논문 2011-48SP-3-10

# 벼 병충해분할을 위한 색채공간의 비교연구

# ( A Comparative Study of Different Color Space for Paddy Disease Segmentation )

앨롬엠디 자한기르\*, 이 효 종\*\*

(Alom MD Zahangir and Hyo Jong Lee)

#### Abstract

The recognition and classification of paddy rice disease are of major importance to the technical and economical aspect of agricultural industry over the world. Computer vision techniques are used to diagnose rice diseases and to efficiently manage crops. Segmentation of lesions is the most important technique to detect paddy rice disease early and accurately. A new Gaussian Mean (GM) method was proposed to segment paddy rice diseases in various color spaces. Different color spaces produced different results in segmenting paddy diseases. Thus, this empirical study was conducted with the motivation to determine which color space is best for segmentation of rice disease. It included five color spaces; NTSC, CIE, YCbCr, HSV and the normalized RGB(NRGB). The results showed that YCbCr was the best color space for optimal segmentation of the disease lesions with 98.0% of accuracy. Furthermore, the proposed method demonstrated that diseases lesions of paddy rice can be segmented automatically and robustly.

#### 요 약

전 세계적으로 벼 병충해의 인식과 분류는 농업현장에서 기술적 경제적으로 중요한 요소이다. 컴퓨터 비견 기술은 벼 병충 해를 진단하고 곡물의 효율적인 관리에 유용하다. 영역 분할은 벼 병충해를 조기에 정확하게 탐지하는데 매우 중요한 기술이 다. 가우시안 평균기법을 이용한 새로운 벼 병충해 분할 방식을 다양한 색체공간에서 제안하였다. 사용 색체공간에 따라 벼 병 충해의 분할에 따른 성능은 달라질 것이다. 따라서, 이 수치연구는 어느 색체공간이 벼 병충해를 분할하는데 최적한지를 결정 할 목적으로 수행되었다. 본 연구는 NTSC, CIE, YCbCr, HSV, 그리고 정규화 RGB의 5개의 색체공간을 다루었다. 연구 결과 는 YCbCr 색체공간이 98%의 정확도로 벼 병충해 영역을 최적으로 분할하는 것을 보여주었다. 또한 제안하는 방법은 벼 병충 해의 영역을 자동화에 의하여 강건하게 분할할 수 있다는 것을 증명하였다.

Keywords: Polar mean, circular features, Gaussian mean, rice disease

#### I. Introduction

Paddy rice diseases may cause agricultural industry millions of dollar in damage. Ever since humans started farming, diseases have been one of the major obstacles in maximizing production. There are numerous disease of paddy such as fungi, bacteria, viruses and nematode. Some diseases occur whenever paddy is grown. Some diseases reach epidemic proportion and cause serious crop losses where others cause only negligible crop losses. These diseases result in great economic losses to farmers every year<sup>[1]</sup>. It is important to note that disease symptoms of paddy rice diseases, such as paddy bacterial leaf blight, paddy sheath blight and paddy blast, appear initially as spots around the infected area. Hence, the detection of these diseases mainly

<sup>\*</sup> 학생회원, 전북대학교 전기전자컴퓨터공학부 (Dept. of Electric, Electronic and Computer Engineering)

 <sup>\*\*</sup> 평생회원-교신저자, 전북대학교 컴퓨터공학부, 영상정보신기술연구센터
 (Div. of Computer Science and Engineering, Chonbuk National University, CAIIT)

접수일자: 2010년12월27일, 수정완료일: 2011년3월2일

relies on their spots. At present, farmers or technicians mainly rely on their experiences, guide books and experts to identify these diseases. It is suitable for detecting some familiar and typical diseases. However, some different diseases can cause similar spots and the same disease can cause different spots because of different paddy varieties and local conditions. It increases the complexity of identifying paddy diseases. Misidentification usually leads to some inappropriate control measurements, such as indiscriminate and untimely use of pesticides

In agricultural areas, numerous image-processing based computerized tools have been developed to help farmers to monitor the proper growth of their crops. Some computer tools are used to discriminate between plants and other objects present in the field. However, most researches have been focused on the treatment and control of weeds the area of disease control. For detecting paddy diseases early and accurately, some researches have used image processing and classification techniques for detecting diseases of agricultural or economic  $\operatorname{crops}^{[2\sim 12]}$ . Nawala and Moratuwa<sup>[2]</sup> applied an image processing technique to detect crop diseases, but was weak in handling noise and difficult to find a proper threshold. General approaches of these researches had similar steps. First, disease images are acquired using cameras or scanners. Second, lesions are segmented from non lesions part of the leaf. Finally, performance analysis is conducted where different color space with respect to the various kinds of paddy disease shape or texture features are segmented. There are two major difficulties in rice disease detection researches. One is that there are not many researches actively progressed in this area due to the complexity of the task. The other is the performance of those methods are not acceptable.

In this study, an accurate and optimal threshold-based segmentation method was developed to segment the five types of paddy rice diseases including brown spot (BS), cold damage (CD), paddy leaf blast (PLB), paddy bacterial leaf blight (PBLB),

and paddy sheath blight (PSB). The rest of this paper is organized as follows. Related researches with the proposed work are described in Section II. Section III represents the proposed method for paddy disease segmentation for different color spaces. The proposed system results and discussion are discussed in Section IV. The conclusions of this work and about the future works are described in Section V.

