ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

*International Journal of Advanced Research in Computer and Communication Engineering*
*Vol. 4, Issue 9, September 2015*

# Plant Disease Identification using Segmentation Techniques

**S. Nagasai[1], S. Jhansi Rani[2]**

M. Tech CST, Dept. of Computer Science & Systems Engg, Andhra University, Visakhapatnam, India [1]

Assistant professor, Dept. of Computer Science & Systems Engg, Andhra University, Visakhapatnam, India[2]

**Abstract:** Plant diseases are the major cause of production in agricultural and medicinal industry worldwide. It leads to economic losses. Monitoring of health and detection of diseases in plants and trees is a critical issue. This thesis describes a method for identification of diseased rose plants based on some important features extracted from its leaf images. The most significant part of research on plant disease to identify the disease based on CBIR that is mainly concerned with the accurate detection of disease of rose plant. We present an approach where the leaf is identified based on its leaf features such as Color, shape using Color histogram and edge histogram. The combination of CBIR, Canny edge detector and HSI Color model identifies the disease accurately.

**Keywords:** CBIR, Segmentation, Edge detection, canny algorithm, Edge histogram, HSI histogram, SVM classifier.

## I. INTRODUCTION

For years, Plant diseases are the main source of plant damage which causes major economic and production losses in agricultural areas. Hence we need to grow more plants. This rejuvenation work requires easy recognition of the diseases in plant leaves.

There are different types of diseases as bacterial, fungal and viral diseases. These diseases can be detected by identifying the symptoms such as spots and damage arising on the leafs. In order to manage these diseases effectively, there is a need of an automated system for identifying the condition of the plant.

A variety of techniques are already developed for identifying the diseases in plants. But they are poor in accuracy. In order to obtain accuracy, this paper presents a method which uses various techniques of image processing. In training phase, edge detection, one of the effective methods of image segmentation is performed to get the edge features and then they are used for classification. Here, the CBIR plays a vital role in identifying the disease of the leaf. Content Based Image retrieval acts like a query extension. CBIR retrieves the related images from larger datasets effectively. Depending on the features of the query leaf, the CBIR fetches the database leaf image containing the similar features such as the diseased portion, spots and damage. The structure of the paper is organized as follows: Sec.2 is an overview of related work in this field; Sec.3 explains the segmentation techniques applied, Sec.4 explains the implementation, Sec.5 presents the experimental results, and finally Sec.6 concludes the chapter.

## II. RELATED WORK

K.Padmavathi present a study was completed to investigate the use of computer vision and image processing techniques in agricultural applications. In biological science, sometimes thousands of images are generated in a single experiment. These images can be required for further analysis of plant diseases [1]. In order to analyze these images there is a need to fetch the features of these images, which include segmentation techniques.

There are different techniques for segmentation proposed and developed so far. One of the segmentation techniques here applied is edge detection. There are two categories of operators for identifying edges. Some of them are listed here. First order derivatives and second order derivatives.

In First order derivative [2] the input image is convolved by an adapted mask to generate a gradient image in which edges are detected by thresholding. Most classical operators like sobel, prewitt, Robert [4] are the first order derivative operators. These operators are also said as gradient operators. These gradient operators detect edges by looking for maximum and minimum intensity values.

In Second order derivatives [2], these are based on the extraction of zero crossing points which indicates the presence of maxima in the image. In this, image is first smoothed by an adaptive filters [5].Since the second order derivative is very sensible to noise, and the filtering function is very important. These operators are derived from the Laplacian of a Gaussian (LOG), and proposed by Marr and Hildreth [5], in this, the image is smoothed by a Gaussian filter. A significant problem of LoG is that the localization of edges with an asymmetric profile by zero-crossing points introduces a bias which increases with the smoothing effect of filtering [8].In order to serve this problem the solution was proposed by Canny [9], which says in an optimal operator for step edge detection includes three criteria: good detection, good localization, and only one response to a single edge. Edge histogram is an effective representation of the finalized edges in order identifies the structure of leaf.

Color histogram is a best way of identifying the color of the infected portions. HSI color model [3] is used here because it is very close to human eye perception of colors.

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

IJARCCE

*International Journal of Advanced Research in Computer and Communication Engineering*
*Vol. 4, Issue 9, September 2015*

## III. SEGMENTATION METHODS

### A. Canny edge detector

Canny edge detector is the optimal edge detection operator. Edge detection is used to obtain the useful structural information of the image by reducing the data to be processed. Canny has found that, the requirements for the application of edge detection on different computer vision systems are relatively the same. An optimal edge detector should satisfy the following criteria,

i. Detection**:** The probability of detecting real edge points should be maximized while the probability of falsely detecting non-edge points should be minimized. This corresponds to maximizing the signal-to-noise ratio.
ii. Localization**:** The detected edges should be as close as possible to the real edges.
iii. Numberof responses: One real edge should not result in more than one detected edge.

