Performance Evaluation of Segmentation of Frog-Eye Spot Lesions on Tobacco Seedling Leaves

P.B. Mallikarjuna and D.S. Guru

Department of Studies in Computer Science, Manasagangothri,
University of Mysore, Mysore-570006, Karnataka, India
pbmalli@yahoo.com, dsg@compsci.uni-mysore.ac.in

Abstract. In this paper, a new algorithm for segmentation of frog-eye spot lesions on tobacco seedling leaves is proposed. Segmentation algorithm consists of mainly two steps. First step is to approximate lesion extraction using contrast stretching transformation and morphological operations such as erosion and dilation. Second step refines the outcome of first step by color segmentation using CIELAB color model. We have also conducted a performance evaluation of segmentation algorithm by measuring the parameters such as Measure of overlapping (MOL), Measure of under-segmentation (MUS), Measure of over-segmentation (MOS), Dice similarity measure (DSM), Error-rate (ER), Precision (P) and Recall (R). In order to corroborate the efficacy of the proposed segmentation algorithm, an experimentation is conducted on our own dataset of 400 segmented areas of tobacco seedling leaves which are captured in uncontrolled lighting conditions. Experimental results show that our proposed segmentation algorithm achieved best average DSM and MOL accuracy as compared to our previous segmentation algorithm.

Keywords: Image enhancement, CIELAB color model, Lesion area segmentation, Performance measures.

1 Introduction

Agriculture plays an important role in economy of any nation. Economy of agricultural industries is directly depends on the quality production of agriculture. A stable agricultural industry ensures a country of food security, source of income, source of employment. Therefore, to improve agriculture production and its quality, farmers should practice precision agriculture.

Precision agriculture focuses on getting maximum quality output with minimum input. The objectives of precision agriculture are profit maximization, agriculture input rationalization and environmental damage reduction, by adjusting the agriculture practices to the site demands. To achieve these objectives some practices which are site specific application of agrochemicals to remove diseases at seedling (nursery) level and plant level, right time harvesting of crops and grading (quality inspection) of crops are to be adopted. Human intervention in these practices raises many disadvantages such as wrong diagnosis of diseases in crops, wrong quality

N. Meghanathan et al. (Eds.): CCSIT 2012, Part II, LNICST 85, pp. 444–453, 2012.

analysis of crops, man power, labor cost and time consuming. Therefore, we need to automate these practices to increase efficiency and speed using computer vision (CV) algorithmic models.

Diagnosis and detection of diseases at nursery stage is very important in any crop. The emergence and spreading of frog-eye spot lesions have become more common in tobacco at nursery (seedling) phase because of climate and environmental factors. The symptoms of frog-eye spots on a leaf are characterized by small spots of size 2mm with circular central grey or white dead tissue as shown in Fig 2(a). Apart from our work [1] no other attempt can be traced on disease detection on leaf of tobacco seedlings. However few attempts can be traced on other crops. Enhancing color differences in images by means of vectorial normalization was proposed to better separation of diseases [2]. Conversion of a RGB image into H, I3a, and I3b color transformation and segmenting the transformed image by analyzing the distribution of intensities in a histogram was explored to identify plant disease visual symptoms [2]. Color co-occurrence method (CCM) was used in conjunction with statistical classification algorithms to identify diseased and normal leaves under laboratory conditions [4]. An algorithm to classify fall armyworm damaged maize plants and an undamaged maize plant at simplified lighting conditions has also been recommended [5]. A back propagation neural network (BPNN) and a gray level co-occurrence matrix (GLCM) were used to evaluate the texture features of the lesion area in seedling diseases [6]. Fuzzy feature selection techniques are proposed for identifying diseases on cotton leaves [7]. Grading method of leaf spot disease on soya bean leaf using image processing techniques was proposed [8]. A method of recognizing disease in a cucumber leaf based on image processing and support vector machine was developed [9]. Digital image analysis and spectral reflectance data are used to quantify damage by greenbugs in wheat crop [10]. Severity of fungal disease in a spring wheat crop was estimated using hyperspectral crop reflectance data vectors and corresponding disease severity field assessments [11]. Fuzzy feature selection approach was proposed to diagnose and identify diseases in cotton crop. This approach reduces the dimensionality of the feature space so that it leads to a simplified classification scheme [12]. A method of using wavelet transform was developed to detect pests in stored grains [13]. An image processing algorithm was proposed for automatic identification of whiteflies, aphids and thrips in greenhouse. The size and color components were selected as features for automatic identification [14].

