An Image Recognition System for Crop Disease Identification of Paddy fields in Sri Lanka

Conference Paper · December 2009

DOI: 10.1109/ICIINFS.2009.5429828

2 authors:

Gehan Anthonys
The Open University of Sri Lanka

Nalin Wickramarachchi
University of Moratuwa

An Image Recognition System for Crop Disease Identification of Paddy fields in Sri Lanka

G. Anthonys[1], N. Wickramarachchi[2]

[1] Department of Electrical and Computer Engineering
The Open University of Sri Lanka
Nawala, Sri Lanka
ganth@ou.ac.lk

[2] Department of Electrical Engineering
University of Moratuwa,
Moratuwa, Sri Lanka
wick@elect.mrt.ac.lk

Abstract— **The classification and recognition of paddy diseases are of the major technical and economical importance in the agricultural industry. To automate these activities, like texture, color and shape, disease recognition system is feasible. The goal of this research is to develop an image recognition system that can recognize paddy diseases.**

Images were acquired under laboratory condition using digital camera. Three major diseases commonly found in Sri Lanka, Rice blast (*Magnaporthe grisea)*, Rice sheath blight (*Rhizoctonia solani)* and Brown spot (*Cochiobolus miyabeanus)* were selected for this research.

Image processing starts with the digitized a color image of paddy disease leaf. Then a method of mathematics morphology is used to segment these images. Then texture, shape and color features of color image of disease spot on leaf were extracted, and a classification method of membership function was used to discriminate between the three types of diseases.

The analysis of the results showed over 70 percent classification accuracy around 50 sample images.

Index Terms— **Mathematics morphology, CIE *L*a*b** color space, Color texture, Membership function.**

I. INTRODUCTION

HUMANS ar e pr imarily visu al creatur es and their perception has th e capab ility t o acq uire, in tegrate, ar id interpret all these abundant visual information around us. It is ch allenging t o im part s uch cap abilities to a m achine in order to in terpret the visual in formation em bedded in stil l images, graphics and video or moving images in our sensory world. An Image Reco gnition p roblem can b e d efined as a labeling problem based on m odels of known images. It finds objects in the real world from an image, which are known a priori. This task is surprisingly difficult.

Paddy diseases pose a major threat to rice pr oduction in Sri Lanka [1], because millions of rupees in direct loses and loses related to use o f control measures. Ever since humans started farming, diseases h ave b een o ne of the m ajor ob stacles i n maximizing p roduction. Th ere are n umerous d iseases of rice such as fungi, bacteria, viruses and nematode. Some diseases occur wherever rice is gr own. Some diseases reach epidemic proportion and cau se ser ious cr op lo sses wh ere o thers cause only negligible crop losses.

Field problems in rice cultivation could broadly be divided into three major categories viz. insect pests, weeds, and plant diseases. Identification and m anagement of pro blems re lated to insect pests and weeds are not considered in this research. This r esearch f ocused o n accur ate identification o f so me diseases of paddy plants in the laboratory.

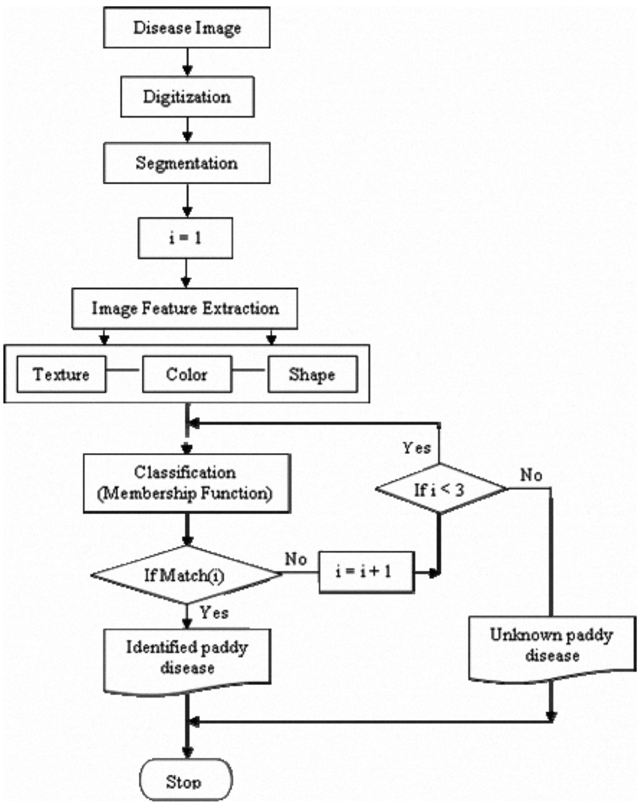

Fig. 1. Modularized architecture of the overall system

Fig.1 sh ows t he ov erall modularization of t he s ystem. Functionalities o f each m odule are ex plained in th e secti ons here after.

