, base, lobe tips, and the remaining which was used to compute margin descriptors.

However, Cope et al. [3] suggested that the use of landmarks has several limitations, including that landmarks on leaves are difficult for automatic extraction. They also mentioned that usually the extraction of landmarks is manually carried out by human experts. Some leaf identification algorithms require this information as the input from the user [13] or allow the user to adjust the position of apex and base detected by the algorithm [12]. The difficulties in detection of apex and base are caused by several reasons. First, there is a wide variety of apex- and base-shapes. As shown in Fig. 1, they can be sharply pointed, rounded, blunt, etc., and can also be either a convex or concave point on the boundary of leaves. Dozens of words have been used by botanists to define the shape of apex and base. Botanists also use the angle of apex and base to describe leaves. However, it can vary from an acute angle (< 90 degrees) to a reflex angle (between 180 and 360 degrees). Second, the intra-class variation of some species of leaves is very high. It is often found that two leaves from the same species have different shapes of apex and base. Figure 2 presents an example of a leaf species that has different base shapes. Third, a leaf may have several other landmarks such as lope tips and teeth which form a set of distinguish points on the leaf boundary. A simple detection of an abrupt change along the contour would include these other landmarks, resulting in necessity of a post-processing method to obtain only the apex and base points.

Until now, although there are very few papers regarding automated detection of apex and base, unfortunately, the performance evaluation of apex/base detection methods has never been reported. Corney et al. [4] proposed an algorithm that specifically detects the apex and base from herbarium specimens of *Tilia* L. leaves for the extraction of blade length. The method used Hough transform to identify the midvein, which links the base to apex. A leaf-symmetry test and local morphology are also used to verify the detected result. Software named LeafAnalyser developed by Weight et al. [12] was used to analyze the variation of leaf shape. The position of leaf tip was used to align a leaf before performing principle component analysis. Although Weight et al. reported that LeafAnalyser could automatically detect the tip of a leaf by examining leaf contour, a clear explanation about the algorithm was not provided. Another work by Im et al. [6, 7] analyzed the curvature of smoothed contour of leaf and finds critical points. They designated as an apex (or lobe tip) if a critical point has a positive curvature. Ma et al. [11] proposed a method to localize the position of leaf apex from 3D data. Cerutti et al. [1] proposed a polygon model for leaf segmentation and identification. During the first stage of segmentation process, a polygonal shape model was evolved to fit the given leaf shape. Then an active contour segmentation based on the evolved model was adopted to extract leaf contour. From the contour, the position of apex, base, and lope tips were located as the point in the corresponding areas with highest curvature.

In this work, we propose a new approach solely using the contour of a leaf image to automatically determine the position of apex and base of the leaf based on contour and symmetry analysis. The leaf image is firstly segmented to obtain the contour, which is then analyzed for its signature. Pairs of contour points with high degree of curvature are then evaluated using leaf symmetry analysis to obtain the optimal pair which represents the apex and base. The accuracy of the proposed algorithm will be measured in the experiments.

The rest of this paper is organized as follows. Section II explains the detail of the proposed method for detection of leaf apex and base. Section III describes the experiment set-up and discusses the experimental results. Section IV concludes the paper.

II. Proposed Method

The proposed apex/base detection method is depicted as a flowchart, as shown in Fig. 3. It can be divided into four main steps, as follows.

A. Candidate point selection

In this step, the contour of leaf is analyzed to find a set of candidate points, which are chosen based upon the absolute difference of tangent angles of the contour. In particular, the tangent angle θ_i of a contour point i is estimated as the angle of a line segment from the point i to the point $i + step$ on the contour, where $step$ is a step size used for estimation of tangent (equal 5 in this work). For each point i on the contour, the absolute difference of tangent angles $|\Delta\theta_i| = |\theta_i - \theta_{i-1}|$ is computed. All contour points that have this value larger than a predefined threshold are chosen as candidates. The threshold value is 10 in this work.

B. Candidate pairing

Each candidate point obtained from the previous process is then paired with a point on the opposite side of the contour to form a candidate pair of apex and base. Suppose the number of points on the contour is s, the pair of a candidate point i is chosen from contour points in the range $(i + s / 2) \pm r$, where r is a searching range. A contour point j in the range with highest value of absolute tangent angle difference is chosen as the counterpart of i, and forms a pair of candidates (i, j).

