An Implementation of Leaf Recognition System

Kue-Bum Lee[1], Kwang-Woo Chung[2] and Kwang-Seok Hong[1],

[1] College of Information and Communication Engineering, Sungkyunkwan University, 300, Chunchun-dong, Jangan-gu, Suwon, Kyungki-do, 440-746, South Korea
[2] Department of Railway Operation System Engineering, Korea National University of Transportation, 157, Cheoldobangmulkwan-ro, Uiwang-si, Kyungki-do, 437-763, South Korea

leo0608@skku.edu, ckw1201@ut.ac.kr, kshong@skku.ac.kr

Abstract. In this paper, we propose and implement a leaf recognition system using the leaf veins and shape for plant classification. The proposed approach uses the main vein, and frequency domain data by using Fast Fourier Transform (hereinafter, FFT) methods in conjunction with distance measurements between the contours and centroids on detected leaf images. To verify the validity of the approach, images of 1907 leaves were applied to classify 32 kinds of plants. In the experimental results, the proposed leaf recognition system showed an average recognition rate of 97.19%, which was better than that of the existing leaf recognition method.

Keywords: Leaf recognition, Leaf feature extraction, Leaf vein feature, Plant classification, Convex hull

1 Introduction

Leaf recognition technology plays an important role in plant classification, and its key issue lies in whether selected features are stable and have good ability to discriminate different kinds of leaves. Many recent studies exist on plant classification and recognition based on plant components, such as flowers, leaves, and bark. To handle such volumes of information, realization of a quick and efficient classification method has become an area of active study [1-6]. In particular, it is well known that the correct way to extract plant features involves plant recognition based on leaf images. Two features widely used to such ends are color and shape [1, 6]. In conventional color-based study, a simple color similarity between two images can be measured by comparing their color histogram. In conventional shape-based study, region and contour-based simple features are used, and can be considered in the time domain data.

We propose a leaf recognition system for plant classification based on the leaf veins and shape. One of the notable points is that the leaf direction can be decided upon using leaf veins, and frequency domain data extracted using FFT on the distance between contour and centroid in a given leaf image can be used.

2 Leaf direction decision using main vein

2.1 Leaf vein extraction

We convert the input leaf image to a gray scale image and then perform opening operations [1]. We obtain a difference image between the gray scale image and the image subjected to opening operation. Then, we obtain a leaf vein image by converting the difference image to a binary image.

2.2 Main vein extraction and leaf direction decision

We describe the method of main vein extraction and leaf direction decision from a leaf vein image, using a projection histogram in the horizontal and vertical directions, in order to measure the distribution of the leaf veins. In the first step, we extract main vein of a leaf using projections in the horizontal direction. We decide on the main vein using the maximum point of the histogram while the leaf vein image is rotated 180 degrees. After extracting the main vein, we decide on the direction of the leaf through the projections in the vertical direction.

3 Leaf Feature Extraction

The centroid of the detected leaf region was found as follows:

$$C(x,y) = C(\frac{1}{N}\sum_{n=1}^{N} x_n,\ \ \frac{1}{N}\sum_{n=1}^{N} y_n) \tag{3}$$

where *C(x, y)* is the centroid coordinate of the leaf region image, and *N* is the number of pixels on the detected leaf region.

The distance is calculated by measuring the centroid of the leaf region to all points on the leaf contour as follows:

$$D(i) = \sqrt{\left|C_x - E(i)_x\right|^2 + \left|C_y - E(i)_y\right|^2} \tag{4}$$

where *D(i)* is the distance between the centroid of the leaf region and the i^{th} leaf contour pixel. C_x and C_y are the coordinates of the centroid of the leaf region, and $E(i)_x$ and $E(i)_y$ are the coordinates of i^{th} leaf contour pixel. FFT is then performed using calculated distance values. The distance is acquired by measuring the longest distance point from the centroid in a clockwise direction. 10 features are then extracted based on the distance, FFT magnitude, and phase.

We describe the geometric and digital morphological features for leaf feature extraction. We extract four basic geometric features: the leaf length, leaf width, leaf area, leaf perimeter. The leaf length is defined as the longest distance between the centroid and the two ends on the margin of the leaf on opposite sides of the centroid, denoted by *LL*. The leaf width is defined as the distance between the intersection point with *LL* at the centroid and its opposite side on the margin of the leaf, denoted by *LW*. The leaf area is the number of pixels in the leaf region, denoted by *LA*. The leaf perimeter is the number of pixels in the leaf contour, denoted by *LP*.

