

Plant Leaf Recognition using Gabor Filter

Jyotismita Chaki
School of Education Technology
Jadavpur University
Kolkata, West Bengal

Ranjan Parekh
School of Education Technology
Jadavpur University
Kolkata, West Bengal

ABSTRACT

This paper proposes an automated system for recognizing plant species based on leaf images. Plant leaf images of three plant types are analyzed using Gabor Filter by varying the filter parameters. Leaf images are convolved with Gabor filters followed by a separation of the real and imaginary portions of the signal. Absolute difference between the real and imaginary signals form the scalar feature value used for discrimination. Associated parameters like filter size, standard deviation, phase shift and orientation are varied to investigate which combination provides the best recognition accuracies. Classification is done by subtracting the test samples from the mean of the training set. The data set consists of 120 images divided into 3 classes. Accuracy obtained is comparable to the best results reported in literature.

General Terms

Pattern Recognition, Texture, Shape

Keywords

Gabor filter, leaf classification

1. INTRODUCTION

Plants form the majority of all life forms of the world. Plants produce almost all oxygen in the air that humans and other animals breathe, hence without plants it is difficult to think of existence of human life on earth. Plants are also useful because of their medicinal properties and as sources of alternative energy sources like bio-fuel. In recent times however many of the plants species are at the risk of extinction. A first step towards protection of plant species can be cataloging their features and benefits in a plant database. To handle the huge volume of plant diversity a quick and efficient method for plant recognition and classification needs to be developed. There are several ways to recognize a plant. Plants can be recognized by their flower, fruit, root, leaf etc. The quickest and easiest way to identify a plant is by its leaves. In recent years the image processing and pattern recognition techniques have been applied towards automated procedures of plant recognition.

This paper presents an automated system to recognize three species of plants by analyzing the digital images of their leaf. The organization of the paper is as follows: section 2 provides an overview of the previous works, section 3 outlines the proposed approach, section 4 details the experimentations done and results obtained, section 5 analyses the current work vis-à-vis contemporary works, section 6 brings up the overall conclusions and future scopes.

2. PREVIOUS WORK

Many methodologies have been proposed to analyze plant leaves in an automated fashion. Most of these methods use combinations of color, texture and shape features. One of the earliest works [1] uses geometrical parameters such as area,

perimeter, maximum length, maximum width, elongation to differentiate between four types of rice grains. To classify flowered and cactus plants [2] uses statistical discriminant analysis along with color based clustering and neural networks. In [3] the authors use the Curvature Scale Space (CSS) technique and k-NN classifiers to classify chrysanthemum leaves. In [4] both color and geometrical features have been used to detect weeds in crop fields. In [5] the authors propose a hierarchical technique to represent leaf shapes by first their polygonal approximations and then introducing more and more local details in subsequent steps. In [6] to detect weeds in an agricultural field fuzzy logic decision making has been used. In [7] for leaf image retrieval centroid-contour distance (CCD) and the object eccentricity have been used. The centroid-contour distance (CCD) curve and eccentricity along with an angle code histogram (ACH) have been used in [8] for plant recognition. The effectiveness of using fractal dimensions in describing leaf shapes has been explored in [9]. In contrast to contour-based methods, region based shape recognition techniques have been used in [10] for leaf image classification. In [11] to recognize the leaf shape Elliptic Fourier harmonic functions have been used. Leaf venation is used for leaf categorization in [12]. Leaf venations are represented using points selected by the curvature scale scope corner detection method on the venation image and categorized by calculating the density of feature points using non parametric estimation density. Neural Networks have been used in [13] to classify plant based on parameters like size, radius, perimeter, solidity and eccentricity. An accuracy of about 80% is reported. In [14] to model the uneven shapes of leaves Wavelet and Fractal based features have been used. In [15] to improve the accuracy texture features along with shape features have been used. In [16] to model the leaf structure Zernike moment and Polar Fourier transform have been proposed. The accuracy of 64% has been reported. In [17] the classification of leaf images, have been done using Gabor filter and Jeffrey divergence measure. In [18] authors propose Guiding Active Contours for plant leaf segmentation and classification.

3. PROPOSED APPROACH

This paper proposes a scheme for automated detection of 3 classes of plant category by analyzing shapes obtained from a collection of their leaf images using features based on Gabor filter. Classification is done with the help of Manhattan distance.

