Jurnal
Teknologi

Full Paper

CLASSIFY THE PLANT SPECIES BASED ON LOBES, SINUSES AND MARGIN

Article history
Received
3 December 2013
Received in revised form
2 July 2014
Accepted
25 November 2014

Mei-Fern Bong[a], Ghazali Sulong[b], Rosely Kumoi[a], Mohd Shafry Mohd Rahim[b*]

[a]UTM ViCubeLab, Faculty of Computing, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[b]UTM IRDA Digital Media Centre, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia

*Corresponding author
shafry@utm.my

Graphical abstract

Abstract

This paper proposed a novel approach to cluster the species of plants based on their lobes, sinuses and margin. Firstly, all the boundary points in the clockwise or anticlockwise direction were selected. Then, an estimated centre point for leaf boundary points was used to compute the distance between the leaf boundary points and centre point. Next, the peaks and valleys from the distance found stated above were located where peaks represent lobes and valleys represent the sinuses. The number of peak and valleys is calculated to cluster the plant according to the rule-based method. From the results obtained, the accuracy for the plant clustering is up to 100 percent.

Keywords: Leaf classification, centroid contour distance, disk filter, find the peak, find the valleys, lobes, sinuses, teeth, margin

1.0 INTRODUCTION

Malaysia has more plant species compared to Europe and known to have mega-diversity of plant species in the world. The rapid changes to the natural ecosystem in Malaysia cause high risk of plant extinction compared to any other countries in the world. A rough estimation has been made by Onn [1] that there are 8,500 species of vascular plant found in Peninsular Malaysia and approximate 15,000 plant species found in the East Malaysia. However, Malaysia had been addressed to these problems as early as 1975 included this issue in the Third Malaysia Plan (1976-1980) [2]. Robust leaf recognition frameworks are crucial to help Malaysia in conserve and preserve the endangered plants.

There are many significant features which act as the key for identifying the leaf that had been investigated by other researchers. A large percentage of researchers distinguish the plant species using the leaf shape information [3-20]. However, there are many other researchers distinguish the plant species according to the information that extracted from leaf texture [18-21], color [16-21], leaf skeleton [22], leaf vein [16, 18, 20, 23] and leaf edge [24]. Researchers rarely cluster those important information by dividing them into groups for further investigation.

The most famous methods used to get leaf shape information, which are also applied in leaf recognition, are geometric features [4-6, 12-15, 17-19], Centroid Contour Distance [13, 25 28], Moment Invariant [12, 15, 26] and Polar Fourier transform [15, 16, 18].

Many investigators involved in leaf recognition field, but rarely seen them divide the leaf into groups for further identification. The importance of grouping the leaves together can save the computational time and less complexity. In this framework, we divided the leaves in groups according to the numbers of lobes, sinuses and the existence of tooth.

2.0 LEAF IMAGE PRE-PROCESSING

The approach we applied only required the information of leaf boundary, so the RGB leaf images were first converted into binary images to avoid from extensive computational. When the leaf was converted into binary, the scars of the leaf which had others intensity may cause some hole in the binary image. The existence of the hole in binary images may cause failure in tracing the correct leaf boundary point. Therefore, we need to fill up the hole in the binary image first to avoid any others unnecessary failure.

The imperfect margin of the leaves may cause confusion during the process of leaf lobes and sinuses extraction. Therefore, to avoid the disturbance of imperfect leaf margin, circular averaging filter or disk filter is applied to smooth out the margin. In this investigation, the radius with value 20 is used, the higher the radius is applied, the smoother the leaf margin. An example of filter mask or kernel with radius 2 is display in Table 1.

Table 1 Filter mask with radius 2

0	0.017	0.038	0.017	0
0.017	0.078	0.080	0.078	0.017
0.038	0.080	0.080	0.080	0.038
0.017	0.078	0.080	0.078	0.017
0	0.017	0.038	0.017	0

Canny edge detection with threshold low = 0:2 and high = 0:7 is applied to trace the edge of the leaves. Get the position x and y which the intensity is 1 then rearrange the leaf edge position in the clockwise direction which is 1- Dimension.