We extract 10 features based on digital morphological features using four basic geometric features and the study conducted by Wu et al. [1], including the aspect ratio, form factor, rectangularity, perimeter ratio of the leaf length, perimeter ratio of the leaf length and leaf width, and five vein features.

In mathematics, the convex hull of a set *X* of points in the Euclidean plane or Euclidean space is the smallest convex set that contains *X*. For instance, when *X* is a bounded subset of the plane, the convex hull may be visualized as the shape formed by a rubber band stretched around *X* [7]. In this paper, we use the convex hull in order to reflect the complexity of the leaf contour. We compute the rate of extent of the original image and convex hull image in order to extract leaf feature.

4 Experiments and Results

In this paper, we used 1907 leaf images of 32 species collected by Wu et al. [1]. Each plant species has a minimum of 50 and a maximum of 77 sample leaves.

Table 1. Experimental results of proposed leaf recognition system.

Recognition system	Number of Leaf Image	Number of Incorrect recognition	Recognition rate
Existing system [1]	1800	31	90.31%
Existing system [2]	1907	15	95.31%
Proposed system	1907	9	97.19%

The proposed system was implemented using Microsoft Visual C++ 6.0 and the Intel OpenCV library. Because the leaf image size and position of the dataset are not constant, we normalized it to the leaf image.

To evaluate the performance of the proposed leaf recognition system, a recognition model was created using a range of values with 21 features for each plant species. For each kind of plant, 10 pieces of leaves from testing sets are used for testing. Table 2 show the 2 experimental results for the proposed leaf recognition system, and Fig. 3 shows an example of the leaf recognition system. The average recognition accuracy of the proposed system is 97.19%.

5 Conclusions

In this paper, a leaf recognition system based on leaf veins and shape has been proposed and implemented for plant classification. We extract main vein from the input image, and leaf direction is decided using projection histograms of an extracted main vein image. 21 leaf features of distance, FFT, and convex hull were extracted for the leaf recognition.

In the experimental results, the proposed leaf recognition system showed a performance of 97.19%. From the experimental results, we can confirm that the recognition rate of the proposed advanced leaf recognition system was better than that of the existing leaf recognition system.

In future work, we will improve the proposed system and further improve its recognition performance. In addition, we are continuing to perform research in order to find a correct leaf contour extraction method in the complex backgrounds.

6 Acknowledgements

This research was supported by MKE, Korea under ITRC NIPA-2012-(H0301-12-3001) and PRCP through NRF of Korea, funded by MEST (2012-0005861).

References

1. Wu, S. G. F., Bao, S., Xu, E. Y. Y., Wang, X. Y., Chang, F. and Xiang, Q. L.: A Leaf Recognition Algorithm for Plant Classification using Probabilistic Neural Network. The 7th IEEE International Symposium on Signal Processing and Information Technology, pp11--16, IEEE Press, Cairo (2007)
2. Lee, K. B., and Hong, K. S.: Advanced Leaf Recognition based on Leaf Contour and Centroid for Plant Classification. The 2012 International Conference on Information Science and Technology, pp. 133--135, SERSC Press, Shanghai (2012)
3. Kim, J. H., Huang, R. G., Jin, S. H. and Hong, K. S.: Mobile-based flower recognition system. 2009 Third International Symposium on Intelligent Information Technology Application , pp. 580--583, IEEE Press, NanChang (2009)
4. Hsu, T. H., Lee, C. H. and Chen, L. H.: An interactive flower image recognition system, Multimed. Tools Appl. 53:1, 53--73 (2011)
5. Kim, S. J., Kim, B. W. and Kim, D. P.: Tree Recognition for Landscape Using by Combination of Features of its Leaf, Flower and Bark. The SICE Annual Conference 2012, pp. 1147--1151, IEEE Press, Tokyo (2011)
6. Smeulders, A. W. M., Worring, M., Santini, S., Gupta, A. and Jain, R.: Content-based image re-trieval at the end of the early years. IEEE Trans. Pattern Anal. Mach. 22, 1349--1380 (2000)
7. Berg, M., Cheong, O., Kreveld, M., and Overmars, M., Computational Geometry: Algorithms and Applications. Springer-Verlag Press, pp.2-8 (2000)