**ALGORITHM**

This algorithm is broken in to five steps,
1. Smoothing
2. Finding gradients
3. Non-maximum suppression
4. Double thresholding
5. Edge tracking by hysteresis

### B. Edge histogram

Edge histogram is a graphical representation of the tonal distribution over an image. Edge histogram is applied to find the shape of the leaf. Then, it is further used as a parameter for identification of disease.

Parameters involved in the selection of an edge detection operator include:

- Edgeorientation**:** The geometry of the operator determines a characteristic direction in which it is most sensitive to edges. Operators can be optimized to look for horizontal, vertical, or diagonal edges.
- Noiseenvironment**:** Edge detection is difficult in noisy images, since both the noise and the edges contain high-frequency content. Attempts to reduce the noise result in blurred and distorted edges. Operators used on noisy images are typically larger in scope, so they can average enough data to discount localized noisy pixels. This results in less accurate localization of the detected edges.
- Edge structure: Not all edges involve a step change in intensity. Effects such as refraction or poor focus can result in objects with boundaries defined by a gradual change in intensity.

### C. HSI-Color Histogram

Color histogram represents the color distribution over all pixels of image. Generally, Color histogram will divide all the pixel colors into RED, GREEN and BLUE combinations. This will skip a few colors of diseased portions. In order to avoid performance deficiency due to this, here the implementation of Color histogram in HSI color model is introduced.

HSI color histogram defines the colors with respect to normalized Red, Green and Blue values, given in terms of RGB primaries as,

$$r = \frac{R}{R+G+B} \qquad (1)$$

$$g = \frac{G}{R+G+B} \qquad (2)$$

$$b = \frac{B}{R+G+B} \qquad (3)$$

From the above equations, it is concluded that

$$r + g + b = 1 \qquad (4)$$

For the R, G and B values of each pixel the hue, saturation and intensity values can be determined as follows,

$$I = \frac{1}{3(R+G+B)} \qquad (5)$$

The value of I lies between {0, 1}

$$S = 1 - \frac{3}{R+G+B}[\mathrm{Min}(R, G, B)] \qquad (6)$$

The value of S ranges between {0, 1}.

And,

$$H = \cos^{-1}\left\{\frac{\frac{1}{2}*[(R-G)+(R-B)]}{\sqrt[2]{[(R-G)^2+(R-B)(G-B)]}}\right\} \qquad (7)$$

The H value will vary in between [0, 360°].

For every pixel of strong edges, the three h, s, i values are calculated and their mean and median values along with the image name a feature vector is stored in to database for every training and testing images.

## IV. EXPERIMENTAL SETUP & PROCEDURE

The proposed methodology gives the identification of plant disease based on its edge and color features. First, the color images are resized to 200*200 pixels and converted to gray scale. Canny edge detector is applied on the gray scale image to obtain the strong edges of the leaf.

### A. DatasetPreparation

The dataset for this work is prepared manually. Rose leaf samples were acquired from web. Every leaf image is pre-processed and resized to 200*200. Then they will be categorized into healthy and diseased using SVM classifier.

### B. Segmentation

Segmentation is a crucial step in our process. Segmentation separates the diseased portion from the leaf image. For this purpose, we have chosen Canny edge detector.

The edge detection performs as follows,

i. For smoothing, Gaussian filter is applied with the filter size 5*5. General filter is of size 2*2, it cause the detection of useless edges. The larger the filter size, detection of useful and smoother edges.
ii. The gradients in X and Y directions is calculated where the 3*3 window.
iii. Non-maximum suppression is an edge thinning technique. The pixel whose edge value is greater than its neighbor's is considered as strong.
iv. Double thresholding is applied on the result after suppression. The decision about the threshold value decides the result of potential edges.
v. Hysteresis tracks the potential edges which are connected by suppressing the remaining pixel values.

Thus the final strong edges are fetched from the image.

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

IJARCCE

*International Journal of Advanced Research in Computer and Communication Engineering*
*Vol. 4, Issue 9, September 2015*

### C. Disease identification

After evaluation of edges, the H, S and I values are calculated. Feature vectors are calculated and stored along with the image and disease name. When a testing image is given as an input, a feature vector for it also will be created and stored into database. Euclidean distance for the testing feature vector and all the dataset feature vectors is calculated. After comparing the testing and training feature vectors the result disease along with the HSI histograms of testing and matched leaf will be displayed.
The proposed system is a step-wise process is illustrated in Fig-1.