# II. Related Works

The automatic paddy diseases analysis and recognition system is a developing task in the field of computer vision and machine learning. Agriculture is sometimes referenced as a weak theory domain, in which a large part of the reasoning knowledge is vague and described differently by various experts. Though the entry points in reviewing a case among different experts could not be the same, the conclusion should be similar. The precedence factors related to outcome from different researcher's viewpoint also varies. During that time some expert system already developed by researcher. But all the existing system requires different input parameters manually that is related to color, texture, shape and others environment related informations. We are presenting some expert system for tomato<sup>[13]</sup>, wheat<sup>[14]</sup>, lime<sup>[15]</sup>, cucumber<sup>[16]</sup>, and orange<sup>[17]</sup> have been built using the developed methodology related to the machine learning.

Nawala and Moratuwa proposed CIE color model based system for crop disease<sup>[2]</sup>. Finding a proper threshold value was the most difficult part in their method. If the threshold value is too high, it could not classify pixels having small gray level difference. In opposite cases the method became vulnerable to noise. Yao and Yang presented a system, which deployed shapes, colors and textures to detect paddy disease<sup>[3]</sup>. Otsu applied two threshold values to segment disease spots from the rice leaf<sup>[10]</sup>. RGB and HSV color models were used in Zhoo and Yao<sup>[4]</sup>. RGB color space is used for segmenting the rice blast spots. In their method 2G-R-B images and spots were segmented from leaves base on R-G channels. CIELAB color spaces was also used to detect brown spot<sup>[5]</sup>. Various color spaces were used by the several proposed system.

Our goal is to develope a new method which dose not depend on complicate input parameters. The concepts of such kind of system discussed in [2–6]. In this paper the best color space is found for developing an accurate and robust system for paddy diseases. We have considered five color space including YCbCr and YIQ/NTSC. A Gaussian mean based segmentation method using a single threshold is proposed. A procedure to calculate the automatic threshold value is described in next section.

## III. Proposed Method

1. Different Color Spaces and their Relationship As the segmentation of paddy rice diseases depends on color representation, it is important to understand the relationship among the different kinds of color space. Color spaces are different bases for representing intensity and color information in color images. Color spaces have multiple components or channels for representing all possible color and intensity information. For example RGB is a commonly used three dimensional color space with color channels of red, green, and blue. Different combinations of these three channels are capable to represent millions possible colors. In this experiments following five color spaces were considered including NTSC, CIE, YCbCr, HSV and NRGB (normalized RGB). The mathematical relationship of each of these color space was presented using RGB. Using the relationship of different color spaces to RGB, we can compute their relationship amongst each other such as tried in [13]. HSV describes colors in terms of the three quantities-hue, saturation, and value (HSV).

The NTSC color space system consists of three components: luminance(Y), hue(I), and saturation(Q), where the choice of the letters YIQ is conventional.

The relationship with the RGB color space is given Equation (1):

$$\begin{bmatrix} Y\\I\\Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114\\0.596 & -0.274 & -0.322\\0.211 & -0.523 & 0.312 \end{bmatrix} \begin{bmatrix} R\\G\\B \end{bmatrix}$$
(1)

The fundamental color space of the Commission internationale de l'Eclairage(CIE) and is the standard based on color matching experiments on human observers. It has a linear conversion from the RGB color.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7151 & 0.0721 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$
(2)

YCbCr color space used by most image compression standards like JPEG, H.261 and MPEG has the following relation with the RGB color space.

$$\begin{bmatrix} Y\\C_r\\C_b \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098\\0.439 & -0.368 & -0.071\\-0.148 & -0.291 & 0.439 \end{bmatrix} \begin{bmatrix} R\\G\\B \end{bmatrix} + \begin{bmatrix} 16\\128\\128 \end{bmatrix}$$
(3)

The mathematical relation between HSV and RGB is given by Equation (4)

$$H = \begin{cases} \theta & \text{if } B \le G\\ 360^{\circ} - \theta, & \text{otherwise} \end{cases}$$
(4)

where

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

$$S = 1 - \frac{3}{R + G + B} \left[ \min(R, G, B) \right]$$
(5)

$$V = \frac{1}{3}(R + G + B)$$
(6)

Relationship between HSV and RGB is non-linear demonstrated in [14].

Normalized RGB used by [13] to detect foreground objects and remove shadows is obtained as follows.

$$r = \frac{R}{R+G+B}$$

$$g = \frac{G}{R+G+B}$$

$$b = \frac{B}{R+G+B}$$
(7)

### 2. Gaussian Mixture Model

An image is a matrix which each element is a pixel. The value of the pixel is a number that shows intensity or color of the image. Gaussian Mixture Model (GMM) was applied to calculated the mean and deviation of paddy rice images in those five color spaces. For a probability model determination, we can suppose to have mixture of Gaussian distribution as the following form

$$f(x) = \sum_{i=1}^{k} p_i N\left(x/\mu_i, \sigma_i^2\right) \tag{8}$$

where k is the number of regions and  $p_i > 0$  are

weights such that  $\sum_{i=1}^{k} p_i = 1$ .

$$N(\mu_i, \sigma_i^2) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}}$$
(9)

where  $\mu_i$  and  $\sigma_i$  are mean and standard deviation. For a given image X, the lattice data are the values of pixels and GMM is applied to the pixel-based model. The proposed system calculated the mean and deviation in linearly except for HSV color space. We have calculated polar mean for converting the non linear hue pixels elements to linear form.

