Volume 5, Issue 7, July 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

A Method of Feature Extraction from Leaf Architecture

[1]V. R. Patil, [2]R. R. Manza
[1]Department of MCA, KKWIEER, Nashik, India
[2]Department of Computer Science and IT, Dr.Babasaheb Ambedkar Marathwada University, Aurangabad, Maharashtra, India

Abstract-- Plants are very important for human being life. All food people eat comes directly or indirectly from plants. To effectively use plants one must learn plant identification. So it is necessary to establish database by information technology as soon as possible. The leaves of the plant carry a lot of information about plant species. These features are extracted and used as a basis for plant identification. In the proposed work, basic geometrical and digital morphological features are extracted. These features are independent of leaf growth and image translation , rotation and scaling and are studied. The output shows the result for features of selected leaves which are further useful for plant identification and classification.

Keywords-- Digital morphological features, feature extraction, geometrical features, leaf, image preprocessing

I. INTRODUCTION

Plants play a very important role in the life of human being. They are useful for human as well as animals too. Plants provide food, medicine, oxygen and much more. It is necessary to know which plants are useful and which are harmful to human being. There are many challenges for researchers to try other approaches for better performance in order of feature extraction. A particular method may produce accurate result in specific samples of leaves, but does not guarantee to perform good result for other ones. In order to recognize plant species there have been many recent studies on plant identification. One way to recognize the plant is from their leaves because every plant has a unique leaf. It serve as a tool to plant biologists and botanists for distinguishing plant species. It is very difficult to analyze the shape of flower and seedling because of their 3D structure and it increases complexity. Whereas plant leaves are two dimensional in nature. Thus, they are most suitable. Classification of plants based on leaves is the fastest and simpler way to identify a plant. These are the most visible and variable aspect of a plant from trees. They can be easily found everywhere. So it is necessary to establish database by information technology as soon as possible. Sampling leaves and photoing them is expensive in cost. We can easily transfer the leaf image to a computer and then the computer can extract necessary features automatically using image processing techniques and subsequently can recognize the plant leaf using machine learning.

A. Organization of the Paper

This paper is organized as follows, In Section 2 gives and introduction of leaf image preprocessing. In section 3, digital morphometric feature extraction is discussed. The experimental setup for the same is in section 4. The concluding remarks are given in section 5.

II. LITERATURE SURVEY

Global feature and local descriptor are two categories for features of leaf, as stated by Shabanzade et al(2011)[1]. The shape,color and skeleton are basic features for plant classification according to Jing el al 2009 [12]. According to C. L. Lee , S. Y. Chen [2] past research in recognizing objects can be broadly classified into two categories : a) contour based and b) region based approaches. The disadvantage of the contour based feature is the difficulty on finding the correct curvature points. Based on the contour of leaf, features were extracted to differentiate species. However contour of leaves have variation even in the same species. For plants identification purpose Wu, et al[3] used shape slimness, defined as ratio of length to width of leaves, shapes roundness, defined as ratio of area of leaf image and perimeter of leaf contour, and shape solidity, defined as ratio of the internal area connecting to valley points and the external area connecting the top points. A paper by Ji- Xiang Du, Xiao-Feng Wang, Guo-Jun Zhang [4] introduce how to extract digital morphological features. The features are extracted from the contours of leaf. The digital morphological features(DMF) generally include geometrical features(GF) and invariable moment features(MF). A paper by Cholhang Im, Hirobumi Nishia and Tosiyasu L. Kunii[5], a method that normalizes shapes of leaves is presented using symmetry of each leaflet with respect to its vein. According to Najjar and Zagrouba [6] had used region based feature for proposed method in order to classifying the leaf. According to C.S Sumathi and A.V.Senthil Kumar in plant classification, the leaf shape plays a significant role. In machine intelligence, the most significant part essential for both decision making and data processing is shape recognition. Valliammal and Geethalakshmi [7], who stated in their publications that leaf image could be categorized based on color, texture, shape or combination of these properties. Later Zhang and Zhang[8] was

enhanced that the properties for these features such as surface area, surface perimeter and the disfigurement are inherited from the shape features, variance of red, green and blue channels are belonging to the color features and texture energy. Other research was also used aspect ratio and other basic geometric features to recognize plants. Du, Wang and Zhang used Aspect ratio, rectangularity, area ratio of convex hull, perimeter ratio of convex hull and sphericity [3].

[9] this paper proposes a method for an automatic identification of plant species from low quality pictures of their leaves created using mobile devices. To avoid segmentation of the images local features are used which are scale and rotation invariant.[10] this paper proposed a method that combines polar fourier transform, color moments and vein features to retrieve leaf images based on leaf image. Du et al [11] captured the leaf shape polygonial approximation and algorithm called MDP(modified dynamic programming) for shape matching.

