An Image Processing Technique to Calculate Percentage of Disease Affected Pixels of Paddy Leaf

Rashedul Islam
Department of ICT
Rajuk Uttara Model College
Sector#06, Uttara, Dhaka-1230, Bangladesh

Md. Rafiqul Islam
Computer Science and Engineering Discipline
Khulna University
Khulna-9208, Bangladesh

ABSTRACT

Bangladesh is an agricultural country. Paddy is a principle crop of it. Millions of people depend for their living on paddy by way of its farming and processing. Blasts in paddy leaves are the most predominant disease which appears as brown spots on the leaves. If not treated on time, it may cause the great loss. Excessive use of pesticide for treatment of plant diseases increases the cost, environmental pollution and decreases the production. So their use must be optimized. This can be achieved by targeting the disease places, with the appropriate quantity and concentration of pesticide by estimating disease severity using image processing technique. In this paper K-means clustering method has been used to segment the image into three images based on color. Among these images unaffected leaf regions and disease affected regions are used to calculated percentage of affected pixels. By calculating percentage of affected pixels disease severity can be observed which leads to take appropriate measure for treatment.

General terms

Image processing, Affected pixels, Disease severity

Keywords

K-Means clustering, Segmentation, Binary image, Matlab.

1. INTRODUCTION

About 80% of population of Bangladesh depends on agriculture. Paddy is the principle crop of Bangladesh. In order to get quality production in optimum range needs technical knowledge of the farmers. It can be ensured by aid of technological support by the government. Close monitoring is required for the management of disease that can affect production significantly and subsequently the post-harvest life. The naked eye observation of experts is the main approach adopted in practice for detection of plant diseases [1]. However this requires continuous monitoring of experts which might be prohibitively expensive in large farms. Automatic detection of plant diseases is an essential research topic as it may give benefits in monitoring

large fields of crops and thus automatically detect the symptoms of diseases as soon as they appear on plant leaves. Therefore looking for fast, automatic, less expensive and accurate method to detect disease by calculating leaf area through pixel number statistics is an essential matter. The

image processing can be used in agricultural applications for following purposes [2]:

1. To detect diseased leaf, stem, fruit.

2. To quantify affected area by disease.

3. To find shape of affected area.

4. To determine color of affected area.

5. To determine size and shape of fruits.

Close observation of leaf area is an essential parameter in evaluating damage caused by leaf diseases and pastes. Leaf area plays an important role in plant growth analysis and photosynthesis. Leaf area is measured using different destructive and non-destructive methods. In destructive methods first the leaf is removed from plant and then measured. In non-destructive methods dimensions of leaf are measured without removing the leaf. Traditionally leaf area is measured by regression equitation, grid count method, gravimetric method and planimeter. Montgomery [3] suggested a formula for leaf area measurement which is, $A=b \times l \times w$. Where b is leaf shape coefficient, l is length of leaf and w is width of leaf. This method is non-destructive, quick and easier. But problem is that formula is not uniform for all plants as coefficient b differs among species.

In grid counting method [4] first the leaf is removed from plant and placed on a grid paper. Outlines of the leaf are drawn by pencil on grid paper. Finally leaf area is measured by counting grids covered by leaf. This method is accurate but laborious and time consuming when applied on large number of leaves. Gravimetric method is also used for leaf area measurement [5]. In this method leaf is first removed from plant and placed on white paper. Paper is cut-out according to the shape of leaf. Now weight of this paper is compared to the weight of known area on the same paper. This method is also laborious and time consuming when applied on large number of leaves. In this paper an image processing based algorithm is implemented to calculate percentage of disease affected pixels of paddy leaf, which leads to get severity of the disease. By understanding the severity of the disease the farmer can take appropriate measure to cure it. The hardware used is a digital camera, PC and a white paper sheet. The leaf is placed on white paper sheet and the image is taken using digital camera. The method is implemented in MATLAB 2012a.