II. METHODOLOGY

A. Image Acquisition

Digital camera was used to collect the digital RGB images. The images were taken in side the laboratory with dark background in order to control the environmental factors such as lighting, illumination, intensity etc.

The image taken from digital camera in the laboratory transferred to the computer. Then system starts Image processing. The steps involved in the processing include image digitization, segmentation, image feature extraction, and classification.

B. Digitization

An image captured by a sensor is expressed as a continuous function $f(x,y)$ of two co-ordinates in the plane. Image digitization means that the function $f(x,y)$ is sampled into a matrix with M rows and N columns. Image quantization assigns to each continuous sample an integer value the continuous range of the image function $f(x,y)$ is split into K intervals. The finer the sampling (larger M and N) and quantization (the larger K), the better the approximation of the continuous image function $f(x,y)$ achieved [2].

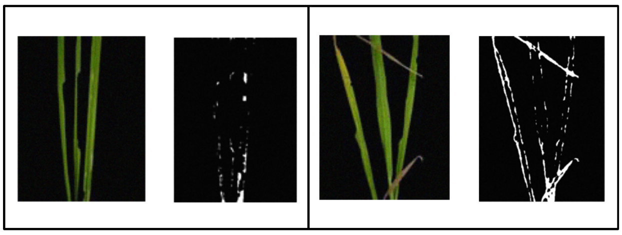

Fig. 2. Originals with Digitize images

C. Segmentation

Image segmentation is the first step in image analysis. The component of paddy disease image is complex, and the disease spot with the light or deep color was arranged on leaf disorderly. Various diseases had different spot color. As the image of paddy disease on leaf shown, disease leaf was composed by disease spot and natural part, so segmentation is a two class problem in essence.

Given discriminate function is linear, the normal formula is $g(x) = w^T x + w_0$, Where x is three dimensions feature vector, w is weight vector, respectively expressed by $x = [RGB]^T$, $w = [w_1, w_2, w_3]$; w_0 is threshold value.

The segmentation result is better, which can classify the disease spot on leaf from natural part basically. There are some lonely spot, stab and small empty hole in paddy disease image after segmentation. We used the *Sobel* method to detect the edges of the Image. This method finds edges using the Sobel approximation to the derivative.

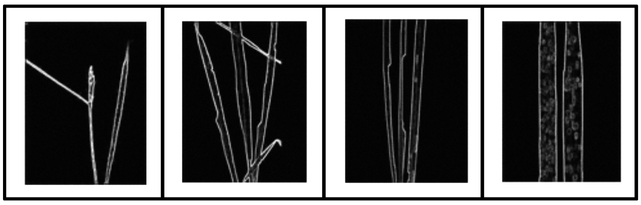

Fig. 3. Segmented images

D. Image Feature Extraction

Extraction of color texture

The repetition of a pattern or patterns over a region is called texture. This pattern may be repeated exactly, or as set or small variations. Texture has a conflictive random aspect: the size, shape, color, and orientation of the elements of the pattern. The main goal identifying different textures in image analysis is to replace them by a unique grey level or color [3].

The selection of texture feature of paddy disease for color image is based on chromatology concept of CIE XYZ color space [4]. In CIE XYZ color space, each pixel produces a pair of (x, y).

The number of edge pixels in a window can be found after applying an edge detector to that window [5]. Then, the density is calculated dividing the number of edge pixels found by the area of the window.