3.1 Gabor Filter

A complex Gabor is defined as the product of a Gaussian kernel times a complex sinusoid. The Gaussian function is called the envelope and the complex sinusoid is called the carrier. A 2D Gaussian curve with a spread of σ in both x and y directions, is represented as below:

$$g(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \cdot \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$$

The complex sinusoid is defined as follows, where u denotes the spatial frequency, θ denotes the orientation and φ the phase shift ($j = \sqrt{-1}$).

$$s(x, y, u, \theta, \varphi) = \exp\{j2\pi(x. u \cos\theta + y. u \sin\theta) + \varphi\}$$

The complex Gabor function can therefore be represented as follows

$$h(x, y, \sigma, u, \theta, \varphi) = g(x, y, \sigma).s(x, y, u, \theta, \varphi)$$

3.2 Feature Values

A grayscale image $I(x, y)$ is convolved with a Gabor filter with experimentally determined filter parameters to produce a set of complex signals.

$$J(x, y) = I(x, y) \otimes h(x, y, \sigma, u, \theta, \varphi)$$

The real and imaginary parts of the signal are separated out and converted to binary values using an appropriate threshold.

$$P(x, y) = \text{Re}\{J(x, y)\} \quad Q(x, y) = \text{Im}\{J(x, y)\}$$

The discrete signal is converted to a scalar representation by summing over all the discrete signal values.

$$R = \sum_x \sum_y P(x, y) \quad I = \sum_x \sum_y Q(x, y)$$

The feature value is generated by subtracting the imaginary value from the real value

$$F = |R - I|$$

Classification is done by dividing the dataset into a training set T and a testing set S each consisting of n samples. The i -th training class T_i is represented by the mean of the feature values of all its component samples.

$$M_i = \frac{1}{n} \{T_{i1}^F + T_{i2}^F + \dots + T_{in}^F\}$$

The j -th test sample S_j with feature value S_j^F is classified to class k if the absolute difference $D_{j,i}$ between the j -th test sample and i -th training class is minimum for $i = k$ i.e.

$$S_j \rightarrow k, \quad \text{if } D_{j,i} = |S_j^F - M_i| \text{ is minimum for } i = k$$

4. EXPERIMENTATIONS & RESULTS

Experimentations are performed by using 120 leaf images from the Plantscan database [19]. The dataset is divided into 3 classes A (Acer Negundo), B (Crataegus Azarolus) and C (Pistacia Lentiscus) each consisting 40 images. Each image is resized to 320 by 320 pixels in dimensions and saved in JPEG format. A total of 45 images are used as Training set and remaining 75 images as the Testing set. Samples of training and testing images are shown below in Figure 1 and Figure 2 :

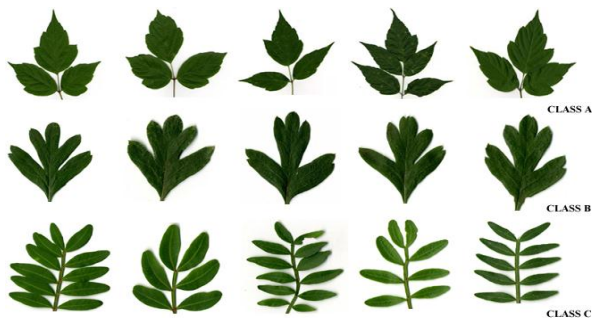


Fig 1: Samples of Training Set Images



Fig 2: Samples of Testing Set Images

Figure 3 shows the variation of the real values T^R , imaginary values T^I and feature values T^F of the samples of the training set for the 3 classes.

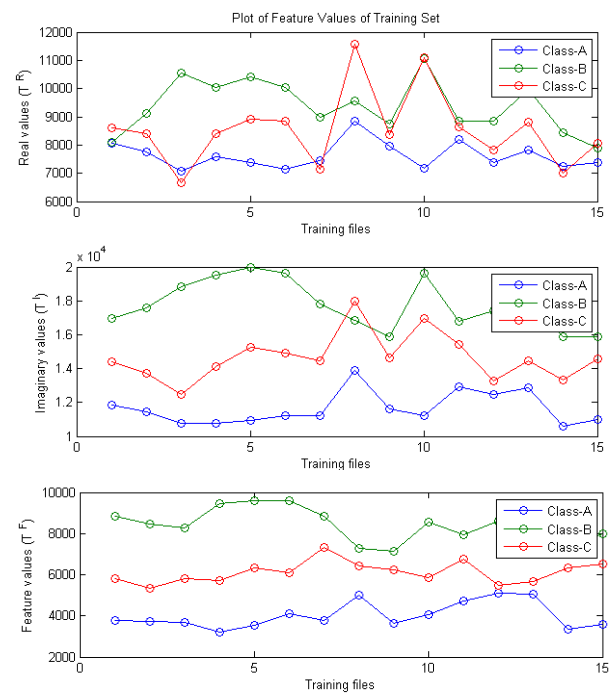


Fig 3: Variation of Feature values of Training set images for 3 classes

Figure 4 shows the variation of the real values S^R , imaginary values S^I and feature values S^F of the samples of the testing set for the 3 classes.

