# Feature Extraction of Leaf Diseases

**Mr. Hrishikesh P. Kanjalkar, Prof. S.S.Lokhande**

*Abstract* -**Vegetables and fruits are most important agricultural products. In order to obtain more value added products, a proper quality control is essentially required. Many studies show that quality of agricultural products may be reduced from many causes. One of the most important factors of such quality is plant diseases. This proposed paper presents some important features of diseased leaves which will help us to find exact disease of plant.**

*Index Terms*— **Diseased leaves.**

## I. INTRODUCTION

Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately. When performing analysis of complex data one of the major problems stems from the number of variables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power or a classification algorithm which over fits the training sample and generalizes poorly to new samples. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy.

Automatic detection of plant diseases is an important research topic as it may prove benefits in monitoring large fields of crops, and thus automatically detect the diseases from the symptoms that appear on the plant leaves. This enables machine vision that is to provide image based automatic inspection, process control and robot guidance. Comparatively, visual identification is labor intensive, less accurate.

When crops suffered disease, if the symptoms are not obvious or more complex, it will be very difficult to distinguish the characteristics so that delay the effective control of crop diseases seriously. We can analyze the image of disease leaves by using computer image processing technology and extract the features of disease spot according to color, texture and other characteristics from a quantitative point of view. The caused and extent of disease can be diagnosed timely and effective, it could be prevented and control comprehensive combined with the specific circumstances of crop. It has a great significance in automatic and intelligent management aspects of crop growth and health with the research of crop diseases using image feature extraction technology.

The quality inspection of leaves consists of two main aspects, internal and external examinations. The internal quality inspection is usually achieved by human sensory, smoking test or chemical analysis, while the external quality inspection is mainly achieved through human vision. It is costly and yet time-consuming to inspect internal quality since leaves contain too many ingredients to be handled. As an alternative, external quality examination is often used instead in the examination of internal quality of leaves, since external features are closely related to internal quality. The external quality inspection of leaves includes judgment of color, maturity, surface texture, size and shape. Human vision, which is inevitably limited by personal, physical and environmental factors, has been the predominant means of inspection.

## II. SEGMENTS AND FEATURES

### *A. Color Transformation Structure*

First, the RGB images of leaves are converted into Hue Saturation Intensity (HSI) color space representation. The purpose of the color space is to facilitate the specification of colors in some standard, generally accepted way. HSI (hue, saturation, intensity) color model is a popular color model because it is based on human perception. Hue is a color attribute that refers to the dominant color as perceived by an observer. Saturation refers to the relative purity or the amount of white light added to hue and intensity refers to the amplitude of the light. Color spaces can be converted from one space to another easily. After the transformation process, the H component is taken into account for further analysis. S and I are dropped since it does not give extra information.

### *B. The HSI color model*

The RGB and CMY color models are not suited for describing colors in terms of human interpretation. Color objects in images can be easily described by its hue, saturation, and brightness (intensity). Hence the HSI color model has been

***Mr. Hrishikesh P. Kanjalkar**, Department of E and TC, Sinhagad College of Engineering, Pune, India.*
***Prof. S.S.Lokhande**, Department of E and TC, Sinhagad College of Engineering, Pune, India.*

presented. The HSI model decouples the intensity component from the color-carrying information (hue and saturation) in a color image. As a result, this model is an ideal tool for developing color image processing algorithms. The hue, saturation, and intensity values can be obtained from the RGB color cube. That is, we can convert any RGB point to a corresponding point in the HSI color model by working out the geometrical formulas.

Converting colors from RGB to HSI

The hue **H** is given by,

$$H = \begin{cases} \theta, & \text{if } B \leq G \\ 360 - \theta, & \text{if } B > G \end{cases} \quad (1)$$

Where,

$$\theta = \cos^{-1}\left\{\frac{1/2[\![(R-G)+(R-B)]\!]}{\sqrt{[\![(R-G)^2+(R-B)(G-B)]\!]}}\right\} \quad (2)$$

The saturation **S** is given by,

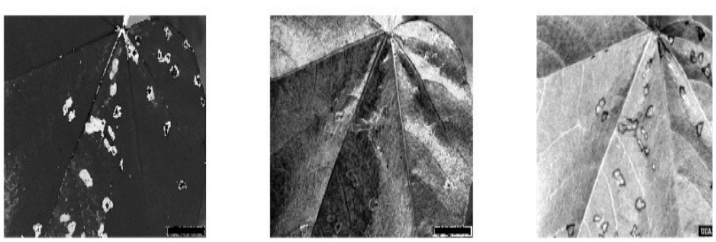
$$S = 1 - \frac{3}{(R+G+B)}[\min(R, G, B)] \quad (3)$$

The intensity **I** is given by,

$$I = \frac{1}{3}(R + G + B) \quad (4)$$

All RGB values are normalized to the range [0, 1].


Fig.1 Original image (angular leaf spot of cotton)



a b c

Fig.2 (a) Hue image (b) Saturation image (c) Intensity image

Above figures clearly shows that, the selection of H S I image will definitely give more accurate results. For segmentation and feature extraction more discriminative image will be useful, therefore Hue image from HIS image is selected which is more discriminative.

At first, all the connected components are extracted from original diseased image which are diseased parts of leaves. Connected component labeling is an algorithmic application of graph theory, where subsets of connected components are uniquely labeled based on a given test. It actually collects all same or near valued pixels, in our case these regions are diseased parts of the leaves.

In the connected component labeling sometimes it labels the unwanted regions, therefore to remove these unwanted areas we use thresholding operation. Here we are removing all connected components (objects) that have fewer than 30 pixels from a binary image, producing another binary image.