C. Symmetry analysis

The goal of this step is to use leaf-symmetry to test which candidate pair is most likely the pair of the apex and base. For many leaves whose shape is symmetry or almost-symmetry, a line segment from apex to base could roughly form an axis of symmetry. Consequently, in this algorithm, each candidate pair of apex/base is tested for its symmetry and gives a symmetry score. A pair with highest value of symmetry score is chosen as the pair of apex and base. The following is the algorithm for calculation of the symmetry score:

Fig. 1. Variety of leaf shapes (images from Flavia dataset [13])

Fig. 2. Leaves from the same species having different shapes of base (images from Flavia dataset [13])

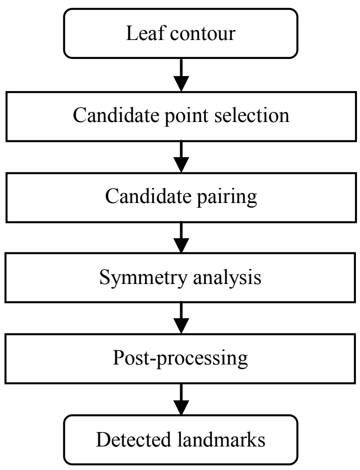

Fig. 3. Flowchart of the proposed method

1) Sampling the line segment between the candidates: Sampling N points (5 in this work) on the line segment between the two candidates with an equal distance.

2) Drawing perpendicular lines: For each sampled point, draw a perpendicular line and find its crossing points with the contour. The distances from the sampled point i to the left and right crossing points are computed (l_i and r_i) (Fig. 4).

3) Calculation of symmetry score: The symmetry score is defined based on the values of l_i and r_i as follows:

$$S_N = W \times \left(1 - \sum_{i=1}^{N} \frac{|l_i - r_i|}{l_i + r_i} \right), \qquad (1)$$

$$W = \begin{cases} 1 & \text{if } \dfrac{l}{w} \geq 1 \\ \dfrac{l}{w} & \text{otherwise} \end{cases}, \qquad (2)$$

where l is the distance between a pair of candidates and w is the average value of (l_i+r_i). Because it is possible that a leaf may have more than one symmetry axes, we use an assumption that the length of a leaf is usually larger than its width to help identify the correct one. If a pair of candidates yields a relatively short distance, a smaller weight for the symmetry score will be applied.

Figure 5 presents some examples of symmetry analysis for different species of leaves. Both ends of the symmetry axis indicate the position of points that should be apex and base.