III. LEAF IMAGE PREPROCESSING

Plant leaf preprocessing is important for digital morphometric feature extraction of plant leaf. In this section we introduce preprocessing procedure on the plant leaf image. There are various parameters of a plant taxonomy which are used by botanist in plant morphological research. Some of these features are color of plant leaf, flower, seed and shape of leaf . Normally the plant leaves are green in color. But there are various shades of color for a single plant. Moreover , the variety of shades for a single plant appears because changes in water volume, nutritional value, atmospheric changes in environment. Based on this we recognize that color feature has low reliability to move forward. Similarly it is very difficult to analyze the shape of flowers and seedling because of their 3 dimensional structure, if we consider only two dimensional images. Thus we decided to recognize plant by their shape. For this we have considered the dataset provided by Intelligent Computing Laboratory, Institute of Intelligent Machine, Chinese Academy of sciences. Firstly we have converted the 2 dimensional image of plant leaf to grey level image . for each conversion of a color image to grayscale is not unique, different weighting of the color channels effectively represent the effect of shooting black-and-white film with different strategy is to match the luminance of the grayscale image to the luminance of the color image. To convert plant leaf color image to a grayscale representation of its luminance, Firstly we have obtained the values of its red, green and blue(RGB)primaries. Finally only gray component for each is computed from the color image by

Gray = 0.2999 * R + 0.578 *G+ 0.114 * B

Where R,G,B correspond to the color of the pixel , respectively. This image is then transformed into a binary image. Further , smoothing filter is applied to the binary image to reduce the noise . The boundaries are gained by applying a laplacian filter. The steps involved in pre-processing technique are illustrated in the following figure

Figure1 : Sample images in leaf database

Figure 2 : Different samples in the same class

Original Image(RGB)→Gray→Binary→Smoothing-→Contours of Leaf

Figure 3: Steps in image pre-processing

IV. FEATURE EXTRACTION

The next step in plant leaf recognition is the digital morphological feature extraction . The aim behind this step is to remove redundancy from the image and to represent by a set of numerical features. These features are extracted from the contours of leaf. The digital morphological features generally include geometrical features and invariable moment features. The digital morphological features contain basic geometric features and morphological features. Our system

extract 5 basic geometric features i.e longest diameter, physiological width, leaf area and leaf perimeter. Our system also extract 12 digital morphological feature are as follows smooth factor, aspect ratio, form factor, rectangularity, narrow factor, perimeter ratio of diameter, perimeter ratio of physiological length and physiological width and 5 vein features : f1,f2,…..f5.

A. Five basic Geometric Features

1. Longest Diameter : it is longest distance between any two points on the contours of leaf. It is denoted as D.
2. Physiological length : Here one needs to select two terminal of the main vein of leaf via mouse click. The distance between these two terminal is defined as the physiological length. It is denoted as L_P
3. Physiological Width : Once physiological length is find out i.e line drawn between two selected terminal on the main vein, one can plot infinite lines orthogonal to that line. We consider two lines are orthogonal if their degree is $90^{0.}$ Physiological width is consider as longest distance between points of those intersection pairs. It is denoted as W_P
4. Leaf Area : Smoothed leaf image is consider to find out leaf area. Number of pixels having binary value 1 is termed as leaf area. It is denoted as A
5. Leaf Perimeter : Leaf Perimeter is calculated by counting the number of pixels consisting leaf margin. It is denoted as P

B. Digital Morphological Features

Based on above mentioned five basic geometric features, we can define 12 digital morphological features

1. Smooth factor : It is defined as the ratio between area of leaf image smoothed by 5 X 5 rectangular averaging filter and area of leaf image smoothed by 2 X 2 rectangular averaging filter.
2. Rectangularity : It is defined as the ratio between physiological length, physiological width and leaf area. Thus, $L_P W_P/A$.
3. Aspect ratio : It is defined as the ratio between physiological length and physiological width. Thus,L_P/ W_P.
4. Perimeter ratio of diameter : It is defined as the ratio between perimeter and diameter. Thus, P/D.
5. Form factor : It is defined as $4 Д A/P^{\cdot}$ where A is area of leaf and P is perimeter of the leaf margin.
6. Narrow factor : It is defined as the ratio between diameter and physiological length. Thus, D / L_P
7. Perimeter ratio of physiological length and physiological width : It is defined as the ratio between perimeter and sum of physiological length and physiological width. Thus, $P / (L_P + W_P)$
8. Vein features : we perform morphological opening on grayscale image with flat, disk shape structuring element of radius 1,2,3,4 and subtract the remain image by the margin. Then gray threshold value is computed and with that threshold image is converted to binary. 8/4 connected components of leaf are find out and measured set of properties of each connected component. The result is look like the vein. Areas of left pixel are denoted as A_{V1}, $A_{V2,}$ $A_{V3,}$ A_{V4} respectively. Then obtain the last five features as
   A_{V1}/A, A_{V2}/A, A_{V3}/A, A_{V4}/A, A_{V1}/ A_{V1}.