#### 3. Polar calculation of HSV color space

The polar mean has been calculated to convert the non-linear to linear system to the hue pixels of HSV images. The algebraic structure of the line and the circle are different and therefore adequate methods of circular data analysis and discussed in [22]. Contrast to the linear domain, only one operation of addition modulo  $2\pi$  is available in the circular domain. Due to the fact that the circle is a closed curve, its natural periodicity must be taken into account. As described in [21] a set of N angular estimates can be represented by N unit phases with arguments equal to the corresponding angular estimates. The mean angle  $\hat{\mu_p}$  is then given by the argument of the phasor sum and this value is independent of the

choice of an origin. The general definitions of the circular mean and variance based on this phasor sum are of the following from.

Definition: Circular sample mean and sample variance: Let  $\{\hat{a}(k), \hat{a}: Z | \rightarrow R\}$  be a set on N observation of a random variable in the circular domain [0, P). Then the circular sample mean  $\hat{\mu_p}$  and circular sample variance  $\hat{v_p}$  are defined by

$$\hat{\mu_p} = \frac{P}{2\pi} \left( \arg \left[ \sum_{k=0}^{N-1} e^{\frac{j2\pi a(k)}{P}} \right] \right)_{2\pi}$$
(10)

and

$$\hat{v_p} = \frac{P^2}{4\pi^2} \left[ 1 - \frac{1}{N} \left| \sum_{k=0}^{N-1} e^{\frac{j2\pi a(k)}{P}} \right| \right]$$
(11)

where ()<sub>2 $\pi$ </sub> denotes reduction modulo  $2\pi$  onto  $[0,2\pi)$ . The circular variance  $\hat{v_p}$ ,  $\hat{v_p} \in \left[0, \frac{P^2}{4\pi^2}\right]$  cannot be compared directly with its linear equivalent  $\sigma^2$  which lies in the domain  $[0,\infty)$ . However, by using the relationship between the normal distribution on the circle and the normal distribution on the line a circular standard deviation in the range  $[0,\infty)$  can be defined as

$$\hat{\sigma_p} = \sqrt{-2\log_n \left(1 - \frac{4\pi^2}{P^2} \, \hat{v_p}\right)} \tag{12}$$

Therefore, when using statistical definition in the context of hue values, we always refer to the above definitions from polar statistics. The calculation procedure of polar mean for hue of the propose system given below:

i) Take the pixel values for hue of

 $f_{hsv}(x_i, y_i), i = 1, 2, \dots, N$  and calculate the

$$\theta = f_{hsv}(x_i, y_i) \% 2\pi, \qquad (13)$$

where  $x = \cos(\theta)$  and  $y = \sin(\theta)$ .

- ii)  $\mu_x$ ,  $\mu_y$  are calculated from  $f_{hsv}(x_i, y_i)$ .
- iii) If data is dimensionless then

$$\hat{\mu_p} = \tan^{-1}(\mu_y, \mu_x) \% 2\pi$$
 (14)

Otherwise,

$$\widehat{\mu_p} = \tan^{-1}((\mu_y, d), (\mu_x, d)) \% \ 2\pi \tag{15}$$

where d refers data dimension. We have calculated  $\mu_{sv}$  and  $\sigma_{sv}$  from  $\mu_s$ ,  $\mu_v$ ,  $\sigma_s$  and  $\sigma_v$  for saturation and value pixels of the respective images. The polar mean  $\hat{\mu_p}$  for hue color pixels of the respective input image is calculated according the Equation (13)~(15). Now take image f(x,y) to generate the new image h(x,y) only for hue pixel as following:

$$h(x,y) = (f(x,y)^* 2\pi - \mu_p + \pi) \% 2\pi$$
(16)

Calculation of  $\mu_h$  and  $\sigma_h$  was carried in a range to fit Gaussian from h(x,y).

$$\mu_h = \left( \left( \mu - \pi + \mu_p \right) \% \ 2\pi \right) / 2\pi \tag{17}$$

$$\sigma_h = \sigma/2\pi \tag{18}$$

The values of  $\mu_h$  and  $\sigma_h$  are in linear format. Thus, the mean and deviation of hue are useful to calculate non-lesion and lesion means  $(\mu_{nl}, \mu_l)$  and their deviations  $(\sigma_{nl}, \sigma_l)$  using GMM for HSV color space.

#### Algorithm

A Gaussian model calculates non-lesion, lesion means  $(\mu_{nl}, \mu_l)$  and their deviations  $(\sigma_{nl}, \sigma_l)$  from all the sample pixels of the input image. The model sets a lower bound and an upper bound that will eliminate pixels that are outside of the norm.