2. RELATED WORK

In this section, a brief review of the literature is presented. S. B. Patil et al. presented the method of measuring leaf disease severity using segmentation and histogram thresholding [1]. V. A. Gulhane et al. presented color image feature extraction technique; this technique is called as color image segmentation [6]. Using this technique it is easy to extract the various features of diseased leaf of cotton image. S. B. Patil has described the process of measuring sugarcane leaf area using pixel number statistics [7]. It is a good approach. P. Chaudhary described the process of calculating leaf area using CIELAB color space [2]. The author has used a reference object to measure leaf area. H. Al-Hiary et al. have described the process of detection and classification of leaves disease by applying K-means clustering [8], Masking green-pixels and some other procedures and finally Configuring Neural Networks for Recognition. Authors of [9]

described the process of calculating leaf area The activities of the authors was focused on to calculate leaf area. Some researchers have used image processing techniques for fast and accurate measurement of leaf area [9,11-13]. Leaf area is measured using following steps [10]: image acquisition, image pre-processing, leaf region segmentation, region filling and area calculation. Some researchers use contour extraction technique [11-13] for leaf region segmentation and others use threshold based segmentation [9]. H. D. Marathe et al. have measured affected leaf area through pixel number statistics[2].Unit pixel in the same digital images represents the same size hence from known reference area and pixel count, unit pixel size can calculate. So that it is easy to calculate leaf area by counting total pixel in leaf area region. Our target is to calculate percentage of pixels of affected area from the leaf regions, which leads to take proper decision for treatment.

3. IMAGE SEGMENTATION AND CLUSTERING

Image segmentation is the process of dividing an image into multiple parts. This is typically used to identify objects or other relevant information in digital images. There are many different ways to perform image segmentation as follows:

- ✓ Thresholding methods such as Otsu's method.
- ✓ Color-based Segmentation such as K-means clustering.
- ✓ Transform methods such as watershed segmentation.
- ✓ Texture methods such as texture filters.

An effective approach to performing image segmentation includes using algorithms, tools, and a comprehensive environment for data analysis, visualization, and algorithm development

3.1 Necessities of segmentation

Segmentation of digital image is very important to work with it. Different types of information can be obtained from the image if it is segmented in proper way. In this paper we have segmented the original image into two basic parts. One contains unaffected leaf regions and the other contains affected leaf regions. So, we can say proper segmentation is very much essential for image processing.

3.2 Clustering

Image clustering and categorization are means for high-level description of image content. The goal is to find a mapping of the archive images into classes (clusters) such that the set of classes provide essentially the same information about the image archive as the entire image-set collection. The generated classes provide a concise summarization and visualization of the image content that can be used for different tasks related to image database management. Image clustering enables the implementation of efficient retrieval algorithms and the creation of a user-friendly interface to the database.

A common approach to image clustering involves addressing the following issues:

- ✓ Image features – how to represent the image.
- ✓ Organization of feature data – how to organize the data.
- ✓ Classifier – how to classify an image to a certain cluster.

3.2.1 K-Means Clustering

Clustering is a method to separate group of objects. In the proposed method we have used K-means clustering in segmentation process. In this regard, K-means clustering method has been described here. K-means clustering is one of the most used clustering methods. It considers each object as having a location in space. In this method an image is partitioned into K mutually exclusive clusters and the index of the target cluster is returned by the function. It creates partition such that objects of each cluster are as close to each other as possible. K-means clustering works on actual observation and produces a single level of clusters. It is more suitable than hierarchical clustering for large amount of data [14]. It requires that we specify the number of clusters to be partitioned and a distance metric to quantify how close two objects are to each other.

K-means algorithm [15]:

The K-means algorithm is used to solve the clustering problem and works as follows. First, decide the number of clusters then:

1. Initialize the center of the clusters

2. Attribute the closest cluster to each data point

3. Set the position of each cluster to the mean of all data points belonging to that cluster

4. Repeat steps 2-3 until convergence

Since the color information exists in the 'a*b*' space, the objects are pixels with 'a*' and 'b*' values. K-means is used to cluster the objects into three clusters using the Euclidean distance metric [14]

Results of k-means are shown by the following figures (Figure 2 to Figure 5).