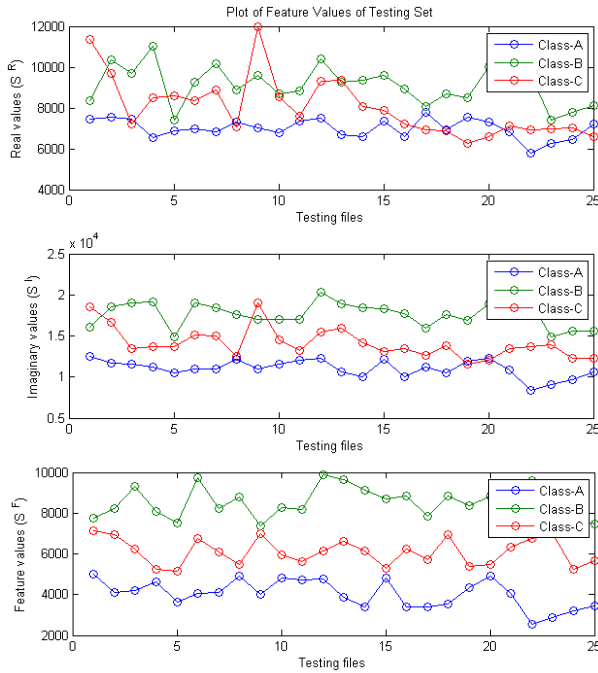


Fig 4: Variation of Feature values of Testing set images for 3 classes

The classifications plots of the testing set images are shown in Fig. 5. Class-A comprise of samples 1-25, Class-B of samples 26-50 and Class-C of samples 51-75. The difference (D) of each of these testing samples with the mean values of the Training sets A, B, C are depicted viz. $D(T_A), D(T_B), D(T_C)$

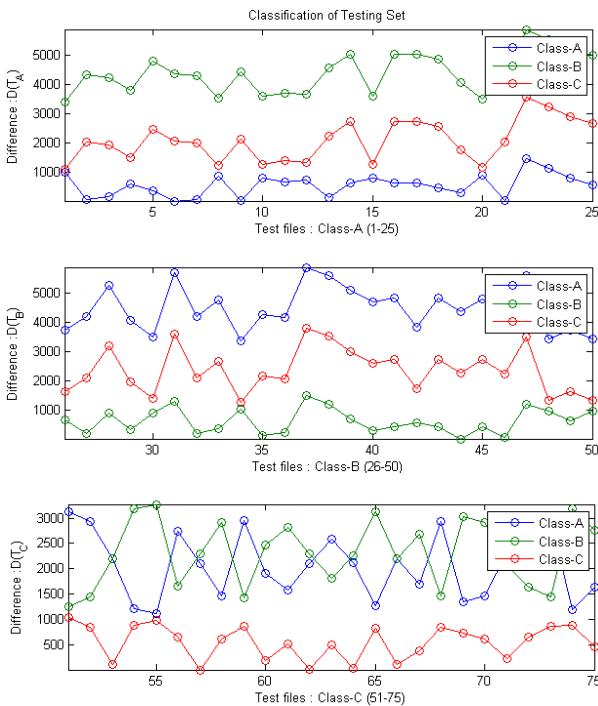


Fig 5: Classification plots of Testing set images

As expected for the first 25 samples difference with Class-A is minimum, for the next 25 files difference with Class-B is minimum and for the last 25 files difference with Class-C is minimum, which indicates their correct classifications.