1) To calculate Gaussian mean in f(x,y) of all color space for pixels we have used smaller data range  $[0.45 \ 0.55]$  to achieve faster calculation. Gaussian non-lesion and lesion means  $(\mu_{nl}, \mu_l)$  and their deviations  $(\sigma_{nl}, \sigma_l)$  were calculated for specific portion of data by using Gaussian Model for NTSC, CIE, YCbCr, and NRGB color spaces. We have used mean and deviation from Equation  $(17) \sim (18)$  to calculate for hue and calculated Gaussian non-lesion and lesion means  $(\mu_{nl}, \mu_l)$  and their deviations  $(\sigma_{nl}, \sigma_l)$  for specific portion of data by using Gaussian Model for HSV color space.

2) We compared with optimal threshold (T) to generate label of the lesions of the paddy leaf for all of the above mentioned color space. Calculation of an optimal threshold T is discussed in next section.

4. Optimal threshold based on Gaussian model Pixel deviation values that are higher than the threshold value considered as lesions, otherwise non-lesions. The input image contains some principle region and the distribution of pixel values in each region follows a Gaussian distribution.

The proposed system uses the following algorithm for selecting the optimal threshold for this system:

- 1) Find the histogram h(z) of the normalized image to be segmented.
- 2) Calculate the probability of a pixel value by the following mixture.

$$P(z) = P_{nl}P_{nl}(z) + P_{l}P_{l}(z)$$
(19)

where  $P_{nl}(z)$  and  $P_l(z)$  are probability distributions of non-lesions and lesions pixels.  $P_{nl}$  and  $P_l$  are priori or a posteriori probabilities of non-lesions and lesions pixels.

3) Calculate overall probability of error:

$$E(T) = P_{nl}E_{l}(T) + P_{l}E_{nl}(T)$$
(20)

where 
$$E_l(T) = \int_{-\infty}^{T} P_l(z) dz$$
 and  
 $E_{nl}(T) = \int_{T}^{\infty} P_{nl}(z) dz$ 

4) Find T to minimize  $E(T) \frac{dE(T)}{dT} = 0$ , such as  $T = \frac{\mu_{nl} + \mu_l}{2} + \frac{\sigma^2}{\mu_{nl} - \mu_l} ln(\frac{P_l}{P_{nl}}) when(\sigma_{nl} = \sigma_l = \sigma)$ , or  $T = \frac{\mu_{nl} + \mu_l}{2}$  when  $P_{nl} = P_l$  or  $\sigma = 0$  (21)

The T found is the optimal threshold value to 
$$T$$

The T found is the optimal threshold value to segment a lesion.

#### 1. Disease types

We have considered common paddy disease of the different areas over the world as experimental images. We have collected our experimental images from the Louisiana State University Auricular Center (www.lsuagcenter.com). We have considered paddy disease images with the environmental parts. The proposed system consists GM for segmentation of paddy diseases like BS, CD, PLB PBLB and PSB. The brief descriptions are given below respectively.

BS (Cochliobolus miyabeanus) : called as brown spot. Round to oval, dark-brown lesions with yellow or gold halo.

CD (Cold temperature) : called as cold damage. White band on leaf blade.

PLB (Pyricularia oryzae) : called as leaf blast. Round dark spots to oval spots.

PBLB (Xanthomonas oryzae pv. Oryzae) : called as bacterial leaf blight. Elongated lesions near the leaf tip or margin.

PSB (Rhizoctonia solani) : called as sheath blight. Wide bands of white, greenish-gray to tan with narrow bands.

#### 2. Segmentation results

To examine the performance of the proposed method, 40 images were cropped from various kinds of diseases. We have considered lesions and real part of leaf which contains normal condition. During the experimental analysis of cropped images we focused lesions rather than normal part of the paddy leaf. The experimental study and the performance analysis of our proposed method as follows:

(1) Images database: Database has been constructed with cropped images for our experimental study. Using this database we analyzed the performance of our propose system. The size of the image was 30x30 and five images for the individual disease. Fig. 1 shows samples of database images.

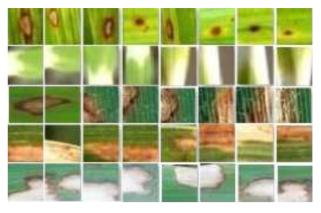


그림 1. 벼 병충해의 실험 영상 Fig. 1. Sample images of paddy disease.

(2) Accuracy calculation: Segmentation results from different color spaces were compared to each other for diseases in database. To measure the performance of various color spaces for paddy disease detection test, the concepts of sensitivity and specificity are employed. These concepts are readily usable for the evaluation of any binary classifier. The accuracy of segmentation can be defined with true positive (TP), true negative (TN), false positive (FP) and false negative (FN) as Equation (22).

$$accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$
(22)

We have calculated accuracy using average pixels of TP, TN, FP and FN for five individual disease images in terms of percentage. Table I represents the average pixels and accuracy of our proposed method with respect to different color spaces and diseases.

The proposed segmentation algorithm was able to segment the disease part of the paddy robustly and accurately. The performance of each color space for NTSC, CIE, YCbCr, HSV and NRGB are 88.2%, 91.1%, 98.0%, 94.3%, and 86.7%, respectively. The average performance of our proposed method with respect to color spaces is 91.7%. We have seen that the performance of YCbCr color space-based method accuracy was up to 98% for some specific cases, which is the highest performance among the all above mentioned color spaces.