3.0 COMPUTE EACH DISTANCE OF LEAF BOUNDARY POINT TO CENTROID POINT

In this research, we assume centroid point (CentroidX;CentroidY) as the mid-point between maximum and minimum point in both x and y axes among the leaf edge point, displayed in Equation 1 and Equation 2 in next page,

$$Centroid_X = \left(\frac{maxX - minX}{2} + minX\right) \quad (1)$$

$$Centroid_Y = \left(\frac{maxY - minY}{2} + minY\right) \quad (2)$$

Then, we compute the Euclidean distance of each leaf boundary point to centroid point as shown in Equation 3.

$$dist_i = \sqrt{(CurX_i - Centroid_X)^2 + (CurY_i - Centroid_Y)^2} \quad (3)$$

Where CurX represents X value from leaf boundary and CurY represents Y value from leaf boundary. While, i represent the current number of boundary point (i = 1; 2; 3; :::;N) , N is total number of boundary point, size (CurX). This method is slightly different compared to regular Centroid Contour Distance (CCD) as regular CCD only computes the distance of leaf boundary point to centroid point in regular angle, which may lose the important information. However, in this method the distance between each leaf boundary point and centroid point were computed to obtain the position and exact value of local maxima and local minima. The local maxima represents the peak of the lobes and the local minima was the sinuses point of the leaves. Figure 1 shows the plot of distance between leaf boundaries versus the count of curve.

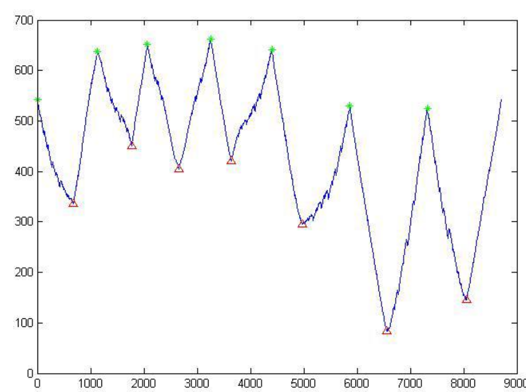

Figure 1 This graph shows the distance between leaf boundary point and centroid point

4.0 FIND THE PEAKS AND VALLEYS FROM THE GRAPH OF THE DISTANCE

After we got every distance of leaf boundary point to centroid point, a graph is plotted in Figure 1. From Figure 1, the total lobes and sinuses were counted. The peaks or the local maxima shown in Figure 1 represent the lobes or the pointed projecting parts in leaf boundary. Inversely, the valleys or the local minima shown in Figure 1 represents the sinuses or a notch or depression between two lobes. The total lobes and sinuses can help us to group them for further distinguished so it may make leaves recognition less computational extensive but fast and efficient when it comes to distinguish millions of plant species compared to previous approach. Those green and red points in the Figure 2 represent the peak of the lobes and the sinuses respectively. Figure 2 shows the plotted graph of distance versus the number of curves with the marked peaks and valleys. Figure 3 depicts the location of peak or lobes and valleys or sinuses in the leaf boundary image. The algorithm to find the peaks and valleys is shown below:

1) Assume first distance point is maxima and also the minima.
2) If the current distance point greater than maxima, replace the current point to maxima. Keep the index of current point.
3) If the current point less than minima, replace the current point to minima. Keep the index of current point.
4) If the current point is less than the distance of maxima a defined threshold, store the maxima point as the peak. Then, replaces the minima with current point.
5) If the current point greater than the total of minima and a defined threshold, store the minima point as valley. Then, replaces the maxima with current point.