The accuracy values for the three classes are tabulated below

Table 1. Recognition Accuracies

Feature	Class A	Class B	Class C	Overall Accuracy
F	100%	100%	100%	100%

5. ANALYSIS

Automated discrimination between three leaf shapes was done using Gabor filter. The overall accuracy obtained is 100% for this experiment. To put the above results in perspective with the state of the art, the best results reported in [8] has a recall rate of 60% for discrimination of chrysanthemum leaves from a database of 1400 color images. Accuracy for classification for 10 leaf categories over 600 images is reported to be 82.33% in [10]. Overall classification accuracy reported in [11] for 4 categories of leaf images obtained during three weeks of germination are around 90%. An overall classification of 80% is reported in [13] for identifying two types of leaf shapes from images taken using different frequency bands of the spectrum. In [17] 95.5% accuracy is reported while considering the Brodatz dataset. Best results reported in [16] are around 93% using Polar Fourier Transform. The accuracy reported in the current paper is therefore seen to be comparable to the best results reported in extant literature.

To arrive at the best combination of the Gabor filter parameters, the size of the Gaussian filter size was varied and accuracy pertaining to each was plotted. Figure 6(a) shows the plot of accuracy as filter size is varied. The best accuracy is observed to be obtained at size 6. For size 6, angle φ is varied from 0 to 2π in multiples of $\pi/8$. Figure 6(b) shows how accuracy changes as the multiple is varied. The best results are obtained when $\varphi = 7 * \pi/8$

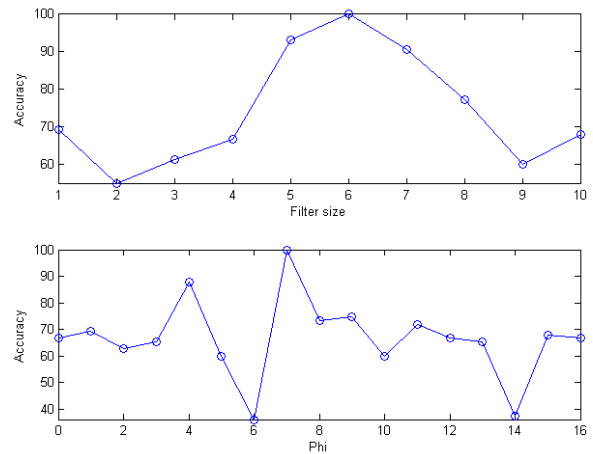


Fig 6: Variation of accuracy with (a) filter size (b) phi

Keeping filter size at 6 and $\varphi = 7\pi/8$ the angle of orientation θ is varied between 1 to 10. Figure 7(a) shows the variation of accuracy with varying values of θ . The best results are obtained at $\theta = 2$. Keeping filter size at 6, $\varphi = \frac{7\pi}{8}$, $\theta = 2$ the value of standard deviation σ is varied between 0.1 to 1. Figure 7(b) shows the variation of accuracy with varying values of σ . The best values are obtained at $\sigma = 0.65$.

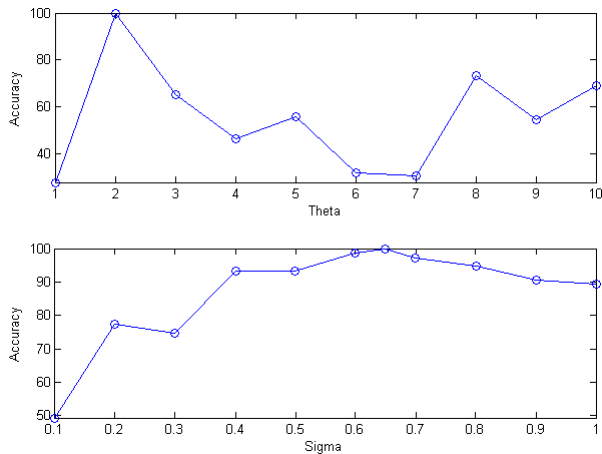


Figure 7: Variation of accuracy with (a) theta (b) sigma

6. CONCLUSIONS & FUTURE SCOPES

This paper proposes an automated system for plant identification using shape features of their leaves. Such automated classification system can prove extremely useful for quick and efficient classification of plant species. The accuracy of the current proposed approach is comparable to those reported in contemporary works. The current approach also based on low complexity data modeling scheme as it involves only scalar values. For less complex shape of leaf images, consisting of a single lobe, standard shape description schemes such as Centroid-Radii model or Moment-Invariants model might be sufficient, as is evident from the current authors' previous works [20] and [21], but these schemes do not provide satisfactory results when the shape becomes more complex, especially when multi-lobed leaves are involved. The current proposed approach provides an alternative for discriminating complex multi-lobed leaf shapes.