Since Since NameTP pixelTN pixelFP pixelFN pixelSe pixelBS2571313230033NTNID1265312301313PLB2176611666697.60PBB28056245513333PBB28056245513333PBB28056245513333PBB10568862235383PBB1336384223799PBB256599378839PBB256599378899PBB256599378899YCC12077055598PBB21067077598PDB27060912199PBB27060912299PBB26060957289PBB28061213399PBB28062333139PBB2888343339PBB28684155339PBB28684155337PBB28684155337PBB285636169 <t< th=""><th></th><th></th><th></th><th>51</th><th></th><th></th><th></th></t<>				51			
NTNCICD126531230133I73.0NTNCPLB21766116697.6PBLB450404361094.9PSB2805624513393.6PSB2805624513393.6CO105688623588.1PLB183638423791.2PBLB450403351291.7PSB256599378895.0PSB256599378895.0PSB2767705598.6PLB420770598.6PLB2116777598.6PLB2116777598.7PSB2206091299.7PSB2706091294.9PSB2706091294.9PSB2706091294.9PSB2706091294.9PSB2706121393.2PSB2506121319.9PSB2506121313PSB2506121315PSB2506121315PSB2506121315PSB2506121315PSB2506121315PSB							Accuracy (%)
NTSCPILB217661166697.6PBLB4504043610094.9PSB28056245313393.6PSB280747942788.1CU105688623588.1PLB183638423791.2PBLB450403351294.7PSB256599378895.0PSB256599378897.6PSB256599378897.6PSB256599378897.6PSB256599378898.8PLB4207705598.8PLB2106777598.6PLB211677775598.7PSB2706091299.7PSB2706091291.9PSB2706091294.9PSB2706091293.0PSB2806121393.2PSB250612131393.2PSB250612231598.7PSB250612231594.7PSB250612231595.7PSB250612231595.7PSB250612231595.7 <td rowspan="5">NTSC</td> <td>BS</td> <td>25</td> <td>713</td> <td>132</td> <td>30</td> <td>82.0</td>	NTSC	BS	25	713	132	30	82.0
PBLB450404361094.9PSB28056245313393.6PSB28056245313388.1CU105688623588.1PLB183638423791.2PBLB45040335512294.7PBLB256599378891.2PSB256599378891.2PSB256599378891.2PSB256599378891.2PSB256599378891.2PSB256599378891.2PSB250700555598.6PLB21167775598.6PSB27060912290.5PSB27060912291.2PSB27060912292.9PSB27060912293.0PSB27060912292.9PSB2606122315393.7PSB2506122315393.7PSB286841553393.7PSB286841553394.9PSB286841553394.9PSB286841553394.9PSB2863078.911595.7 </td <td>CD</td> <td>126</td> <td>531</td> <td>230</td> <td>13</td> <td>73.0</td>		CD	126	531	230	13	73.0
PSB28056245133JPSB280562451388.R3274794278CD105688623588.PLB183638423791.2PBLB450403351294.7PSB256599378891.2PSB256599378897.6PSB256599378897.6CD1207705598.8PLB2116777598.6PBLB2106091297.6PSB2706091298PSB2706091298PSB43803421294.7PSB266609572898.7PSB250612231593.7PSB250612231593.7PSB250612231593.7PSB286841553394.7PSB286841553394.7PSB286841553394.7PSB286841553394.7PSB286841553394.7PSB286841553378.1PSB286347811578.5PS		PLB	217	661	16	6	97.6
No.         No.         No.         No.         No.           BS         32         747         94         27         88.1           CD         105         688         62         35         88.1           PLB         183         638         42         37         91.2           PBLB         450         403         35         12         91.2           PBLB         256         599         37         8         95.0           PSB         256         599         37         8         91.2           YGCb         120         770         5         5         98.8           YGCb         120         770         5         5         98.8           YGCb         120         770         5         9         98.8           PBLB         211         677         7         5         98.8           PBLB         470         404         15         111         9         97.6           PSB         270         609         12         9         98.0         98.0           HSY         ISS         433         803         422         12         94.0		PBLB	450	404	36	10	94.9
BS         32         747         94         27         2           CD         105         688         62         35         88.5           CD         105         688         62         35         91.2           PLB         183         638         42         37         91.2           PBLB         450         403         35         12         91.2           PBLB         256         599         37         8         95.0           PSB         256         599         37         8         91.2           PSB         256         599         37         8         91.0           CD         120         707         5         9         97.6           CD         120         770         5         9         98.6           PLB         211         677         7         5         98.6           PBLB         470         404         15         111         9         97.6           PSB         270         609         12         9         97.6         98.0         98.0           MSV         PLB         206         609         57 <td< td=""><td>PSB</td><td>280</td><td>562</td><td>45</td><td>13</td><td>93.6</td></td<>		PSB	280	562	45	13	93.6
CD105668866235881CD10566886223591.2PLB183638423791.2PBLB450403351294.7PSB25659937895.0PSB25659937891.2PSB25659937891.2CD1207705598.8PLB2116777598.6PBLB2106091297.6PSB27060912997.6PSB27060912997.6PSB27060912997.6PSB27060912997.6PSB27060912997.6PSB27060912997.6PSB27060912997.6PSB2706091299PSB266609572898.6PSB250612231593.7PSB250612231595.7PSB286841553379.1NR68PLB846237811578.5PSB226540924285.1PSB263583361894.0					ŀ	Average	= 88.1
CIEPLB1836384223791.2PBLB450403351294.7PSB25659937895.0PSB25659937891.1PSB4283712997.6CD12077055598.8PLB2116777598.6PBLB4704041551197.6PSB270609129997.6PSB27060912997.6PSB27060912998.0PSB27060912998.0PSB27060912998.0PSB27060912998.0PSB27060912998.0PSB20660912998.0PSB206612231598.0PSB250612231598.0PSB286841553379.1PSB286841553378.1PSB286540924285.1PSB266540924285.1PSB266540924285.1PSB266540924294.0	CIE	BS	32	747	94	27	86.5
PBLB450403351294.7PBLB25659937895.0PSB25659937891.0SUPSB25659937891.0SUPSB4283712997.6CD1207705598.8PLB21167775598.6PBLB47040415511197.1PSB270609122997.6PSB270609122997.6PSB270609122998.0CD13275012898.0PBLB428411481393.2PBLB428411481395.1PBLB226612231553395.1NRGBPLB846237811578.5PBLB22654092421295.1PSB263583361894.0		CD	105	688	62	35	88.1
PSB25659937895.0Average=91.1Average=91.1Average128371298CD1207705598.8PLB2116777598.6PBLB2116777598.6PBLB2106091299.6PBLB27060912997.6PSB27060912994.0CD13275012698.0PLB206609572890.5PBLB4284114813393.2PSB2506122315594.0PLB286841553379.1NRGEPLB846237811578.5PLB226540924285.1PSB236583361894.0		PLB	183	638	42	37	91.2
Average         =         91.1           BS         42         837         12         9         97.6           CD         120         770         5         5         98.8           PLB         211         677         7         5         98.6           PBB         470         404         15         11         97.6           PBB         470         404         15         11         97.6           PBB         270         609         12         9         97.6           PSB         270         609         12         9         97.6           PSB         270         609         12         9         97.6           CD         132         750         12         9         97.6           CD         132         750         12         6         98.0           CD         132         750         12         6         98.0           PBLB         428         411         48         13         93.2           PBL         428         612         23         15         95.7           NRGE         ES         28         684 <td< td=""><td>PBLB</td><td>450</td><td>403</td><td>35</td><td>12</td><td>94.7</td></td<>		PBLB	450	403	35	12	94.7
BS4283712997.6CD1207705598.8PLB2116777598.6PBLB470404151197.1PSB27060912997.6PSB27060912997.6CD13275012994.0CD13275012698.0PBLB206609572890.5PBLB206609572890.5PBLB206612231595.7PSB250612231595.7NRGBPLB286841553379.1PBLB226540924285.1PBLB226540924285.1PSB263583361894.0		PSB	256	599	37	8	95.0
