ISSN 2047-3338

Edge and Texture Fusion for Plant Leaf Classification

C. S. Sumathi and A. V. Senthil Kumar

Abstract—**Automatic digital plant classification and retrieval can be achieved by extracting features from its leaves. There are ample opportunities to improve plant species identification due to computerization through the designing of a convenient automatic plant recognition system. Many different approaches are used to classify the species of plants based on the plants features. In this research, the plants are classified on the basis of shape and texture. In this paper, it is proposed to extract edge and texture features using Gabor filter and fuse them for image classification. Radial Basis Function is used to measure the classification accuracy of the proposed method.**

***Index Terms*— Plant Leaf Classification, Sobel Edge Detector, Gabor Filter, Texture Analysis and Radial Basis Function**

I. INTRODUCTION

PLANTS are important sources for human living and development be it industry, food or medicine. It is also important for environmental protection. The World Wide Fund for Nature (WWF), states that there are currently between 50,000 - 70,000 known species globally. But many plant species are yet unknown and with the environment deteriorating these unknown species might be on the verge of extinction. Hence, it is a must to correctly and quickly identify plant species so that genetic resources are preserved. The plant classification distinguishes different kinds of plants and also builds system to classify the plant. Many different approaches are used to classify the species of plants based on the plants features. In this research, the plants are classified on the basis of shape and texture.

In plant species identification, individual plants are assigned to a species (group of related plants) based on its characteristics [1]. Presently plant taxonomy methods still follow traditional classification systems like morphologic anatomy, cell biology and molecular biological approaches. This work, usually done by botanists, consumes time and is not very efficient in addition to being very troublesome and time consuming. But there are ample opportunities to improve plant species identification due to computerization through the designing of a convenient automatic plant recognition system by identifying plant species based on leaf features. Earlier approaches used k-nearest neighbor (*k*-NN) classifier and adopted Artificial Neural Network (ANN) [2]. But they have disadvantages, in that some are appropriate to only certain species and some methods compare feature similarity when what is needed is the human element to enter queries manually [3].

According to the plant taxonomy theory, plants are essentially classified based on the shapes of leaves and flowers. Leaves are important part of plant and are distinct in shape and texture. A leaf is generally made up of blade, petiole and stipules. The blade is the flat photosynthetic part of leaf, petiole is the stem of leaf and the stipules are leaf-like formation found at the base of the petiole. Leaves are usually two-dimensional while flowers are three-dimensional. Generally, only the blade of the leaf is considered during automatic digital plant classification and retrieval. The blade has a distinct shape and texture. The leaf is used for classification as it is difficult to analyze flower shapes and structures as they have complicated 3D structures [4]. Also while leaves can be easily collected in all season's flowers can be obtained only when they are in bloom. It is for this reason that leaves are used widely for classification through computer aids.

Shape is an important image feature perceived by the human to characterize an object. According to the plant taxonomy, leaf shape is important and effective for classifying the plants. Many morphological differences exist in different kinds of leaves. Variations are seen in the same kind of leaves due to difference in location, direction of rotation of the leaves. Thus, to efficiently classify a leaf, the system should be able extract the leaf shape regardless of location, orientation of the leaf image. Additionally, the texture of the leaf is extracted and combined with the shape feature to improve the efficiency of classification of the plant.

This paper is organized into the following sections. Section II gives a brief literature review on techniques which others have adopted. Section III describes the proposed methodology followed by section IV which presents the findings with discussion. Section V concludes this paper

II. LITERATURE REVIEW

Yanhua Ye, et al., [5] proposed a computerized plant species recognition system (CPSRS). CPSRS is a Web-based application that ensures an efficient way to search and identify

C. S. Sumathi is an Assistant Professor with O/o Dean (SPGS), Tamil Nadu Agricultural University, Coimbatore- 641 003, Tamil Nadu, India (Email: css1@tnau.ac.in).

A.V. Senthil Kumar, Director with Department of Post Graduate and Research in Computer Applications, Hindusthan College of Arts and Science, Coimbatore- 641 028, Tamil Nadu, India.

field plant species. It is built on a Java Web infrastructure to support platform-independent application. Java applets and servlets are adopted to balance the burden of computing in the client and server. The CPSRS architecture introduced and shows its design and working. Two types of plant species retrieval methods, text-based information retrieval and content-based leaf retrieval are discussed. With the former, exact plant species information is retrieved from the database according to input searching criteria provided. For content-based leaf retrieval, experiments reveal a recall rate of about 71.4% when top five returned images are considered.