Figure 1: Original image

Figure 2: Image labeled by cluster index

After segmented the image of Figure 1 by color, following images are obtained.

Figure 3: Object in cluster 1

Figure 4: Object in cluster 2

Figure 5: Object in cluster 3

Within the above three images, Figure 2 shows affected area of the leaf and Figure 5 shows green area(unaffected area) of the leaf. So these two images can be used to count affected and unaffected pixels.

4. PROPOSED APPROACH

Main objective of the proposed method is to calculate percentage of disease affected pixels of a paddy leaf. To do this, a digital camera has been used to capture images of the leaves. But in field condition, it is difficult to keep optical axis of camera vertical with leaf plane. So the leaf is separated and placed horizontally on white background to take the photograph. After that, the overall process is accomplished in two phases.

Phase 1: This phase has three steps:

i) Image acquisition

ii) Read image

iii) Segment the image using k-means clustering

Phase 2: This phase is divided into three steps.

i) Input affected image (obtained from segmentation process).

ii) Convert it into binary image,

iii) Compute the percentage of the affected pixels, P_A as follows:

$$P_L = wp_a + wp_u \qquad (1)$$

$$P_A = \frac{wp_a}{P_L} \times 100\% \qquad (2)$$

Where wp_a is the total number of white pixels in affected region of the leaf, wp_u is the total number of pixel in unaffected regions of the leaf, P_L denotes total pixels of the leaf area.

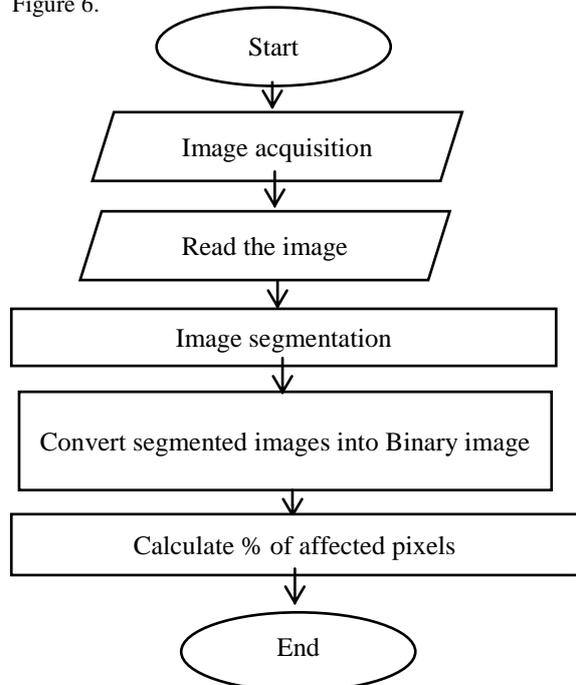The process has been depicted in the flowchart as shown in Figure 6.

Figure 6: Flowchart of the proposed approach

5. EXPERIMENTAL RESULTS
5.1 Phase-1:
5.1.1 Image acquisition
The leaf whose infected area is to be measured is placed on the white background, without any light reflection. The camera is held horizontally to the plane of the leaf. The photograph distance is neither too close nor too far; it is adjusted such a way that the photograph is covering only background.

Figure 7: Original image of paddy leaf

5.1.2 Reading the image

The image is saved in the computer as 'paddyblast.png'. This image is read for further processing by MATLAB.

5.1.3 Segment the Image

In this step, original image is segmented into 3 types of images according to variation of color. From these images two images have been selected that identified affected regions of the leaf and unaffected regions of the leaf. It is done by K-means clustering method. The algorithm of this method is given below:

Step 1: Read Image

Step 2: Convert image from RGB color space to LAB color space.

Step 3: Classify the colors in 'a*b*' space using K-means clustering

Step 4: Label every pixel in the image using the results from K-means

Step 5: Create images that segment the original image (Figure 7) by color.

After applying the above algorithm following three segmented images (Figure 8 to 10) were obtained.