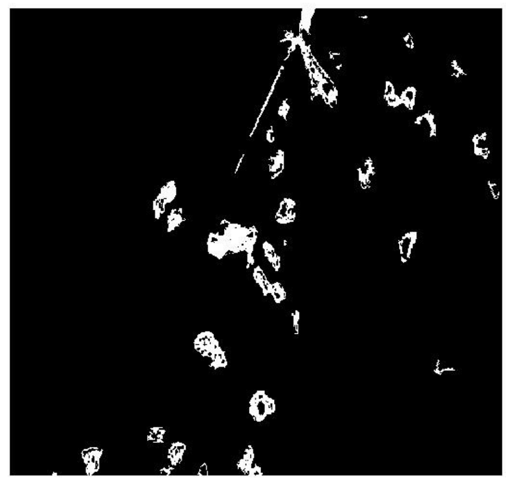


Fig. 3: Extracted diseased spots.

*C. Image Feature Extraction*

The features selected here are considered according to the leaf diseases which will be more discriminative. In proposed approach total 10 features are taken, i.e. 8 color features and size of disease (size of diseased spot), distances of diseased spots from each other.

Color features from images are taken by sum and average technique. We have 768x768 size image database, therefore from these images we are getting 8 distinct values. In extraction process, first we take sum of all columns and then average of 100 pixels gives 8 values.

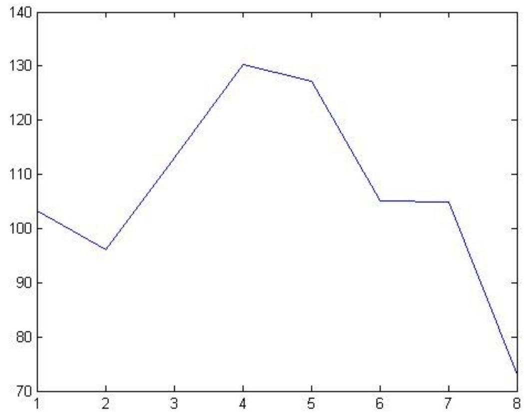



Fig. 4: Color feature, 8 values.

If the centroid is defined, it is a fixed point of all isometries in its symmetry group. In particular, the geometric centroid of an object lies in the intersection of all its hyperplanes of symmetry. The centroid of many figures (regular polygon, regular polyhedron, cylinder, rectangle, rhombus, circle, sphere, ellipse, ellipsoid, superellipse, superellipsoid, etc.) can be determined by this principle alone.

The centroid of a finite set of k points $X_1, X_2 ... X_k$ in $R^n$ is

$$C = \frac{X1+X2+\cdots+Xk}{k} \qquad (5)$$

Distance from centroid to centroid is calculated. These distances are calculated from each centroid to other remaining centroids, which then considered for average distance of centroids.

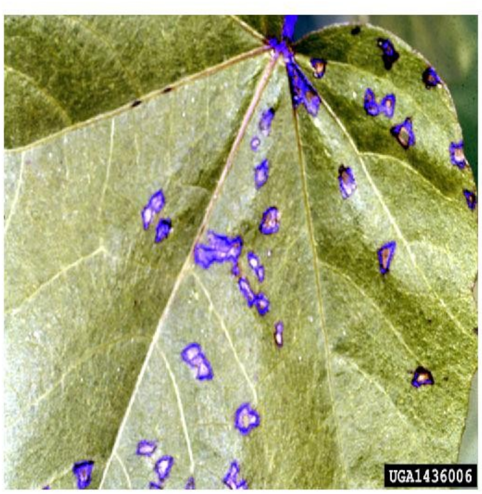



Fig. 5: Detected diseased area in original image.

## III. CONCLUSION

The proposed methodology uses leaf features for disease detection where feature extraction is done on segmented diseased area. Hue image from HSI gives clear discrimination of diseased spots, and which is more helpful for extracting size, color and centroids.

## REFERENCES

[1] Dheeb Al Bashish, Malik Braik, and Sulieman Bani-Ahmad, “A Framework for Detection and Classification of Plant Leaf and Stem Diseases,” Department of Information Technology, *IEEE International Conference on Signal and Image Processing,* 2010.

[2] Santanu Phadik ar and Jaya Sil, “Rice Disease Identification using Pattern Recognition Techniques,” 11th International Conference on Computer and Information Technology (ICCIT 2008),*IEEE Proceedings*, 2008.

[3] S. Ananthi, S. Vishnu Varthini, “Detection and classification of plant leaf diseases,” *International Journal of Research in Engineering & Applied Sciences*, Volume 2, Issue 2 (February 2012), ISSN: 2249-3905, 2012.

[4] Jiazhi Pan, Yong He, “Recognition of plants by leaves digital image and neural network*,” International Conference on Computer Science and Software Engineering,IEEE Computer Society*, 2008.

[5] Xinhong Zhang, Fan Zhang, “Images Features Extraction of Tobacco Leaves,” *Congress on Image and Signal Processing, IEEE Computer Society*, 2008.

[6] A. Meunkaewjinda, P. Kumsawat, K. Attakitmongcol and A. Srikaew,“Grape leaf disease detection from color imagery using hybrid intelligent system,” *IEEE Proceedings of ECTI-CON*, 2008.

[7] Zulkifli Bin Husin, Abdul Hallis Bin Abdul Aziz, Ali Yeon Bin Md Shakaff, Rohani Binti S Mohamed Farook, “Feasibility Study on Plant Chili Disease Detection Using Image Processing Techniques,” Third International Conference on Intelligent Systems Modelling and Simulation, *IEEE Computer Society*, 2012.

[8] Panagiotis Tzionas, Dimitris Manolakis, Stelios E. Papadakis “Plant leaves classification based on morphological features and a fuzzy surface selection technique.”