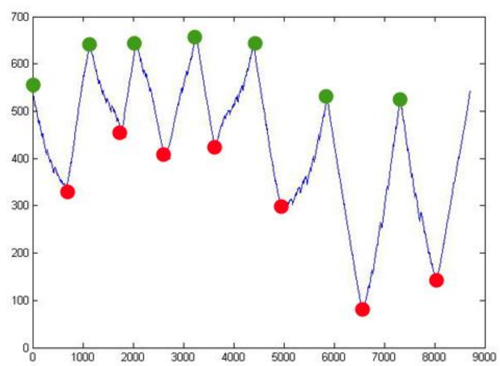

Figure 2 The peak and sinuses are marked with green and red circle respectively in the graph of distance

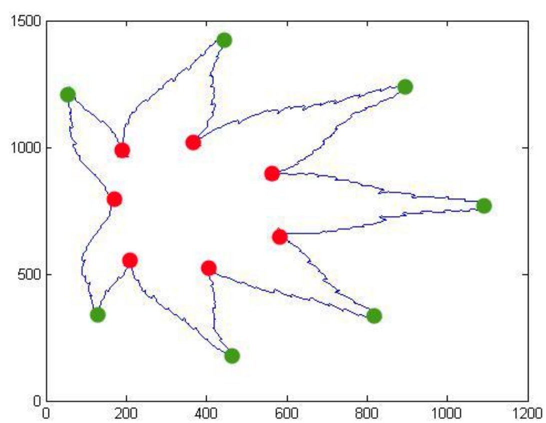

Figure 3 The peak and sinuses are marked with green and red circle respectively in the leaf boundary image

5.0 EXPERIMENTAL RESULT

Experimentation is performed by using 8 kinds of plants with various shapes which were provided by Wu [5] and can be downloaded from the website http://flavia.sourceforge.net/. Figure 4 shows the sample dataset that were used in this investigation. Based on the approach applied, we able to group those plants for further investigate. Table 2 shows the analysis of counting the number of peaks and valleys or the numbers of lobes, numbers of sinuses and the margin type which divided
into group. Figure 5 shows the findings of lobes and sinuses for each leaves.

Table 2 Result of cluster the plant species based on the number of peaks and valleys and the margin type

Group	Plant scientific name
Peak =2, valleys =2, margin = tooth	Koelreuteria paniculata Laxm.
Peak =2, valleys =2, margin = entire	Pittosporum tobira (Thunb.) Ait. f.
Peak =3, valleys =3, margin = tooth	Populus canadensis Moench
Peak =3, valleys =3, margin = entire	Cercis chinensis and Acer buergerianum Miq.
Peak =4, valleys =4, margin = entire	Liriodendron chinense (Hemsl.) Sarg.
Peak =5-7, valleys = 5-7, margin = tooth	Acer Palmatum and Kalopanax septemlobus (Thunb. ex A.Murr.) Koidz.