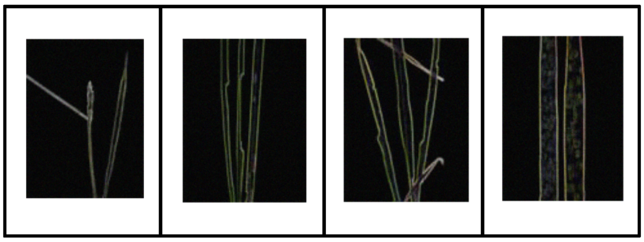

Fig. 4. Extraction of color texture

Extraction of color feature

Every spot in CIE $L^*a^*b^*$ color space can be seen as a spot in three dimension of CIE $L^*a^*b^*$. It can be derived from the CIE XYZ tri-stimulus values (x, y, z). The $L^*a^*b^*$ space consists of a luminosity 'L^*' or brightness layer, chromaticity layer 'a^*' indicating where color falls along the red-green axis, and chromaticity layer 'b^*' indicating where the color falls along the blue-yellow axis [6].

The differences of two kinds of color can be measured by Euclidean distance of them. The color can be compared directly in CIE $L^*a^*b^*$ color space, so it is effective in measure of small color difference. Hence the value of CIE $L^*a^*b^*$ in CIE $L^*a^*b^*$ color space is adopted as the color feature of disease recognition in this research. For example, if the distance between a pixel and the red color marker is the smallest, then the pixel would be labeled as a red pixel.

The color difference ΔE^*_{ab} between two colors, each expressed in terms of L^*, $a^*$and b^* is given by the Euclidean metric

$$\Delta E^*_{ab} = \left[(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2\right]^{\frac{1}{2}}$$

Where h^*_i is the hue of the i^{th} point and C^*_i is the chroma of the i^{th} point.

In the cylindrical representation, the Euclidean distance between two colors (L^*_1, h^*_1, C^*_1) and (L^*_2, h^*_2, C^*_2), is

$$\Delta E^*_{ab} = \left[(\Delta L^*)^2 + C^*_1 + C^*_2 - 2C^*_1 C^*_2 \cos(h^*_1 - h^*_1)\right]^{\frac{1}{2}}$$

Then we can find the distances between the pixel and each a^* and b^* separately (say D_a and D_b in respectively). If $D_a < D_b$ then we called the particular pixel falls along the *red-green* axis, otherwise the pixel falls along the *blue-yellow* axis.

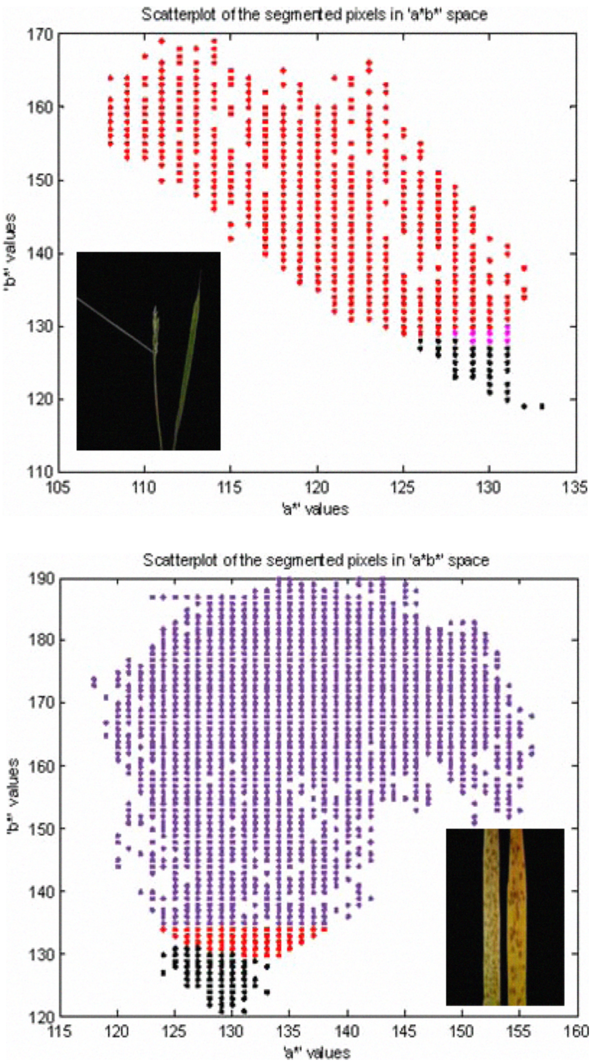

Fig. 5. Extraction charts of color feature in a^* b^* space

Extraction of shape feature on disease spot

Shape features were normally formed by area, roundness, shape complexity, extending length and concavity, equivalent rate of longer axis and shorter axis of ellipse and so on.