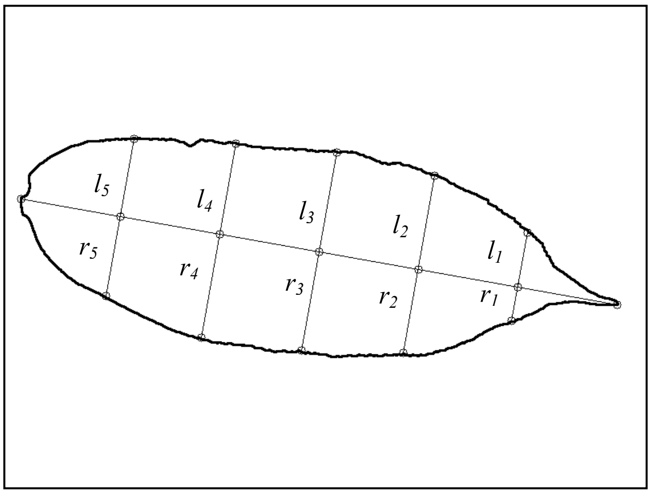

Fig. 4. Symmetry analysis

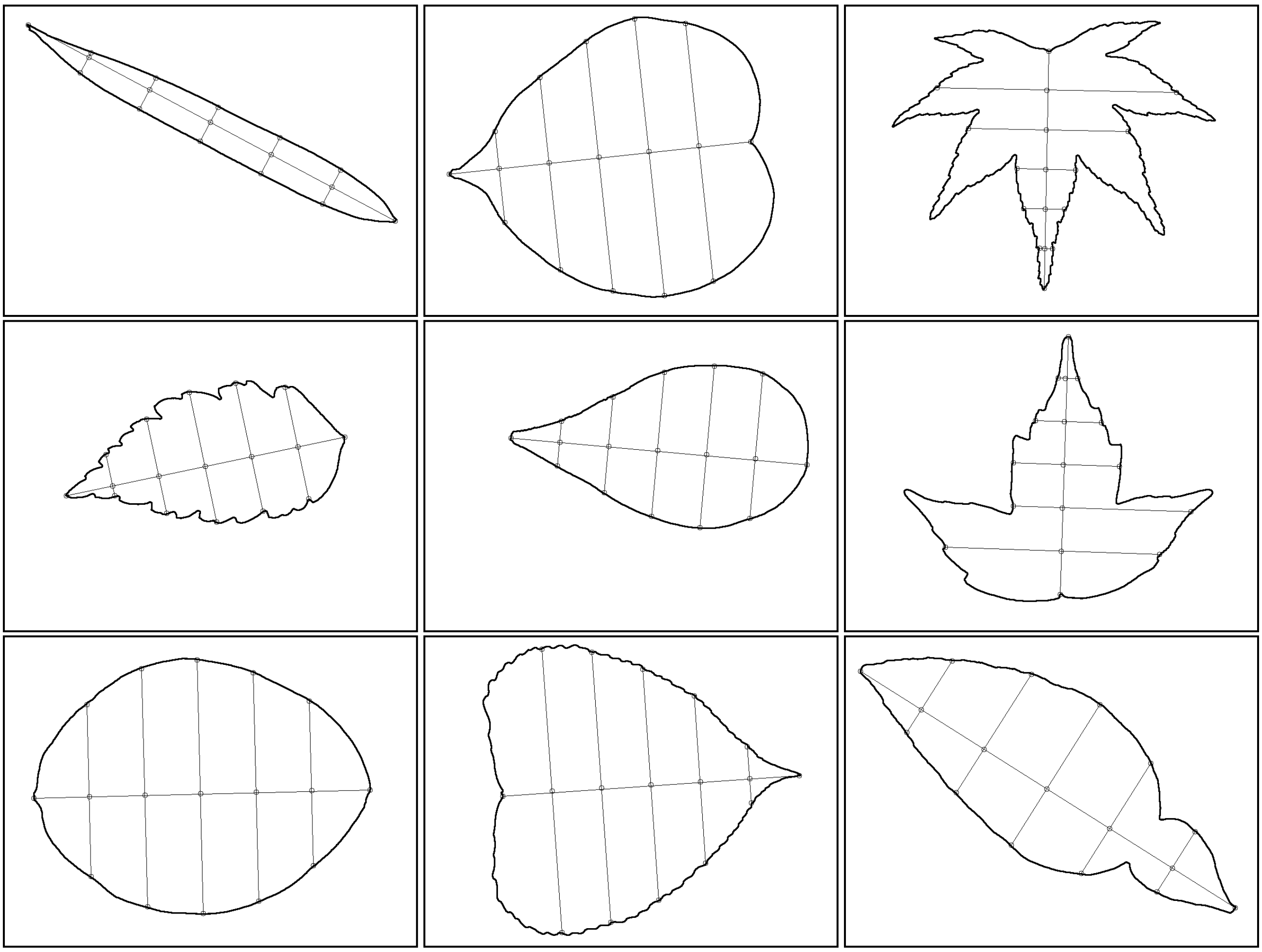

Fig. 5. Examples of symmetry analysis results

D. Post-processing

After a pair of apex/base is found, it undergoes a post-processing which verifies the optimality of points by comparing with nearby points. This step, sometimes, also helps us dealing with leaves that are not symmetric. Let i be the index of a detected landmark, all points of the contour with an index in the range $[i-P, i+P]$ are considered as the nearby points. A nearby point with highest value of absolute tangent difference is chosen to replace the detected landmark. Note that the value of P is 5% of the contour length in this work.

III. EXPERIEMENTAL RESULTS AND DISCUSSION

Experiments have been conducted in order to measure the effectiveness of the proposed method for detection of apex and base.

A. Image Dataset

In this work, a set of 320 leaf images (32 species, 10 images from each) from Flavia dataset [13] was used for the evaluation. The leaf images in Flavia dataset were scanned individually on a white background at 300 dpi, resulting in images of size 1600×1200 pixels. However, in the experiments, the size of images was reduced to 800×600 pixels (i.e., 150 dpi) to increase the computation speed. Therefore the actual size of a pixel in the test images was around 25.4 / 150 = 0.1693 mm.

B. Experiment Set-up

Contour extraction was performed as follows (Fig. 6). First, an input leaf image was converted into a gray-scale image. Second, a global thresholding was applied to convert the gray-scale image into a binary image representing the areas of a leaf and background. Two morphological operations, i.e., opening and closing, were then applied to remove noise, fill holes, and connect the area of leaf together. Next, the contour of leaf was obtained by finding the outer boundary of the object in the image.