Figure 4 Samples of leaves

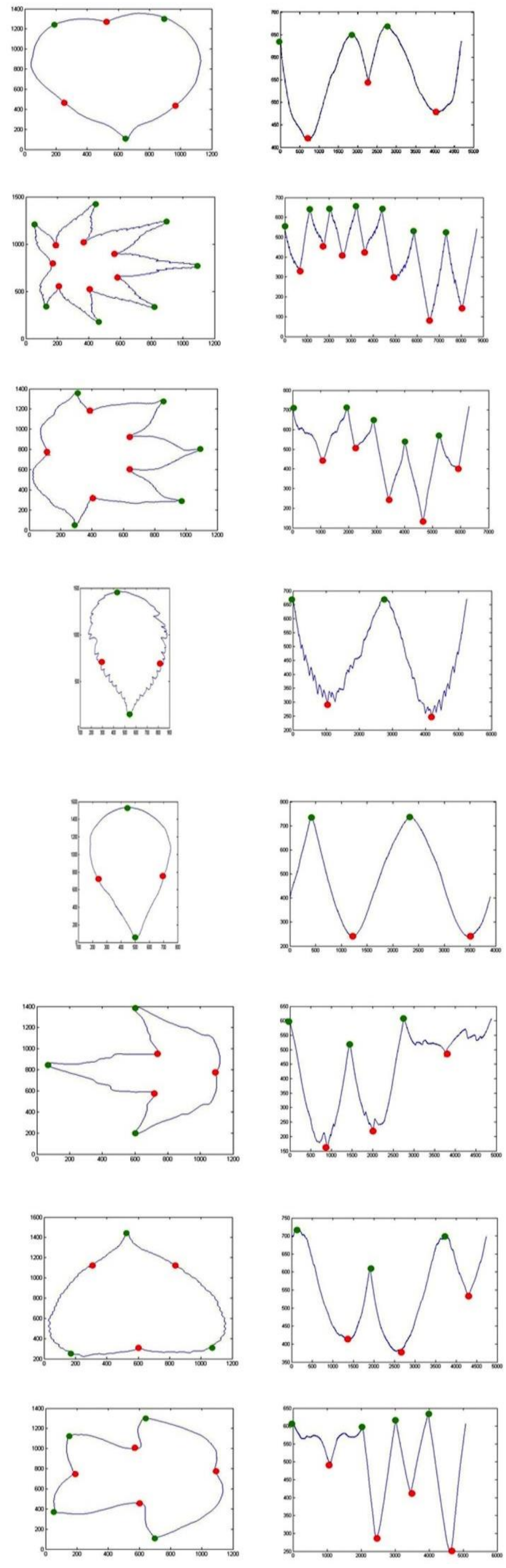

Figure 5 The location of peaks and valleys in the leaf samples

6.0 CONCLUSION

This research has achieved its objective by classifying the plant species based on the number of lobes, sinuses and the type of margin. This investigation can be used as the first hierarchy to differentiate the plant species based on the leaf shape according to their lobes, sinuses and the type of margin. Therefore, further investigation with this method can be done with less computational expensive, less complexity, to avoid over training when using classifier, more efficient and less computational time. This approach is more robust when it comes to recognizing huge number of plant species.

Acknowledgement

This research was funded by the Ministry of Education through Fundamental Research Grant Scheme (FRGS) and Research University Grant (RUG) under Universiti Teknologi Malaysia, and manage by Universiti Teknologi Malaysia under Vot No. J130000.7828.4F299 and Vot Q.J130000.2428.02G28.

References

[1] Oon, H. 2008. *Globetrotter Wildlife Guide Malaysia*. New Holland Publishers. 11. ISBN 1-84537-971-3.

[2] Ministry of Natural Resources and Environment. 2006. Biodiversity in Malaysia Putrajaya: Ministry of Natural Resources and Environment.

[3] Chaki, J. and R. Parekh. 2012. Designing an Automated System for Plant Recognition. *International Journal of Advances in Engineering and Technology (IJAET)*.

[4] Shayan, H. and G. Sajeevan. 2013. Plant Recognition from Leaf Image through Artificial Neural Network. *International Journal of Computer Applications*.

[5] Wu, S. G., F. S. Bao, E. Y. Xu, Y. X. Wang, Y. F. Chang, and Q. L. Xiang. 2007. A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network. *IEEE 7th International Symposium on Signal Processing and Information Technology*. Cairo. 2007.

[6] Du, J. X., X. F Wang, and G. J. Zhang. 2007. Leaf Shape Based Plant Species Recognition. *Applied Mathematics Computation*. 185: 883-893.

[7] Chia, L. L., and S. Y. Chen. 2003. Classification for Leaf Images. *IPPR Conference on Computer Vision, Graphics and Image Processing*.

[8] Du, J. X., D. S. Huang, X. F. Wang and X. Gu. 2006. Computer-aided Plant Species Identification (CAPSI) Based on Leaf Shape Matching Technique. *Transactions of the Institute of Measurement and Control*. 28(3): 275-284.

[9] Neto, C., G.E. Meyer, D. D. Jones and Ashok K. Samal. 2006. Plant Species Identification Using Elliptic Fourier Leaf Shape Analysis. *Computers and Electronics in Agriculture*. 50(121134).

[10] Nam, Y., E. Hwang, and K. Byeon. 2005. *ELIS: An Efficient Leaf Image Retrieval System*. Lecture Notes in Computer Science. 3687/2005: 589-597.