We filled any holes, so that region can be used to estimate the area enclosed by each of the boundaries. We concentrated only on the exterior boundaries. Then we estimated each spot's area and perimeter. Finally we used those results to form a simple metric indicating the roundness of the spot.

metric = (4*pi*area)/ (perimeter* perimeter)

This metric is equal to one only for a circle and it is less than one for any other shape.

III. CLASSIFICATION

In this research we defined membership function for each class, which are then used in a nearest neighbor classification. The format of the Membership Function (MF) is

MF (RGB range, shape, length, width, diameter . . .)

In Table I tabulated the Membership Functions for each type of disease. The membership function allows us to specifically define the criteria; a disease must meet to belong to a class.

TABLE I. MEMBERSHIP FUNCTIONS FOR EACH TYPE OF DISEASE

Disease Type	**Membership Function (MF)**
Rice blast	(RGB range, spots, [1.0 – 1.5], [0.3 – 0.5], [], …)
Rice sheath blight	(RGB range, oval/ ellipsoidal, [2.0 – 3.0], [1.0 – 2.0], [< 1.0], …)
Brown spot	(RGB range, circular, [0.4 – 1.0], [0.1 – 0.2], [0.05 – 0.1], …)

The class hierarchy was used to create the knowledge base for classification of diseases. The class hierarchy contained all classes used for discriminating the diseases and non disease in the analysis. These classes could be organized in a hierarchical structure. A disease(s) in the image was assigned to the class whose evaluation returns the highest membership value.

After the process of classification, each disease was assigned to certain (or no) class and thus connected with the class hierarchy. Then the system returns the appropriate disease type using above Membership Functions.

IV. RECOGNITION

Distribution of affected plants in a field is studied to understand if the problem is developing a pattern. Symptoms of affected plants hold the key to an accurate diagnosis of the field problem in question. Symptoms and signs of already recorded rice diseases have been very well described and recorded [7], and therefore one should use such available literature in relation to present field problem in question. In table II tabulated the signs and symptoms of each disease.

TABLE II. SIGNS AND SYMPTOMS FOR EACH TYPE OF DISEASE

Type of disease	Signs and Symptoms
Rice blast කොළ පාලුව 	Elliptical/ spindle whitish to gray Lesions wide in center and pointed to end Leaf collar, culms, panicle neck No seeds when neck is rotten
Rice sheath blight දිලීර කොපු අංගමාරය 	Water soaked to greenish gray Grayish white with brown margin Lesions near waterline on leaves
Brown spot දුඹුරු පුල්ලි 	Oval spots, dark brown dots Spots are small, circular Brown with gray/ whitish, blackish lesions

V. SYSTEM CONFIGURATION

The following equipments were used to develop and implement our system:

Hardware: Intel Pentium IV CPU with at least 1 GB RAM, SONY Cyber – shot 3.2 Mega Pixels: Digital still camera DSC-P5,

Software: The MathWorks MATLAB.

The format of the images that we captured is only JPEG (.jpg). Since the size of the JPEG images are very small, (relatively to GIF, PNG, TIF, BMP) it can be opened in the MATLAB in very quick. Therefore the system respond time is less.

VI. RESULTS

The GUI which is used by the User for the recognition during the process is shown in Fig. 6. The final outcome of the system is shown in Fig. 7. It gives brief introduction about that disease and the MF (properties) of the selected image.