Fu, et al., [6] proposed an ontology-based leaf classification system, where machine learning techniques are used for automatic classification. Leaf classification, indexing and retrieval are an integral part of a computerized plant identification system. The leaf contour classification used a scaled CCD code to categorize the shape and margin type of a leaf by using a principle usually adopted by botanists. Then a trained neural network is employed to recognize detailed tooth patterns. An unlobed leaf is also measured according to methods employed by botanists. For leaf vein recognition, the vein texture is extracted through a combined thresholding and neural network approach so that more leaf details are made available. Compared to past studies, the proposed method integrates low-level image features and specific domain knowledge (ontology) of botany, and thus ensures a user friendly system. Primary experiments showed promising results, proving the system's feasibility.

Wang, et al., [7] proposed a classification framework for leaf images with complicated backgrounds. Classification of plant leaves had always been difficult as their complicated backgrounds included interferents and overlapping phenomena. An automatic marker-controlled watershed segmentation method along with pre-segmentation and morphological operations was introduced to segment leaf images which had complicated backgrounds. Based on earlier shape information, seven Hu geometric moments and sixteen Zernike moments were extracted as shape features from the segmented binary images after leafstalk removal. In addition, a moving center hyper-sphere (MCH) classifier capable of compressing feature data was designed to handle mass high-dimensional shape features. In the end the study results on practical plant leaves revealed that the proposed classification framework worked well in classifying leaf images having complicated backgrounds. A total of 20 practical plant leaves were successfully classified with correct classification averaging 92.6%.

Ji-Xiang Du, et al., [8] proposed computer-aided plant species identification (CAPSI) approach using a shape matching technique. A modified dynamic programming (MDP) algorithm for shape matching was used for plant leaf recognition and the study revealed that it trotted up a high accuracy rate of over 92 % as it was better than others in the recognition of blurred leaf images. When it came to classification of partial leaves, the MDP algorithm used proved to be more efficient in comparison to other methods. Similarly for retrieval of leaves this procedure was far better than those of competitors. The superiority of the proposed method over traditional approaches to plant species identification was demonstrated by experiments which showed that the algorithm used was suitable for the recognition of not only intact but also partial, distorted and overlapped plant leaves due to its robustness.

Rashad, et al., [9] introduced an approach to plant classification based on characterization of texture properties. A combined classifier learning vector quantization was utilized. The advantage of the proposed system was its ability to classify and recognize a plant from a small part of the leaf without needing to depend either on the shape of the full leaf or its color features. This was due to the fact that the system essentially depended on the part's textural features. Hence, the system can be used by researchers of Botany who need to identify damaged plants, as it can now be done from a small available part. Similarly when researchers need to classify a plant and have only a portion available that is in itself enough to do the needful as the proposed system textural features are required and not color features which are liable to change with the changing season. The results revealed that the combined classifier method produced high performance far superior to other tested methods as its correct recognition rate was 98.7% and hence applicable.

In this paper, it is proposed to extract features using the Sobel edge detection and Gabor filters. The features are fused to create the dataset. The proposed feature extraction method is tested against Classification and Regression Tree (CART) classifier and Radial Basis Function (RBF) classifier.

III. METHODOLOGY

The process of detecting the pixels in the image that represent the edges of the object in the image is termed edge detection. The edge detection process consists of three steps; filtering, enhancement and detection. Noise in the image due to random variation in intensity value is removed during filtering. Enhancement intensifies the pixels where there is change in local intensity. Edges are detected using thresholding. Robert edge detection, Sobel edge detection, Prewitt edge detection and Canny edge detection are the most commonly used detection methods. In this paper, it is proposed to use Sobel edge detector.

The Sobel edge detectors find the approximate absolute gradient magnitude at each point to detect edge. Regions of high spatial frequency corresponding to edge are obtained by 2-D gradient measurement. A series of gradient magnitudes are created using a simple convolution kernel. The convolution can be mathematically represented as:

$$N(x,y)=\sum_{k=-1}^{1}\sum_{j=-1}^{1}K(j,k)\,p(x-j,y-k)$$

For easy computation, the Sobel detector uses two convolution kernels for detecting changes in horizontal contrast (h_y) and vertical contrast (h_x).

Gabor filters have been extensively used in image processing for its ability to perform multi-resolution decomposition. This feature in Gabor filter is made available because of its localization in spatial and spatial frequency domain. This feature makes Gabor filters ideal for texture segmentation as it requires simultaneous measurements in both the spatial and the spatial-frequency domains. Filters with lower bandwidths are a desirable factor in the spatial-frequency domain as they allow making finer distinctions among different textures. To locate texture boundaries filters that are localized in the spatial domain are desired. The relations between the two are inversely related based on the uncertainty principle.