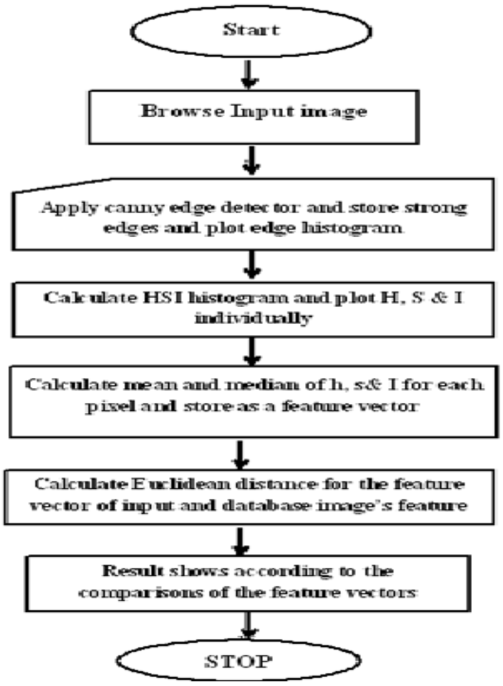


Fig 1: Procedure for Disease identification

## V. RESULTS & DISCUSSIONS

The proposed work implementation is performed using MATLAB. The canny edge detection applied to a testing leaf shown in fig 2 will display each and every stage of the algorithm very clearly.

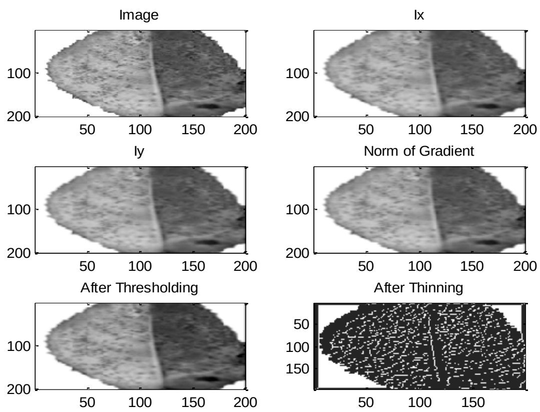


Fig 2: result of Canny edge detector

The edge histogram of the testing and matched training set leaf respectively arranged as fig 3,

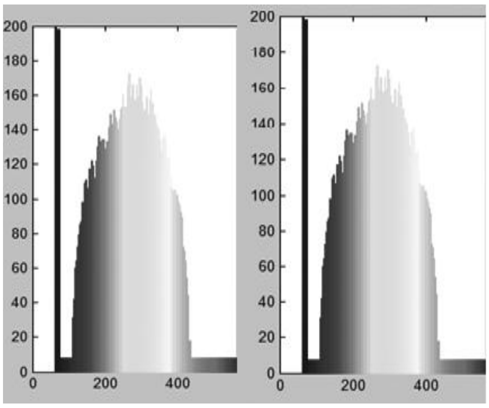

Fig 3: Edge histograms of test and training leaf

The final result after comparing the testing image with the training set the result displayed can be viewed as fig 4,

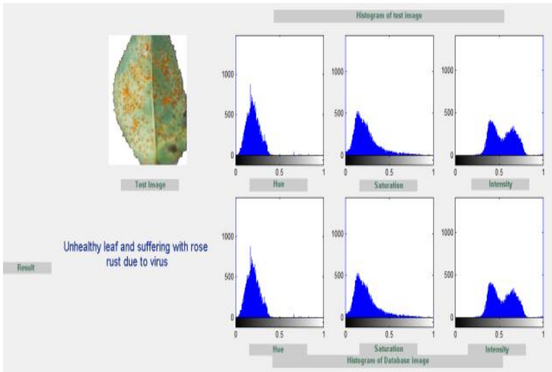


Fig 4: The rested result of rose leaf with better accuracy.

## VI. CONCLUSION& FUTURE SCOPE

Canny edge detector is implemented efficiently and the results are very accurate. The combination of Canny edge detector and HSI color histogram fetched maximum resourceful features, which have been classified using SVM classifier resulted in very better detection of disease.
In future to achieve high tolerance and greater accuracy, we can combine canny algorithm and neural network with back propagation for edge detection for all types of plants. This method is a little complex but achieves both the advantages of canny edge detection and neural networks.

## REFERENCES

[1] K. Padmavathi, "Investigation and monitoring for leaves disease detection and evaluation using image processing", International Research Journal of Engineering Science, Technology and Innovation (IRJESTI), Vol. 1,Issue 3, pp. 66-70, June 2012.

[2] M. Juneja, P. Sandhu, "Performance evaluation of edge detection Techniques for images in spatial domain" International Journal of Computer theory and Engineering, Vol. 1,No. 5,December, 2009 1793- 8201.

[3] Derivation of HSI –to- RGB and RGB-to-HSI conversion Equations, Gonzalez and Woods, Digital Image Processing, 1st ed. Addison-Wesley, 1992.

[4] William K Pratt, Digital image processing, 4th edition , John WileyInc,2007.

[5] MARR, HILDRETH. Theory of edge detection. Proc.R. Soc. Lond. (1980), B 207, pp. 187-217.

[6] HARALICK (R. M.), LEE (J.). Context dependent edge detector. Proc.CVPR 88, Ann Arbor, Michigan (1988), pp.223-228.

[7] J F CANNY," A computational approach to edge detection." 1EEE Trans. PAMI (1986), 8, n ~ 6, pp. 679-698.