Future work would involve research along three directions:

- (1) Combining other shape based techniques like Hough transform and Fourier descriptors
- (2) Combination of colour and texture feature with the shape feature
- (3) Making the recognition process invariant to rotation and scaling

7. REFERENCES

- [1] Sakai.N., Yonekawa.S. and Matsuzaki.A. 1996. Two-dimensional image analysis of the shape of rice and its applications to separating varieties. *Journal of Food Engineering*. 397-407.
- [2] Timmermans.A. J. M and Hulzebosch.A. A. 1996. Computer vision system for on-line sorting of pot plants using an artificial neural network classifier. *Computers and Electronics in Agriculture*. 41-55.
- [3] Abbasi.S, Mokhtarian.F and Kittler.J. 1997. Reliable classification of chrysanthemum leaves through curvature scale space. *Lecture Notes in Computer Science*. 284-295.
- [4] Perez.A. J, Lopez.F Benlloch.J. V and Christensen.S. 2000. Color and shape analysis techniques for weed detection in cereal fields. *Computers and Electronics in Agriculture*. 197-212.
- [5] Im.C, Nishida.H., and Kunii.T. L. 1998. A hierarchical method of recognizing plant species by leaf shapes. *IAPR Workshop on Machine Vision Applications*. 158-161.
- [6] Yang.C-C, Prasher.S. O, Landry.J-A, Perret. J and Ramaswamy.H. S. 2000. Recognition of weeds with image processing and their use with fuzzy logic for precision farming. *Canadian Agricultural Engineering*. 195-200.
- [7] Wang.Z, Chi.Z, Feng.D and Wang.Q. 2000. Leaf image retrieval with shape feature. *International Conference on Advances in Visual Information Systems (ACVIS)*. 477-487.
- [8] Wang.Z, Chi.Z and Feng.D. 2003. Shape based leaf image retrieval. *IEEE Proceedings on Vision, Image and Signal Processing (VISIP)*. 34-43.
- [9] Camarero.J. J., Siso.S., and Pelegrin.E.G. 2003. Fractal dimension does not adequately describe the complexity of leaf margin in seedlings of *Quercus* species. *Anales del Jardín Botánico de Madrid*. 63-71.
- [10] Lee.C-L and Chen.S-Y. 2003. Classification of leaf images. *16th IPPR Conference on Computer Vision, Graphics and Image Processing (CVGIP)*. 355-362.
- [11] Neto.J. C., Meyer.G. E., Jones.D. D. and Samal.A. K. 2006. Plant species identification using elliptic Fourier leaf shape analysis. *Computers and Electronics in Agriculture*. 121-134.
- [12] Park.J.K, Hwang.E.J, and Nam.Y. 2006. A venton – based leaf image classification scheme. *Alliance of Information and Referral Systems*. 416-428.
- [13] Pan.J and He Y. 2008. Recognition of plants by leaves digital image and neural network. *International Conference on Computer Science and Software Engineering*. 906 – 910.
- [14] Wang.Q.P, Du J.X, and Zhai.C-M. 2010. Recognition of leaf image based on ring projection wavelet fractal feature. *International Journal of Innovative Computing, Information and Control*. 240-246.
- [15] Beghin.T, Cope.J. S., Remagnino.P., and Barman.S. 2010. Shape and texture based plant leaf classification. *International Conference on Advanced Concepts for Intelligent Vision Systems (ACVIS)*. 345-353
- [16] Kadir.A, Nugroho.L.E, Susanto.A, and Santosa.P.I. 2011. A comparative experiment of several shape methods in recognizing plants. *International Journal of Computer Science & Information Technology (IJCSIT)*. 256-263.
- [17] Cope.James S, Remagnino.Paolo, Barman.Sarah, and Wilkin.Paul. 2010. Plant Texture Classification Using Gabor Co-occurrences. *Springer*. 669-677.
- [18] Cerutti.G, Tougne.L, Mille.J, Vacavant.A., and Coquin.D. 2011. Guiding active contours for tree leaf segmentation and identification. *Cross-Language Evaluation Forum (CLEF)*.
- [19] Plantscan database (URL: http://imedia-ftp.inria.fr:50012/Pl@ntNet/plantscan_v2/).
- [20] Chaki. J. and Parekh.R. 2011. Plant Leaf Recognition using Shape based Features and Neural Network classifiers. *Int. Journal of Advanced Computer Science and Applications (IJCSA)*. 41-47.
- [21] Chaki. J. and Parekh. R. 2012. Designing an Automated System for Plant Leaf Recognition. *Int. Journal of Advances in Engineering & Technology (IAET)*. 149-158.