A Gabor function in the spatial domain is a sinusoidal modulated Gaussian. For a 2-D Gaussian curve with a spread of σ_x and σ_y in the x and y directions, respectively, and a modulating frequency of u_0, the real impulse response of the filter is given by

$$h(x,|\,y) = \frac{1}{2\pi\sigma_x\sigma_y}\exp\left\{-\frac{1}{2}\left[\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right]\right\}.\cos\left(2\pi u_0 x\right)$$

Radial Basis Function (RBF) is a variant of neural networks, which are better at interpolation, cluster modeling. RBF are embedded in two layer neural network as shown in figure 1 where the radial activated function is implemented in the hidden layer. RBF networks take in nonlinear inputs and give out linear outputs which help mapping complex models. The RBF is similarly trained as a neural network with training set. During training the network parameters are optimized by fitting the network outputs to the given inputs. Cost function is used to evaluate the fit of parameters; the cost function is usually assumed to be the square error. In pattern classification, the Gaussian activation function is used. The Gaussian activation function for RBF is given by:

$$\phi_j\left(X\right) = \exp\left[-\left(X-\mu_j\right)^T \sum\nolimits_j^{-1}\left(X-\mu_j\right)\right]$$

for j=1,…,L, where X is the input feature vector, L is the number of hidden units, μ_j and $\sum_j$ are the mean and the convariance matrix of the jth Gaussian function.

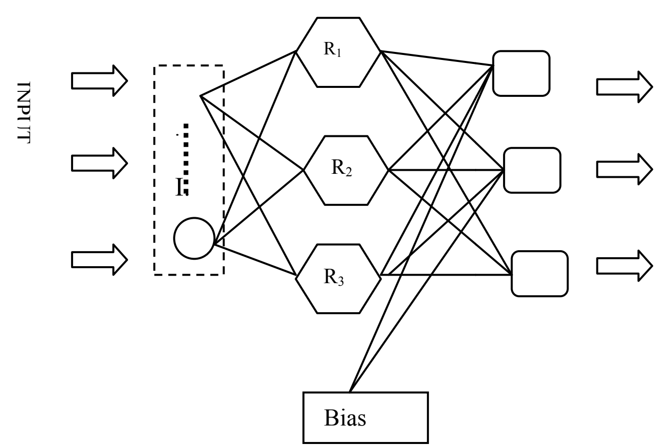

Fig. 1: A Typical Radial Basis Function Network

The output layer implements a weighted sum of hidden-unit outputs:

$$\psi_k\left(X\right) = \sum_{j=1}^{L}\lambda_{jk}\varphi_j\left(X\right)$$

for k=1,…,M where λ_{jk} are the output weights, each corresponding to the connection between a hidden unit and an output unit and m represents the number of output units.

The output of the RBF is limited to the interval (0,1) by a sigmoidal function as follows:

$$Y_k\left(X\right) = \frac{1}{1+\exp\left[-\psi_k\left(X\right)\right]}$$

for k=1,…,M.

IV. RESULT AND DISCUSSION

Nine species of plant leaves were selected [10] with 15 samples for each plant species. Sample image of the plant leaves used is shown in Fig. 2.

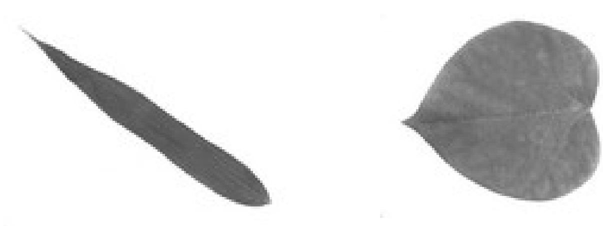

Fig. 2: Leaf samples used in this work

Matlab was used to extract the proposed features and the individual features fused. The extracted feature using Sobel edge detector is shown in Fig. 3.

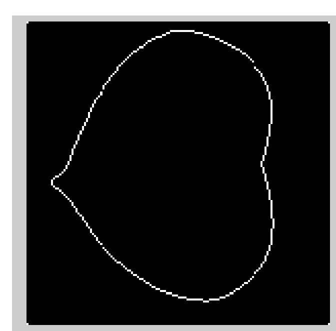

Fig. 3: The extracted features

The features extracted using Gabor filter with DCT is shown in Fig. 4 and Fig. 5 respectively.