For each image, the position of apex and base were manually marked and were compared with the landmarks detected by the proposed method. Note that the proposed method can only deliver a pair of landmarks without knowing which one is apex or base. Consequently, in the experiments, the actual position of apex (and base) was compared with the closer detected landmark to calculate the distance error, which is the distance from a detected landmark to its corresponding point. The percentage of correctly detected landmarks, defined as points whose distance error is smaller than a given range of acceptable error, was then computed and reported.

C. Results and Discussion

Table I shows the percentage of correctly detected points, given different levels of distance error ranging from 6 to 30 pixels (around 1 – 5 mm). If the distance error of a detected point was smaller than this given value, that detected point was counted as a correct detection; otherwise, an incorrect detection. The results indicate that the proposed method could detect apex and base with an accuracy of 85.63% for a distance error less than 5 mm, and suggest that 60% of detections were very precise (with error less than 1 mm). Note that these values of distance error were relatively small, compared to the physiological length of leaves (defined as the distance between the apex and base) which ranged from 43.21 to 154.67 mm (255.16 to 913.40 pixels), with the average of 116.18 mm (686.12 pixels).

In addition, the results suggest that detection of apex seem to be easier than detection of base points. According to our verification of the results, the reason would be that apex often had a sharply pointed shape, which was easier than a round or blunt shape to be chosen as a candidate. On the other hands, base points, which were more often rounded or blunt, were sometimes missing from the candidate point selection.

Table II presents the percentage of correctly detected apex and base points, separated by species (given a distance error of 30 pixels or around 5 mm). The proposed method could accurately detect (with an accuracy of 80% or higher) the apex and base of leaves for 25 out of 32 species. Among them the apex and base of 13 species could be perfectly detected. However, there were a few species of leaves on which the proposed method could not perform well. Two possible causes of the failures have been found. Firstly, the set of chosen candidate points did not include the apex or base of a leaf because their shapes were rounded or blunt, resulting a very small curvature. Secondly, a few species of leaves in the dataset were not symmetric and we model the symmetry axis of a leaf as a line segment. In this case, the proposed method could often correctly detect only one of them and missed the other. Although the post-processing step could sometimes correct the position of a detected apex/base point, in some cases, in which a detected point located far away from the real one, it still could not locate the correct position.

As a future work, a modification of symmetry analysis should be further studied. Rather than using a straight line, a polygonal line should be used to model the symmetry axis of leaves. Moreover, texture features formed by veins and veinlets should be analyzed. They would contain much information that can be used as a clue to help us confirm whether detected apex and base points are correct.

Fig. 6. Contour extraction process

TABLE I. PERCENTAGE OF CORRECTLY DETECTED APEX AND BASE POINTS (%)

Level of errors in pixels	Landmarks		
	Apex	*Base*	*Total*
≤ 6 (1.016 mm)	68.44	51.88	60.16
≤ 12 (2.031 mm)	80.94	77.81	79.38
≤ 18 (3.047 mm)	87.19	82.81	85.00
≤ 24 (4.063 mm)	87.50	83.44	85.47
≤ 30 (5.080 mm)	87.50	83.75	85.63

TABLE II. PERCENTAGE OF CORRECLTY DETECTED APEX AND BASE POINTS (WITH ERRORS ≤ 30 PIXELS) SEPARATED BY SPECIES (%)

Leaf species	**Apex**	**Base**	**Total**
Phyllostachys edulis (Carr.) Houz.	100	100	100
Aesculus chinensisy	100	100	100
Berberis anhweiensis Ahrendt	100	100	100
Cercis chinensis	100	100	100
Indigofera tinctoria L.	100	100	100
Acer Palmatum	40	50	45
Phoebe nanmu (Oliv.) Gamble	100	100	100
Kalopanax septemlobus (Thunb. ex A.Murr.) Koidz.	90	70	80
Cinnamomum japonicum Sieb.	100	100	100
Koelreuteria paniculata Laxm.	70	90	80
Ilex macrocarpa Oliv.	90	90	90
Pittosporum tobira (Thunb.) Ait. f.	100	80	90
Chimonanthus praecox L.	100	90	95
Cinnamomum camphora (L.) J. Presl	100	100	100
Viburnum awabuki K.Koch	90	90	90
Osmanthus fragrans Lour.	90	90	90
Cedrus deodara (Roxb.) G. Don	100	100	100
Ginkgo biloba L.	80	60	70
Lagerstroemia indica (L.) Pers.	90	100	95
Nerium oleander L.	100	100	100
Podocarpus macrophyllus (Thunb.) Sweet	100	100	100
Prunus serrulata Lindl. var. lannesiana auct.	90	100	95
Ligustrum lucidum Ait. f.	100	90	95
Tonna sinensis M. Roem.	100	20	60
Prunus persica (L.) Batsch	100	100	100
Manglietia fordiana Oliv.	100	100	100
Acer buergerianum Miq.	90	80	85
Mahonia bealei (Fortune) Carr.	40	50	45
Magnolia grandiflora L.	100	90	95
Populus ×canadensis Moench	60	50	55
Liriodendron chinense (Hemsl.) Sarg.	30	40	35
Citrus reticulata Blanco	50	50	50
Total	87.50	83.75	85.63