Figure 8: Leaf regions
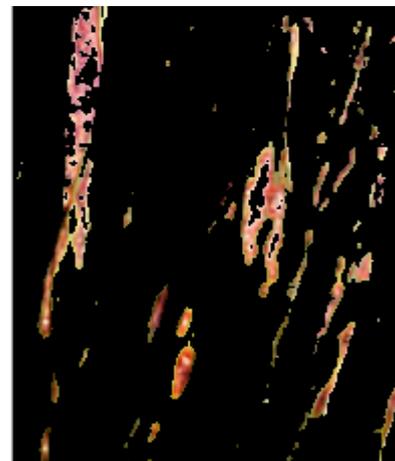

Figure 9: Unaffected leaf regions

Figure 10: Disease affected regions

From the above 3 images, Figure 9(Unaffected leaf regions) and Figure 10 (Disease affected regions) have been selected to calculate percentage of affected pixels.

5.2 Phase-2:

5.2.1 Read the image (Figure 9)

This image indicates unaffected regions as foreground image. So, at first this image has to read to count total number of pixels of unaffected regions.

5.2.2 Convert the image into Binary image.

After reading the image at first this image must be converted from RGB to Gray scale image then the gray scale image will be converted into binary image. In this image white colored regions indicates unaffected part of the leaf as shown in the following figure (Figure 11.). It helps to calculate total number of pixels of unaffected regions of the leaf.

Figure 11: Binary image of Figure 9.

5.2.3 Read and Convert the image shown in Figure 10 into Binary image.

After converting the image of Figure 10into binary image, following image (Figure 12.) is obtained where white colored regions indicate affected regions of the leaf. It helps to calculate total number of pixels of affected regions of the leaf.

Figure 12: Binary image of Figure 10.

5.2.4 Pixel Calculation

i) From Figure 11.

WP_u =28565

BP_1=18,527

Total Pixels, PL=47,092

Where, BP denotes black pixels.

ii) From Figure 12

WP_a=3614

BP_2=43,478

Total pixel, *PL*=47,092

Where, *BP* means black pixels.

Now according to equation (1), total pixels of the leaf area that is obtained is shown below:

P_L=3614+28,565

$\qquad$ =32,179

By using the equation (2), Percentage of affected pixels can be obtained by the following way:

P_A= 3614/32179*100%

$\quad$ = 11.23%

So, P_A=11.23%

6. RESULTS AND DISCUSSION

The results obtained from the above operation show that about **11.23%** of the original image is infected by disease. The calculation of counting pixels of the image is is 100% accurate. Because after counting black and white pixels, total number of pixels, that has got is =**47,092**.

Table 1: Percentage of infection and corresponding measures [16].

% of infection	Measures
Less than 10%	Close observation and flooding the whole field with huge water.
10% to 15%	i. Adjust planting time. Sow seeds early, when possible, after the onset of the rainy season. ii. Split nitrogen fertilizer application in two or more treatments. Excessive use of fertilizer can increase blast intensity.
15% to 25%	Silicon fertilizers (e.g., calcium silicate) can be applied to soils that are silicon deficient to reduce blast. However, because of its high cost, silicon should be applied efficiently
More than 25%	Systemic fungicides like triazoles and strobilurins can be used judiciously for control to control blast. A fungicide application at heading can be effective in controlling the disease.

Total number of pixels of the original binary image of resolution 193×244 is also equal to **47,092**(193*244=47,092).

Following Table 1 shows how to manage the disease based on obtained results.

Now, the experimental result on different images will be shown using the proposed method (PM) as well as the Standard method (SM), which is implemented using MATLAB function bwarea [17]. The method is as follows:

i. Segment the original image into two images. One represents affected leaf area and the other represents unaffected leaf area.

ii. Convert the above two images into binary image

ii. Calculate total area of the objects of converted binary image in pixels.

In this method total number of white pixels is treated as the pixels of the object. As for example
total = bwarea(BW)

Where BW is the binary image of the original image and total is the total number of white pixels of the objects of the binary image.