In our previous work [1], we segment the lesions on leaf of tobacco seedlings and classify them in to three classes: Anthracnose, Frog-eye spot and Uninfected area. In our current work we improved the accuracy in segmenting the Frog-eye spot lesions. We compared our current segmentation algorithm with previous segmentation algorithm by measuring the performance evaluation measures such as Dice coefficient (DC), Error-rate (ER), Measure of overlapping (MOL), Measure of under-segmentation (MUS), Measure of over-segmentation (MOS), Precision (P) and Recall (R).

2 Proposed Segmentation Algorithm

The proposed segmentation algorithm consists of two steps. The first step involves approximate lesion extraction using contrast stretching transformation and morphological operations such as erosion and dilation. Second step refines the outcome of first step by color segmentation using CIELAB color model.

We propose the following method to segment the lesion area. The RGB image (Fig. 1(a)) of tobacco seedling leaf is transformed into a B-channel gray scale image (Fig. 1(b)). The gray scale image is enhanced using a contrast stretching transformation [15] with an adjustable parameter (m) and is given by

$$G = f(g) = 1/[1+(m/g)^E] \tag{1}$$

where,

g *represents the intensities of the input image*

m *represents threshold*

E *controls the slope of the function*

The contrast stretching transformation compresses the values greater than m into a narrow range of dark levels in the output image G, similarly it compresses the values less than m into a narrow band of light levels in the output image G. The enhanced gray scale image (Fig. 1(c)) is transformed into binary image using threshold T_1. The erosion and dilation operations using a disk structuring element of radius one is applied on the obtained binary image (Fig. 1(d)) to remove uninfected areas. A suitable threshold T_2 is used to remove small uninfected areas left in Fig 1(e), i.e., the areas which are less than T_2 pixels are removed as shown in Fig. 1(f). However, selecting a suitable threshold is a challenging task. If small threshold values are selected there are chances of retaining uninfected areas of small size and if large threshold values are selected there are chances of eliminating the lesion areas. Hence, fixing up a suitable threshold T_2 such that there are less probability for lesion areas to be missing. Fig 2 shows the extracted lesion areas of Fig. 1(a). While extracting lesion areas we have considered k pixels around the seed point of lesions. However, selecting a suitable k value is a challenging task. If small k value is selected there are chances of losing lesion information and if large k value is selected there are chances of extracting adjacent lesions information.

From Fig 2 it is observed that extracted lesions include lesion area and also healthy area. Therefore CIELAB color model is used to segment the lesion area from the healthy area. CIELAB is an approximately uniform color system. Its values are calculated by non-linear transformations of CIE XYZ. In this system, Y represents the brightness (or luminance) of the color, while X and Z are virtual (or not physically realizable) components of the primary spectra. The CIE XYZ tristimuli are standardized with values corresponding to the D65 white point: $X_0 = 95.047$, $Y_0 = 100$, $Z_0 = 108.883$. It is then transformed into the standardized tristimuli to the CIELAB Cartesian coordinate system using the following metric lightness function.

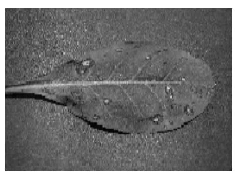
$$L = \begin{bmatrix} 166\times(Y/Y_0)^{1/3} - 16 & for\,(Y/Y_0)^{1/3} > 0.00856 \\ 903.3\times(Y/Y_0) & otherwiswe \end{bmatrix} \quad (2)$$

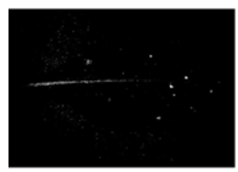
The chromacity coordinates a* and b* are derived using:

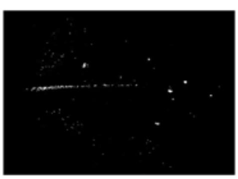
$$\begin{aligned} a^* &= 500\times[(X/X_0)^{1/3} - (Y/Y_0)^{1/3}] \\ b^* &= 200\times[(Y/Y_0)^{1/3} - (Z/Z_0)^{1/3}] \end{aligned} \quad (3)$$

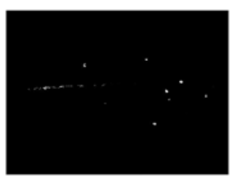
The chromacity coordinates represent opponent red-green scales ($+a$ red, $-a$ greens) and opponent blue-yellow scales ($+b$ yellows, $-b$ blues). Since color of the healthy area in the extracted lesions varies from light green to dark green, the chromacity coordinate a^* is used to segment lesion from the healthy area. The a^* value is calculated for each extracted lesion image pixel. If the value of a^* is greater than a predefined threshold then the corresponding pixel is considered as a lesion pixel else it is considered as a healthy pixel. Segmented Frog-eye spot lesion areas of Fig 2 using CIELAB color model are shown in Fig 3.