IV. CONCLUSION

This paper proposed a method to detect the apex and base of leaves based on contour and symmetry analysis. The method analyses the contour of leaf and determines the candidate points of apex/base. Leaf symmetry is then tested to find the best pair of apex and base. A post-processing step is applied to adjust the position of apex and base. The proposed method was tested with 320 leaves images (32 species) from Flavia dataset. The experimental results indicate that the proposed method could detect the pair of apex and base with an accuracy of 85% with a distance error smaller than 5 mm.

REFERENCES

[1] G. Cerutti, L. Tougne, J. Mille, A. Vacacant, and D. Coquin, “Guiding active contours for tree leaf segmentation and identification,” CLEF 2011 Labs and Workshop, Notebook Papers, 2011.

[2] G. Cerutti, L. Tougne, J. Mille, A. Vacacant, and D. Coquin, “Understanding leaves in natural images: a model-based approach for tree species identification,” Computer Vision and Image Understanding, vol.117, no.10, pp.1482-1501, 2013.

[3] J.S. Cope, D. Corney, J.Y. Clark, P. Remagnino, and P. Wilkin, “Plant species identification using digital morphometrics: a review,” Expert Systems with Applications, vol. 39, pp. 7562-7573, 2012.

[4] D. Corney, J.Y. Clark, H.L. Tang, and P. Wilkin, “Automatic extraction of leaf characters from herbarium specimens,” Taxon, vol. 6, no. 1, pp. 231-244, 2012.

[5] B. Ellis, D.C. Daly, L.J. Hickey, K.R. Johnson, J.D. Mitchell, P. Wilf et al., Manual of Leaf Architecture, Cornell University Press, 2009.

[6] C. Im, H. Nishida, and T.L. Kunii, “Recognizing plant species by leaf shapes: a case study of the Acer family,” Proceedings of International Conference on Pattern Recognition (ICPR’98), vol.2, pp.1171-1173, 1998.

[7] C. Im, H. Nishida, and T.L. Kunii, “A hierarchical method of recognizing plant species by leaf shapes,” Proceedings of IAPR Workshop on Machine Vision Applications (MVA’98), pp.158-161, 1998.

[8] T. Iwata and T. Saitoh, “Tree recognition based on leaf image,” Proceedings of SICE Annual Conference 2013, pp.2489-2494, 2013.

[9] N. Kumar, P.N. Belhumeur, A. Biswas, D.W. Jacobs, W.J. Kress, et al., “LeafSnap: a computer vision system for automatic plant species identification,” Proceedings of 12th European Conference on Computer Vision (ECCV 2012), vol.2, pp.502-516, 2012.

[10] H. Liu, D. Coquin, L. Valet, and G. Cerutti, “Leaf species classification based on a botanical shape sub-classifier strategy,” Proceedings of 22nd International Conference on Pattern Recognition (ICPR 2014), pp.1496-1501, 2014.

[11] W. Ma, H. Zha, J. Lie, X. Zhang, and B. Xiang, “Image-based plant modeling by knowing leaves from their apexes,” Proceedings of 19th International Conference on Pattern Recognition (ICPR 2008), pp.1-4, 2008.

[12] C. Weight, D. Parnham, and R. Waites, “LeafAnalyser: a computational method for rapid and large-scale analyses of leaf shape variation,” The Plant Journal, vol.53, pp.578-586, 2008.

[13] 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 neural network,” Proceedings of IEEE 7th Interational Symposium on Signal Processing and Information Technology, pp.11-16, 2007.