Table 2: Experimental results using two methods.

Sl. No	Sample Image	% of infection (SM)	% of infection (PM)	Deviation
1	Figure A-1	12.5307	11.2309	1.2998
2	Figure A-2	48.6548	48.6531	0.0017
3	Figure A-3	18.8208	18.6593	0.1615
4	Figure A-4	3.2042	2.0311	1.1731
5	Figure A-5	3.2440	3.1640	0.0800
6	Figure A-6	3.4647	3.4014	0.0633
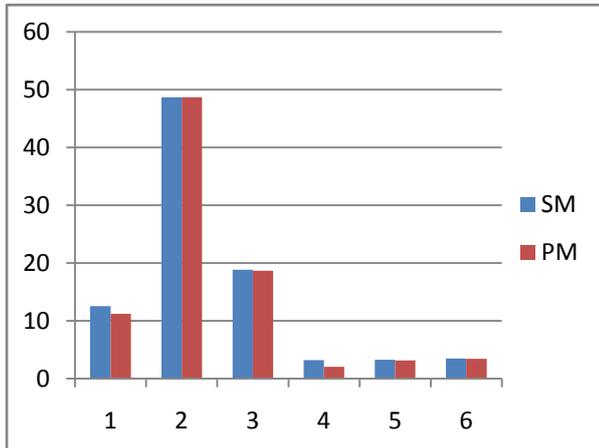
Figure 13: Graphical Analysis on % of infection using Standard Method(SM) and Proposed Method (PM)

7.   CONCLUSION

Image processing technique is very much essential to observe the intensity of disease. As the open eye observation may result poor accuracy and it may vary person to person. In this paper with the help of K-means clustering, image of paddy leaf has been segmented. By counting the disease affected pixels and unaffected pixels the percentage of disease affected pixels has been calculated. According the percentage of disease affected pixels the severity of disease can be understood; consequently appropriate measure can be taken to cure the disease.

In future a common system can be designed using the proposed approach to measure disease severity of the leaves of other crops that will be helpful for the farmers to produce better quality crops.

8.   REFERENCES

[1] Sanjay B. Patil, ShrikantK.Bodhe, Leaf Disease Severity Measurement using Image Processing, International Journal of Engineering and Technology Vol.3 (5), 2011.

[2] Hrushikesh Dattatray Marathe, Prerna Namdeorao Kothe, Leaf Disease Detection Using Image Processing Techniques, International Journal of Engineering Research & Technology (IJERT), Vol. 2 Issue 3, March – 2013.

[3] Montgomery EG, Correlation studies in corn. Nebraska, AgrExpStaAnnu Rep 24, pp. 108–159, 1911.

[4] Ross, J., Ross, V., Phytometrical characteristics of the willow plantation at Toravere", joint Swedish–Estonian seminar,

[5] 24–26Linda Styer Caldas, Carlos Bravo, Homero Piccolo, Carmen Rachel S. M. Faria,   Measurement of Leaf Area With a Hand-Scanner Linked to a Microcomputer, R. Bras. Fisiol. Veg. 4(1), pp. 17-20.

[6] V. A. Gulhane and A. A. Gurjar, Detection of Diseases on Cotton Leaves and Its PossibleDiagnosis, International Journal of Image Processing (IJIP), Volume (5) : Issue (5) : 2011.

[7] Sanjay B. Patil, Image Processing Method To Measure Sugarcane Leaf Area, International Journal of Engineering Science and Technology, ISSN : 0975-5462 Vol. 3 No. 8 August 2011.

[8] H. Al-Hiary, S. Bani-Ahmad, M. Reyalat, M. Braik and Z. ALRahamneh, Fast And Accurate Detection and Classification of Plant diseases, International Journal of Computer Applications (0975 –8887),Volume 17– No.1, March 2011.

[9] Sanjay B. Patil, S.K.Bodhe,?Betel leaf area measurement using image processing,? IJCSE, 2011, pp. 2856-2660.