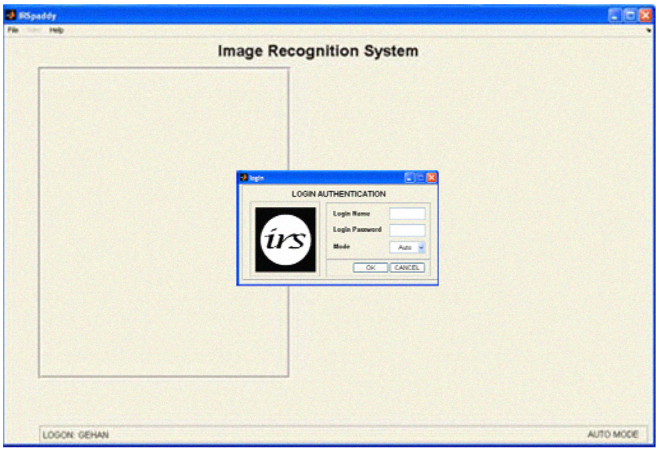

Fig. 6. Main Screen with Login Authentication

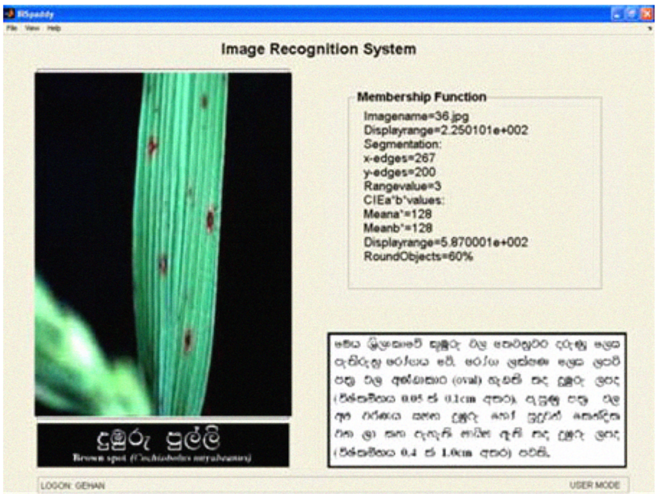

Fig. 7. Result of the System

About 50 images of three different types of diseases were used in this study. Most images were obtained from image resources and others were taken in the laboratory. According to results, the recognition of the diseases 'Rice blast' and 'Brown spot' images was more successful than other two types. Table III summarizes the performance of our system.

TABLE III. PERFORMANCE OF THE SYSTEM

Disease type	Number of captured Images	Success (%)
Rice blast (කොළ පාලුව)	20 80%	
Rice sheath blight (දිලීර කොපු අංගමාරය)	15 60%	
Brown spot (දුඹුරු පුල්ලි)	11 85%	

VII. DISCUSSION

The goal of the research is to recognize some of the paddy diseases in Sri Lanka using image recognition techniques. The results obtained on real data are comparatively satisfactory (more than 70% recognized). The rate of the recognition time is less than 2 seconds.

The images used in the system were 200×267 pixel dimension with 50 dpi. Some images of poor quality (poor contrast) gave the wrong result. To make the system more flexible the image resolution should be higher than the existing value. Also it is better to increase number of images in each type. If we have more number of images of each type then the result of the system may be in accurate than the present situation.

Also the segmentation algorithms have a common disadvantage to noise. Noise effect weakens almost all algorithms. This means most preprocessing to reduce noise effect must be required. Generally threshold is used as this preprocessing but it has a problem. This value should be decided so automated systems may confront this problem. If the threshold value is too high, it cannot classify pixels having small gray level difference. If the threshold value is too small, it becomes vulnerable to noise. It is very difficult to decide a proper threshold value.

The conclusion is that in case of reasonably good images, this approach yields good results. Use of powerful RGB camera would allow higher precision of the image color and segmentation.

ACKNOWLEDGMENT

We wish to thank to Dr. Nimal Disanayaka, Director of Rice Research and Development Institute at Batalagoda, Ibbagamuwa for supporting me to gain information about paddy plants in Sri Lanka. Our thanks extend to Professor Emeritus H. Sriyananda for his great insights and guidance for making this research and success.

REFERENCES

[1] "Rice Diseases", Department of Agriculture, Sri Lanka. 2006. Available: www.agridept.gov.lk

[2] Alberto Martin and Sabri Tosunoglu, "Image Processing Techniques for Machine Vision", Florida Conference on Recent Advances in Robotics, Florida Atlantic University, Boca Raton, Florida, 2000

[3] Ian T. Young, Jan J. Gerbrands and Lucas J. van Vliet, "Fundamentals of Image Processing", Version 2.2, Delft University of Technology, Netherlands, 1998.

[4] A. Hanbury and J. Serra, "Mathematical Morphology in the CIELAB Space - Image Analysis and Stereology", Vol. 21, No. 3, pages 201–206, 2002.

[5] Sonka, M., Hlavac, V., Boyle, R., Image Processing, Analysis, and Machine Vision, 2nd Edition, Pws. Pub. Co., 1998.

[6] Mark Bell, V Balasubramanian, Annie Antonio, "Rice Knowledge Bank - Rice Doctor", International Rice Research Institute (IRRI), Philippines, 2003.

[7] M. S. Swaminanathan, "Field Problems of Tropical Rice", Revised Ed., International Rice Research Institute (IRRI), Philippines, 1983.