1 (a)


1(b)


1(c)


1(d)


1(e)


1(f)

Fig. 1 (a)-1(f). Segmentation of frog-eye spot lesion areas on a tobacco seedling leaf

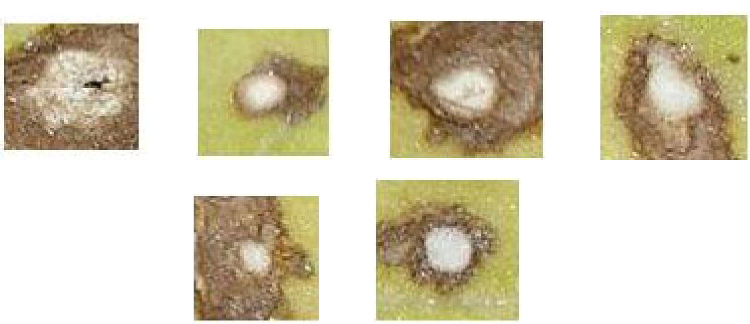


Fig. 2. Extracted Frog-eye spot lesion areas of Fig. 1(a)

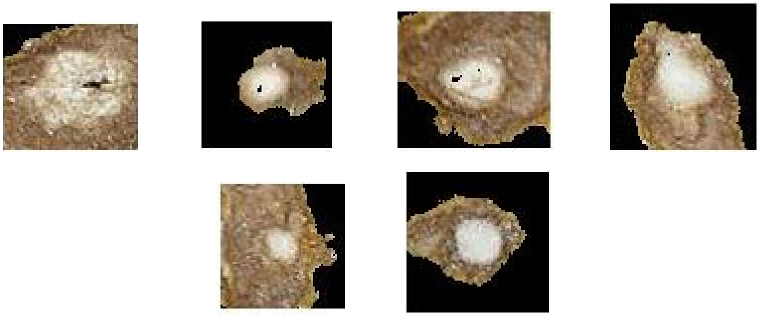


Fig. 3. Segmented Frog-eye spot lesion areas of Fig. 2 using CIELAB color model

3 Evaluation of Proposed Segmentation Algorithm

Measuring the performance of a segmentation algorithm is necessary for two reasons: based on a performance metric, good parameter settings can be found for a segmentation algorithm and the performance of different segmentation approaches can be compared. The region based measures are used when the size and location measurement of the area of the object is essential and is the objective of the segmentation. Therefore, the following region based measures are used to evaluate the proposed segmentation algorithm.

3.1 Measure of Overlap (Jaccard Similarity Measure)

This measure is also known as the area overlap measure (AOM) or the Jaccard similarity measure [16] and is defined as the ratio of the intersection of segmented lesion area S and ground truth lesion area G and the union of segmented lesion area S and ground truth area G. Measure of overlap (MOL) is given in equation 4.

$$MOL = \frac{|S \cap G|}{|S \cup G|} \tag{4}$$

where

$$MOL = Measure\,of\ ovrerlap$$
$$S = Segmented\ area$$
$$G = Ground\,truth\,area$$

3.2 Measure of under Segmentation

This measure is defined as the ratio of the unsegmented lesion area U and the ground truth lesion area G [16]. Measure of under segmentation (MUS) is given in equation 5.

$$MUS = \frac{|U|}{|G|} \tag{5}$$

$$\begin{aligned} &Where \\ &\qquad U = Unsegmented\ lesion\ area \\ &\qquad G = Ground\ truth\ area \end{aligned}$$

$$U = |G \setminus (S \cap G)|$$

$$\begin{aligned} &Where \\ &\qquad S = Segmented\ area \end{aligned}$$

3.3 Measure of over Segmentation

This measure is defined as the ratio of the segmented non-lesion area V and the ground truth area G [16]. Measure of over segmentation (MOS) is given in equation 6.