[10] Piyush Chaudhary, Sharda Godara, A. N. Cheeran, Anand K. Chaudhar, Fast and Accurate Method for Leaf Area Measurement, ,International Journal of Computer Applications (0975 – 8887,) Volume 49– No.9, July 2012.

[11] ChaohuiLü, HuiRen, Yibin Zhang, YinhuaShen, Leaf Area Measurement Based on Image Processing, IEEE International Conference on Measuring Technology and Mechatronics Automation, pp. 580-582, 2010.

[12] Hassan Hajjdiab, AbdellatifObaid," A Vision-based Approach for Nondestructive Leaf Area estimation", 2nd Conference on Environmental Science and Information Application Technology, IEEE, pp.53-56, 2010.

[13] Tan Feng, Wang Chun, Calculating the Leaf-Area Based on Non-Loss Correction Algorithm, International Conference of Information Science and Management Engineering, IEEE, pp.75-78, 2010.

[14] k-means clustering, Available at: http://www.mathworks.com/help/stats/kmeans.html.(Accessed on 25April 2015)

[15] K-means clustering, Available at: http://www.onmyphd.com/?p=k-means.clustering

[16] (Accessed on 22 May 2015)   Blast (leaf and collar) - IRRI Rice Knowledge Bank, Available at: http://www.knowledgebank.irri.org/training/factsheets/pest management/diseases/item/blast-leaf-collar. (Accessed on 25 April 2015)

[17] bwarea, Area of objects in binary image, Available at: http://www.mathworks.com/help/images/ref/bwarea.html (Accessed on 22 May 2015)

9.   AUTHOR PROFILE

Rashedul Islam received B.Sc. degree in Computer Science and Engineering (CSE) from Khulna University, Khulna, Bangladesh in 2002 and received M.Sc. degree in Computer Science and Engineering (CSE) from Uttara University, Dhaka, Bangladesh in 2011. Currently he is working as an Assistant Professor under the department of Information and Communication Technology (ICT) of Rajuk Uttara Model College, Uttara, Dhaka, Bangladesh. He is the author of the

book *Higher Secondary Information and communication Technology*-for class XI-XII (English version).He has one publication in the proceedings of 5th International Conference on Computer and Information Technology (ICCIT), 2002, Dhaka, Bangladesh. His present research interest includes Image processing, Artificial Neural Network, Bio-metrics etc.

Md. Rafiqul Islam obtained Ph.D. in Computer Science from Universiti Teknologi Malaysia (UTM) in 1999 and a combined Master (MS) and Bachelor Degree in Engineering (Computers) from Azerbaijan Polytechnic Institute (Azerbaijan Technical University at present) in 1987. He was a visiting fellow (a postdoctoral researcher) in Japan Advance Institute of Science and Technology (JAIST) in 2001. He worked as head of the Discipline of Computer Science and Engineering of Khulna University and as the Dean of the School of Science, Engineering and Technology of Khulna University. He worked as a Professor in the Department of Computer Science of American International University-Bangladesh (AIUB) from 2009 to 2015. Currently he is a senior professor of Computer Science and Engineering Discipline of Khulna University, Khulna, Bangladesh. He has 26 years of teaching and research experiences. He has supervised more than 40 undergraduate and MS theses. He was external examiner of MS and Ph.D. theses of several universities of Bangladesh. He has published more than 88 papers, which have been published in international and national journals as well as in referred international conference proceedings published by IEEE, Springer and others. He has reviewed papers submitted to International journals of IEEE, Wiley, Springer, papers submitted to journals of different universities of Bangladesh and papers submitted to several International Conferences. He is a current reviewer of IEEE system journal and journal of Security and Communication Networks (SCN) of Willey. His research areas include design and analysis of algorithms in the area of image processing, secure cloud computing, external sorting, Information security, Network security data compression, bio-informatics, grid computing and data mining etc.

10. APPENDIX

Some sample images for experiment

Sample Images		
Figure A-1
Figure A-2
Figure A-3
Figure A-4 |
Figure A-5 |
Figure A-6 |