CD         120         770         5         5         98.8           YGCb         PLB         211         677         7         5         98.6           PBLB         211         677         7         5         98.6           PBLB         470         404         15         11         97.1           PSB         270         609         12         9         97.6           PSB         270         609         12         9         97.6           CD         132         609         12         9         97.6           CD         132         750         12         6         98.0           RS         43         803         42         12         94.0           CD         132         750         12         6         98.0           MSV         PLB         206         609         57         28         90.5           PBLB         428         411         48         13         93.2           PSB         250         612         23         15         95.7           NRGE         BS         28         684         155         33 <t< td=""><td></td><td></td><td></td><td></td><td></td><td>Average</td><td>= 91.1</td></t<>						Average	= 91.1
YCCDPLB2116777598.6PBLB470404151197.1PSB27060912997.6PSB27060912998.0HSVES43803421294.0CD13275012698.0PLB206609572898.0PBLB428411481393.2PSB250612231595.7PSB250612231595.7NR68ES286841553379.1PLB8462378115785.1PBLB226540924285.1PSB263583361894.0	YCrCb	BS	42	837	12	9	97.6
PBLB470404151197.1PSB27060912997.6PSB27060912998.0UUUU132750421294.0CD13275012698.0PLB206609572890.5PBLB428411481393.2PSB2506122315595.7PSB286841553379.1CD14372921796.8NRGBPLB8462378115785.1PBLB226540924285.1PSB263583361894.0		CD	120	770	5	5	98.8
PSB         270         609         12         9           PSB         270         609         12         9         98.0           Image: Second Se		PLB	211	677	7	5	98.6
BS         43         803         42         12         94.0           CD         132         750         12         6         98.0           HSV         CD         132         750         12         6         98.0           PLB         206         609         57         28         90.5           PBLB         428         411         48         13         93.2           PSB         250         612         23         15         95.7           PSB         250         612         23         15         95.7           PSB         250         612         23         15         95.7           CD         143         729         21         7         96.8           NRGB         PLB         84         623         78         115         78.5           PBLB         226         540         92         42         85.1           PSB         263         583         36         18         94.0		PBLB	470	404	15	11	97.1
BS         43         803         42         12         94.0           CD         132         750         12         6         98.0           PLB         206         609         57         28         90.5           PBLB         428         411         48         13         93.2           PBLB         428         411         48         13         93.2           PBLB         250         612         23         155         95.7           PBLB         250         612         23         155         95.7           PSB         250         612         23         155         95.7           CD         143         729         21         75         96.8           NRGB         PLB         84         623         78         115         785.7           PBLB         226         540         92         42         85.1           PSB         263         583         36         18         94.0		PSB	270	609	12	9	97.6
CD         132         750         12         6         98.0           HSV         PLB         206         609         57         28         90.5           PBLB         428         411         48         13         93.2           PSB         250         612         23         15         93.2           PSB         250         612         23         15         93.2           Res         28         684         155         33         79.1           CD         143         729         21         7         96.8           NRGB         PLB         84         623         78         115         78.5           PLB         226         540         92         42         85.1           PSB         283         583         36         18         94.0	Average = 9						
HSV         PLB         206         609         57         28         90.5           PBLB         428         411         48         13         93.2           PSB         250         612         23         15         95.7           PSB         250         612         23         15         95.7           Average         =         94.3           PCD         143         729         21         7         96.8           PLB         84         623         78         115         78.5           PBLB         226         540         92         42         85.1           PSB         263         583         36         18         94.0	HSV	BS	43	803	42	12	94.0
PBLB         428         411         48         13         93.2           PSB         250         612         23         15         95.7           PSB         250         612         23         155         94.3           Verage         9         94.3         155         33         79.1           CD         143         729         21         7         96.8           PLB         84         623         78         115         785.5           PBLB         226         540         92         42         85.1           PSB         263         583         36         18         94.0		CD	132	750	12	6	98.0
PSB         250         612         23         15         95.7           Average         =         94.3           RRA         28         684         155         33         79.1           CD         143         729         21         7         96.8           PLB         84         623         78         115         78.5           PBLB         226         540         92         42         85.1           PSB         263         583         36         18         94.0		PLB	206	609	57	28	90.5
Average         =         94.3           BS         28         684         155         33         79.1           CD         143         729         21         7         96.8           PLB         84         623         78         115         78.5           PBLB         226         540         92         42         85.1           PSB         263         583         36         18         94.0		PBLB	428	411	48	13	93.2
BS         28         684         155         33         79.1           CD         143         729         21         7         96.8           PLB         84         623         78         115         78.5           PBLB         226         540         92         42         85.1           PSB         263         583         36         18         94.0		PSB	250	612	23	15	95.7
CD         143         729         21         7         96.8           PLB         84         623         78         115         78.5           PBLB         226         540         92         42         85.1           PSB         263         583         36         18         94.0						Average	= 94.3
NRGB         PLB         84         623         78         115         78.5           PBLB         226         540         92         42         85.1           PSB         263         583         36         18         94.0	NRGB	BS	28	684	155	33	79.1
PBLB         226         540         92         42         85.1           PSB         263         583         36         18         94.0		CD	143		21	7	96.8
PSB 263 583 36 18 94.0		PLB	84	623	78	115	78.5
		PBLB	226	540	92	42	85.1
Average = 86.7		PSB	263	583	36	18	94.0
						Average	= 86.7