$$MOS = \frac{|V|}{|S|} \tag{6}$$

$$\begin{aligned} &Where \\ &\qquad V = Segmented\ non-lesion\ area \\ &\qquad S = Segmented\ lesion\ area \end{aligned}$$

$$V = |S \setminus (S \cap G)|$$

$$\begin{aligned} &Where \\ &\qquad G = Ground\ truth\ lesion\ area \end{aligned}$$

3.4 Dice Similarity Measure (DSM)

Dice similarity measure (DSM) is derived from a reliability measure known as the kappa statistic [17] and computes the ratio of the intersection area divided by the mean sum of each individual area. Let C denote the contour of the segmented area, C_r the reference contour, $A(C)$ the set of pixels enclosed by contour C, $A(C_r)$ the set of pixels enclosed by contour C_r and $|C|$ the cardinality of set C. Then the Dice similarity measure (DSM) is defined as.

$$DSM = \frac{2 \times |A(C) \cap A(C_r)|}{|A(C) + A(C_r)|} \tag{7}$$

3.5 Error Rate (ER)

The error rate ER is defined as the normalized agreement of segmentation results and the ground truth [17]. Let C denote the contour of the segmented area, C_r the reference contour, $A(C)$ the set of pixels enclosed by contour, $|C|$ the cardinality of set C, and $\oplus$ exclusive or logical operation. The error rate ER is given in equation 8.

$$ER = \frac{|A(C) \oplus A(C_r)|}{|A(C) + A(C_r)|} \tag{8}$$

3.6 Precision

Precision is the fraction of segmented areas that are relevant to the query. It is given by

$$Precision = \frac{|\{\mathrm{Re}levant\ areas\}| \cap |\{Segmented\ areas\}|}{|\{Segmented\ areas\}|} \tag{9}$$

3.7 Recall

Recall is is the fraction of the areas that are relevant to the query that are successfully segmented. It is given by

$$\mathrm{Re}call = \frac{|\{\mathrm{Re}levant\ areas\}| \cap |\{Segmented\ areas\}|}{|\{\mathrm{Re}levant\ areas\}|} \tag{10}$$

4 Experimental Results

4.1 Dataset

Images of tobacco leaves in colour are acquired using a Sony digital colour camera in an uncontrolled real tobacco field. The colour signals from the camera are transferred as a 24 bit RGB colour image data (1632 × 1224) using a personal computer. The leaves used for imaging are randomly selected from the tobacco seedling bed at Central Tobacco Research Institute (CTRI), Hunsur, Karnataka, India and captured at uncontrolled lighting condition. A total of 400 areas are extracted from 50 frog-eye infected leaves of tobacco seedlings using the segmentation algorithm proposed in section 2. To set up the threshold values of parameters m, T_1, T_2 and k, we conducted experimentation by varying the values of parameters. Experimentally it is found that the values of T_1, T_2, m and k are 0.5, 60, 200 and 25 respectively. i.e., for these values proposed segmentation algorithm has achieved best segmentation performance results.

4.2 Results

The segmentation performance is calculated based on the comparison between the manually segmented ground truth G for lesion area and segmentation result S generated by the image segmentation approach. In order to corroborate the efficacy of the proposed segmentation algorithm, six performance measures such as Measure of overlapping (MOL), Measure of under-segmentation (MUS), Measure of over-segmentation (MOS), Dice similarity measure (DSM), Error-rate (ER), Precision (P) and Recall (R) are calculated for 400 segmented areas of 50 frog-eye spot infected tobacco seedling leaves. Segmentation performance results of proposed segmentation algorithm and the existing segmentation algorithm [1] for representative sample shown in figure 1 are tabulated in Table 1 and Table 2 respectively. Precision and recall results of proposed segmentation algorithm for representative samples are tabulated in Table 3. Average segmentation performance results of 400 extracted lesions for both proposed segmentation algorithm and the existing segmentation algorithm [1] are calculated and tabulated in Table 4. A segmentation algorithm is said to be superior when performance parameters MOL, DSM, Precision and Recall are high and MUS, MOS and ER are low. From Table 4 it is understood that the proposed segmentation algorithm has achieved high value for MOL, DSM, Precision and Recall and low value for MUS, MOS and ER when compared to the existing segmentation algorithm [1].