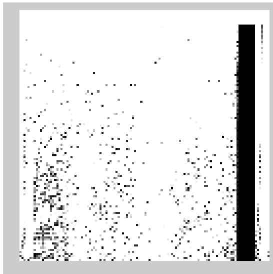

Fig. 4: The Gabor features

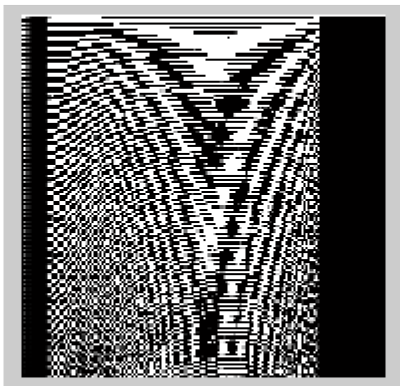

Fig. 5: Representation in frequency domain using DCT

The classification was done using 10 fold cross validation. The classification accuracy of CART and RBF are shown in Fig. 6.

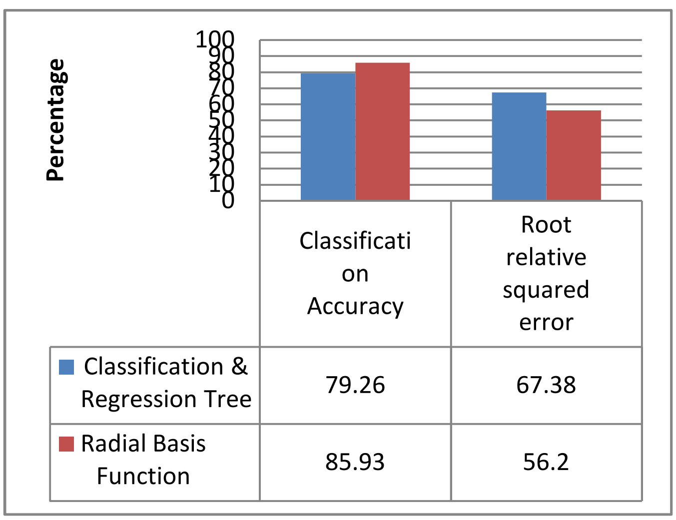

Fig. 6: Classification accuracy and Root relative squared error in percentage

V. CONCLUSION

In this paper a feature fusion technique was proposed using the Gabor filter in the frequency domain and fusing the obtained features with edge based feature extraction. The extracted features were trained using 10 fold cross validation and tested with CART and RBF classifiers. The output obtained using RBF is promising with low relative error for a nine class problem. Further work needs to be done to improve the classification accuracy by proposing feature reduction techniques.

REFERENCES

[1] Du, J. X. e*t al*,. "Computer-aided plant species identification (CAPSI) based on leaf shape matching technique" in *Transaction of the Institue of Measurement and Control*. 2006. Vol 28: 275-284.

[2] Wu, S. G. *et al*.,. "A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network", *IEEE International Symposium on Signal Processing and Information Technology*. 15-18 December 2007. Giza. 11-16

[3] Heymans, B. C. *et al*., "A neural network for opuntia leaf-form recognition" *Proceedings of IEEE International Joint Conference on Neural Networks*, *1991*. 18-21 November 1991. Singapore. Vol 3: 2116-2121

[4] C.L. Lee and S.Y. Chen "Classification for leaf images", in Proc. 16th IPPR Conf. Comput. Vision Graphics Image Process. 2003, Pg: 355–362.

[5] Y. Ye, C. Chen, C.-T. Li, H. Fu, and Z. Chi, "A computerized plant species recognition system," in Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing, Hong Kong, October 2004, PP: 723 – 726.

[6] H. Fu, Z. Chi, D. Feng, and J. Song, "Machine learning techniques for ontology-based leaf classification ," in IEEE 2004 8th International Conference on Control, Automation, Robotics and Vision, Kunming, China, Issue Date: 6-9 Dec. 2004, On page(s): 681 – 686 Vol. 1 2004.

[7] Xiao-Feng, Wang, De-Shuang, Huang, Ji-Xiang, Dua,HuanXu, LaurentHutte, "Classification of plant leaves with complicated background", Applied Mathematics and Computation,2008. Vol.205 pp.916–926.

[8] Ji-Xiang Du, De-Shuang Huang, Xiao-Feng Wang and Xiao Gu; "Computer-Aided Plant Species Identification (CAPSI) Based on Leaf Shape Matching; Transactions of the Institute of Measurement and Control", 2006 Vol. 28, 275–284.

[9] M. Z. Rashad, B.S.el-Desouky, and Manal S .Khawasik " Plants Images Classification Based on Textural Features using Combined Classifier" International Journal of Computer Science & Information Technology, 2011. Vol 3, No 4.

[10] Stephen Gang Wu, Forrest Sheng Bao, Eric You Xu, Yu-Xuan Wang, Yi-Fan Chang and Chiao-Liang Shiang, "A Leaf Recognition Algorithm for Plant classification Using Probabilistic Neural Network" in IEEE 7th International Symposium on Signal Processing and Information Technology, Dec. 2007, Cario, Egypt