표 1. 5종의 병충해 분할에 대한 각 컬러공간의 성능 Table 1. Segmentation performance of different color spaces for five types of diseases.

# 3. Comparison with other works

Due to the complexity of paddy disease images, the segmentation of disease lesions was not a trivial task. Furthermore, rice diseases are always mixed with even more complicated environmental images. Other researchers applied a linear function to segment the disease spot. Their main target was also limited to a fewer number of diseases. Although various methods were presented, their performance was not efficient. The recognition accuracy of previous systems was around 70% that is depending on their segmentation method and disease types, such as RRLB or RSB. However, a rice blast disease considered in Zhou and Yao was an exceptional case, which accuracy was 99.8%<sup>[4]</sup>. Our proposed method shows 91.64% average accuracy for five different types of paddy diseases. High accuracies of 98% and 95.7% were demonstrated for YCrCb and HSV color spaces, respectively. A proposed method may help other researchers to provide more accurate automatic disease analysis system in the sector of agriculture.

#### **V. CONCLUSIONS**

We have investigated five different color spaces of NTSC, CIE, YCbCr, HSV, and NRGB for their suitability in segmenting paddy rice diseases. A noble method of paddy disease segmentation method was presented. This work has empirically compared the suitability of different color spaces for segmentation of paddy disease. Every color space showed reasonably good performance from  $86 \sim 98\%$ . Among those five color spaces, the YCbCr color space demonstrated the best result and proved itself as a suitable solution for paddy disease segmentation robustly and accurately. The Gaussian mixture model was proved to be a reliable method to find automatic threshold values for segmentation of paddy diseases.

The proposed method not only identified the different types of paddy diseases, but it could also be suitable for finding other diseases. In the future by using this segmentation results we will be able to provide a paddy disease diagnosis system, which will be robust and accurate. The next step for this research is to design an effective algorithm to find out the lesions of paddy rice disease as well as analyze the performance with existing segmentation algorithms like Otsu<sup>[10]</sup>.