[11] Du, J., D. Huang, X. Wang, and X. Gu. 2005. Shape Recognition Based on Radial Basis Probabilistic Neural Network and Application to Plant Species Identication. *Proceedings of 2005 International Symposium of Neural Networks, Ser. LNCS 3497*. Springer. 2005.

[12] Wang, J. X. Du, and G. J. Zhang. 2005. Recognition of Leaf Images Based on Shape Features Using a Hypersphere Classifier. *In Proceedings of International Conference on Intelligent Computing, Ser. LNCS*.

[13] Ye, Y., C. Chen, C. T. Li, H. Fu, and Z. Chi. 2004. A Computerized Plant Species Recognition System. *In Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing*. Hong Kong. October 2004.

[14] Pornpanomchai, C., C. Kuakiatngam, P. Supapattranon and N. Siriwisesokul. 2011. Leaf And Flower Recognition System (e-botanist). *IACSIT International Journal of Engineering and Technology*. August 2011. 3(4).

[15] Kadir, A., L. E. Nugroho, A. Susanto and P. I. Santosa. 2011. A Comparative Experiment of Several Shape Methods in Recognizing Plants. *International Journal of Computer Science and Information Technology*. June 2011.

[16] Kadir, A., L. E. Nugroho, A. Susanto and P. I. Santosa. 2011. Foliage Plant Retrieval Using Polar Fourier Transform, Color Moments and Vein Features. *Signal and Image Processing.* September 2011.

[17] Pornpanomchai, C., S. Rimdusit, P. Tanasap and C. Chaiyod. 2011. Thai Herb Leaf Image Recognition System (THLIRS). *Kasetsart J. (Nat. Sci.)*.

[18] Kadir, A., L. E. Nugroho, A. Susanto, and P. I. Santosa. 2011. Neural Network Application on Foliage Plant Identification. *International Journal of Computer Applications*.

[19] Kadir, A., L. E. Nugroho, A. Susanto, and P. I. Santosa. 2011. Leaf Classification Using Shape, Color, and Texture Features. *International Journal of Computer Trends and Technology*. July to Aug Issue 2011, August 2011.

[20] Bama, B. S., S. M. Valli, S. Raju and V. A. Kumar. 2011. Content Based Leaf Image Retrieval (CBLIR) Using Shape, Color and Texture Features. *Indian Journal of Computer Science and Engineering (IJCSE)*.

[21] Zhang, X. and F. Zhang. 2008. Images Features Extraction of Tobacco Leaves. *Congress on Image and Signal Processing*.

[22] Gu, X., J. X. Du, and X. F. Wang. 2005. Leaf Recognition Based on the Combination of Wavelet Transform and Gaussian Interpolation. *In Proceeding of International Conference on Intelligent Computer*. Hefei, China. 23-26 August 2005. 253-262.

[23] Singh, K., I. Gupta, and S. Gupta. 2010. SVM-BDT PNN and Fourier Moment Technique for Classification of Leaf Shape. *International Journal of Signal Processing, Image Processing and Pattern Recognition*. 3(4): 67-78.

[24] Qi, H., and J. G. Yang. 2003. Sawtooth Feature Extraction of Leaf Edge Based on Support Vector Machine. *In Proceedings of the Second International Conference on Machine Learning and Cybernetics*. November 2003.

[25] Fu, H., Z. Chi, D. Feng, and J. Song. 2004. Machine Learning Techniques for Ontology-based Leaf Classification. *In IEEE 2004 8th International Conference on Control, Automation, Robotics and Vision*. Kunming, China. 2004.

[26] Chaki, J. and R. Parekh. 2011. Plant Leaf Recognition Using Shaped Based Features and Neural Network Classifiers. *International Journal of Advanced Computer Science and Applications (IJACSA)*. 2011.

[27] Knight, D., J. Painter, M. Potter. 2010. *Automatic Plant Leaf Classification for a Mobile Field Guide*. Rapport rechnique, Universite de Stanford.

[28] Beghin, T., J. S. Cope, P. Remagnino, and S. Barman. 2012. *Shape and Texture Based Plant Leaf Classification*. Digital Imaging Research Centre.