V. EXPERIMENTAL RESULT

We use Flavia dataset . This database contains 32 kind of different plant species. Each species includes 20 sample images. Hence there are 400 images with the database. The representative sample images for the different plant species in the database are shown in Fig. 1 and 2. The table 1 below shows the results obtained after extraction of 5 basic geometrical features:

Table 1: Results after extracting 5 basic geometrical features

Leaf/Feature	Diameter	Physiological Length	Physiological Width	Leaf Area	Leaf Perimeter
	877.4058	766.5445	523.2447	7.0955e+004	1.7424e+003
	791.0209	529.3611	534.0206	2.2625e+005	1757
	751.8045	423.8629	531.7377	1.1498e+005	4523e+003
	842.0665	806.2175	385.5412	204990	1.9344e+003

	746.8829	464.9033	690.1884	2.1268e+005	1.9494e+003
	838.9857	776.4044	501.1846	290166	2.0861e+003

Table 2 : Results after extracting 12 digital morphological features

Leaf\Feature	Smooth Factor	Aspect Ratio	Form Factor	Rectangularity	Narrow Factor	Perimeter ratio of Diameter	Perimeter ratio of $L_{P\ and}W_P$	Vein Feature				
								F1	F2	F3	F4	F5
	0.9915	1.4650	0.2937	5.6527	1.6769	1.9858	1.3509	0.1585	0.3109	0.3359	0.3919	2.4722
	0.9977	0.9913	0.9210	1.2494	1.4813	2.2212	1.6523	0.1674	0.2703	0.3109	0.3667	2.1902
	0.9947	0.7971	0.1212	0.9601	1.4139	4.5920	3.6126	0.0595	0.0137	0.0147	0.0160	0.2683
	0.997	2.0911	0.6884	1.5163	2.1841	2.2972	1.6231	0.1433	0.2171	0.2656	0.2766	1.9302
	0.9971	0.6736	0.7033	1.5087	1.0821	2.6100	1.6876	0.2520	0.3756	0.4166	0.4870	1.9323
	0.9971	0.6736	0.7033	1.5087	1.0821	2.6100	1.6876	0.2520	0.3756	0.4166	0.4870	1.9323

VI. CONCLUSION

In this paper an approach for automatically extracting features of leaf is presented. First the images are transformed into grayscale and normalized. Then converting grayscale image to binary and binary to smoothing. Using Laplacian filter boundaries are enhanced from binary smooth image. Further 5 basic geometrical features and 12 digital morphological features are extracted from the contours of leaf. These features are further useful for plant identification.

ACKNOWLEGEMENT

We would like to thank the "Intelligent Computing Lab., China" for their support and for providing the images.

REFERENCES

[1] Shabanzade, M.M.Zahedi and S.A.A ghvami, 2011. Combination of local descriptors and global features for leaf recognition signal and image processing. International Journal, vol. 2 : 23-21. DOI : 10 .5121/SIPIJ

[2] C.L.L EE, S.Y. Chen,(2003). "Classification for Leaf Image" , 16th IPPR Conference on Computer Vision , downloaded at www.csie.mcu.edu.tw/`yklee/CVGIP03/CD/Paper/PR/PR-19.pdf.

[3] Q.W.Wu, C.Zhou,C.Wang,(2006) " Feature Extraction and Automatic Recognition of Plant Leaf using Artificial Neural Network ", Advances in Ciencias de la Computation , pp. 5-12.

[4] Ji-Xiang Du, Xiao-Feng Wang,Guo-Jun Zhang. "Leaf Shape based plant species recognition", Applied Mathematics abd Computation 185 (2007) 883-893.

[5] Cholhong Im, Hirobumi Nishida, Tosiyasu L. Kunii . " Recognizing Plant Species by Normalized Leaf Shapes". Vision Interface '99, Trois-Rivieres,Canada, 19-21 May.

[6] Najjar, A. and E.Zagrouba, 2012. Flower image segmentation based on color analysis and a supervised evaluation. Proceddings of the International Conference on Communications and Information Technology (ICCIT) , Jun. 26-28, IEEE Xplore Digital Library, pp : 397-401. DOI : 10.1109/ICCITechnol.2012.6285834

[7] N. Valliammal and S.N. Geethalakshmi, (2011). Automatic Recognitin System Using Preferential Image Segmentation for Leaf and Flower Images, Computer Science and ENGINEERING : An International Journal (CSEIJ),Vol. , No. 4, pp.13-25.

[8] Stefan Fiel and Robert Sablantnig ."Leaf classification using local features."

[9] Abdul Kadir,Lukito Edi Nugroho, Adhi Susanto and Paulus Insap Santosa , " Foliage plant retrieval using polar fourier transform, color moments and vein features", An international Journal(SIPIJ) Vol. 2, No.3, September 2011.

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

[11] Jing, L.Z.Shawan and L.Jiandu , 2009. "A method of Plant Leaf Recognition based on Locally Linear Embedding and Moving Center Hypersphere Classifier". Procedding of Intelligent Computing Technology and Application Springer-verlag (ASV'09), Uslan, South Korea,PP : 645-651. DOI : 10.1007/978-3-642-04020-769

[12] R.C.Gonzalez, R.E.Woods and S.L.Eddins, Digital Image Processing Using MATLAB. Prentice Hall , 2004.

[13] S.G.Wu,F.S.Bao,E.Y.Xu,Yu-Xuan Wang,Yi-Fan Chang and Qiao-Liang Xiang. "A leaf recognition algorithm for plant classification using probabilistic neural network. In proc.IEEE International Symposium on Signal Processing and Information Technology , Pages 11-16,2007