### REFERENCES

[1] M.L.Gullino, and J.P. Stack, Crop Biosecurity Netherlands: pp. 25–36,Springer 2008.

- [2] Nawala, and Moratuwa, "An Image Recognition System for Crop Disease Identification of Paddy fields in Sri Lanka," Fourth International Conference on Industrial and Information Systems, ICIIS 2009, Sri Lanka, 28–31 Dec. 2009.
- [3] Qing Yao, and Baojun Yang, "Application of support vector machine for detecting paddy disease using shape and color texture features," International Conference on Engineering Computation 2009. ICEC 2009.
- [4] Yingfeng Zhou, and Qing Yao, "Segmentation of Paddy Disease Spots Based on Improved BPNN," IASP10, pp.575–578, 2010.
- [5] Libo Liu, and Guomin Zhou, "Extraction of the Rice Leaf Disease Image Based on BP Neural Networks," International conference on. Issue ,pp. 1 - 3, China, 11-13 Dec. 2009.
- [6] Santanu Phadikar and Java Sil, "Rice Disease Identification using Pattern Recognition Techniques,"Proceedings of 11 International Conference on Computer and Information Technology (ICCIT 2008), Khulna, Bangladesh., 25-27 Dec, 2008.
- [7] D.Moshou, and H. Ramon, "Automatic detection of 'yellow rust' in wheat using reflectance measurements and neural networks," Comput. Electron. Agric., vol 44, no.3,pp.173–188, Sept. 2004.
- [8] P.M. Granitto, and H.A. Ceccatto, "Large-scale investigation of weed seed identification by machine vision," Comput. Etectron. Agric., vol. 47no.1, pp.15-24, Apr.2005.
- [9] R. Pydipati, and W.S. Lee, "Identification of citrus disease using color texture features and discriminant analysis," Comput. Electron. Agric, vol52,no.1, pp.49–59, June 2006.
- [10] N. Otsu, "A threshold selection method from gray-level histogram," IEEE Trans. CSAE, vol.23, no. 6, pp. 62–66, Jan. 1979.
- [11] R.M. Haralick, and I. Dinstein, "Texture features for image classification," IEEE Trans. Syst. Man Bybern., vol.3, no.6,pp.610–621, Nov. 1973.
- [12] K.Y. Huang, "Application of artificial neural network for detecting phalaenopsis seedling disease using color and texture features," Comput. Electron. Agric., vol.57, no.1, pp.3–11, May 2007.
- [13] EI-Shshtawy T.,A. and El-Az. E, "From Deperdence Networks to KADS: Implementation Issue," in Proceedings of the 2nd Workshop on AI in Agriculture, Netherlands, 1995.

- [14] Kamel A., and Sticklen J. "An Integrated Wheat Crop Management System Based on Generic Task Knowledge Based Systems and CERES Numerical Simulation," AI Applications, vol. 9, no. 1, 1995.
- [15] Refea M. and Rafea A., "LIMEX: An Integrated Multimedia Expert System," In Proceedings of the International Conference on Multimedia Modeling, Singapore, 1997.
- [16] Rafea A.,and Mahmoud M.,"Experience with the Development and Deployment of Expert Systems in Agriculture," in Proceeding of IAAI-95 Conference, Montreal-Canada, 1995.
- [17] Salah A., and Farahat H., "CITEX: An Expert System for Citrus Crop Management," in Proceedings of ESADW 93, MOALR, Cairo, Egypt,1993.
- [18] Pankaj Kumar, and Surendra Ranganath "A Comparative Study of Different Color space for Foreground and Shadow Detection for Traffic Monitoring System," IEEE 5th International Conference on Intelligent Transportation System, Singapore, September 2002.
- [19] A. Camargo and J.S. Smith, "An image-processing based algorithm to automatically identify plant disease visual symptoms," Biosystems Engineering, vol.102, pp.9-21, 2009.
- [20] Gonzales and R. E. Woods. "Digital Image Processing," Prentice Hall, 2002.
- [21] K.V. Mardia, "Statistics of directional data.," Academic Press, London, 1972.
- [22] B.C. Lovell, and R.C Williamson. "The circular nature of discrete-time frequency estimates," In IEEE International Conference on ASSP, pp. 3369–3372, Toronto, May 1991.
- [23] J. N. Kapur, and A.K.C. Wong, "A new method for gray-level picture threshodling using the entropy of the histogram," Computer Vision Graphics Image Process. Vol 29, pp.273–285, 1985.

액롬엠디 자한기르(학생회원) 2008년 방글라데쉬 University of Rajshahi 컴퓨터공학과 학사졸업. 2010년~현재 전북대학교 컴퓨터 공학과 석사과정 재학중.

<주관심분야 : Image processing, computer vision, communications>

- 저 자 소 개 -



 이 효 종(평생회원)-교신저자
 1986년 미국 유타대학교 컴퓨터 과학과 학사 졸업.
 1988년 미국 유타대학교 컴퓨터 과학과 석사 졸업.
 1991년 미국 유타대학교 컴퓨터 공학과 박사 졸업.

<주관심분야 : Image processing, computer vision, medical imaging, parallel processing>