Table 1. Segmentation performance results of proposed segmentation algorithm for representative sample shown in Fig.1

Lesions	MOL	MUS	MOS	DSM	ER
1	0.6682	0.2366	0.0952	0.8011	0.1999
2	0.7792	0.0104	0.2104	0.8759	0.1241
3	0.6925	0.1906	0.1169	0.8183	0.1817
4	0.7052	0.0228	0.2720	0.8271	0.1729
5	0.6849	0.0328	0.2823	0.8130	0.1870
6	0.7861	0.0219	0.1920	0.8803	0.1197

Table 2. Segmentation performance results of the existing segmentation algorithm [1] for representative sample shown in Fig.1

Lesions	MOL	MUS	MOS	DSM	ER
1	0.4285	0.5651	0.0064	0.5999	0.4001
2	0.3804	0.1066	0.5130	0.5511	0.4489
3	0.4847	0.4733	0.0420	0.6529	0.3471
4	0.6412	0.0427	0.3161	0.7814	0.2186
5	0.5871	0.3286	0.0843	0.7398	0.2602
6	0.5237	0.0171	0.4592	0.6874	0.3126

Table 3. Precision and recall results of proposed segmentation algorithm for representative samples

Samples	Precision	Recall
T1	1	1
T2	0.8571	0.5455
T3	0.7143	0.6667
T3	1	0.7500
T4	1	1
T5	1	1
T6	1	1
T7	0.9091	0.7692
T8	0.8333	0.7143
T9	1	0.9167
T10	1	1

Table 4. Performance results of proposed segmentation algorithm and the existing segmentation algorithm [1]

Segmentation performance measures	Proposed segmentation algorithm	Segmentation algorithm [1]
Average MOL	0.7371	0.4915
Average MUS	0.0952	0.1851
Average MOS	0.1677	0.3234
Average DSM	0.8270	0.6496
Average ER	0.1730	0.3504
Average Precision	0.9472	0.9472
Average Recall	0.8122	0.8122

5 Conclusion

In this work a model for segmenting frog-eye spots lesions in tobacco seedlings has been developed. Lesion areas are segmented efficiently using image enhancement, dilation and erosion operations and color segmentation using CIELAB color model. Proposed segmentation algorithm has evaluated by performance measures. Proposed segmentation algorithm is superior when compared to the existing algorithm [1]. In future we will extend this work to other tobacco seedlings diseases.

References

1. Guru, D.S., Mallikarjuna, P.B., Manjunath, S.: Segmentation and Classification of Tobacco Seedling Diseases. In: COMPUTE 2011 Proceedings of the Fourth Annual ACM Bangalore Conference (2011)

2. Vizhanyo, T., Felfoldi, J.: Enhancing colour differences in images of diseased mushrooms. Computers and Electronics in Agriculture 26, 187–198 (2000)
3. Camargo, A., Smith, J.S.: An image-processing based algorithm to automatically identify plant disease visual symptoms. Biosystem Engineeing 102, 9–21 (2009)
4. Pydipati, R., Burks, T.F., Lee, W.S.: Identification of citrus disease using color texture features and discriminant analysis. Computers and Electronics in Agriculture 52, 49–59 (2006)
5. Sena, D.G., Pinto, F.A.C.: Fall Armyworm Damged Maize Plant Identification using digital Images. Biosystem Engineeing 85(4), 449–454 (2003)
6. Huang, K.: Application of artificial neural network for detecting phalaenopsis seedling diseases using color and texture features. Computers and Electronics in Agriculture 57, 3–11 (2007)
7. Zhang, Y., Mao, H., Hu, B., Li, M.: Features selection of cotton disease Leaves Image Based on Fuzzy Feature selection techniques. In: Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition, pp. 124–129 (2007)
8. Weizheng, S., Yachun, W., Zhanliang, C., Hongda, W.: Grading Mehod of Leaf Spot Disease Based on Image processing. In: 2008 International Conference on Computer Science and Software Engineering, pp. 491–494 (2008)
9. Youwen, T., Tianlai, L., Yan, N.: The recognition of Cucumber Disease Based on Image Processing and Support Vector Machine. In: 2008 Congress on Image and Signal Processing, pp. 262–267 (2008)
10. Mirik, M., Michels Jr., G.J., Kassymzhanova-Mirik, S., Elliott, N.C., Catana, V., Jones, D.B., Bowling, R.: Using digital image analysis and spectral reflectance data to quantify damage by greenbug (Hemitera: Aphididae) in winter wheat. Computers and Electronics in Agriculture 51, 86–98 (2006)
11. Muhammed, H.H.: Hyperspectral Crop Reflectance Data for characterizing and estimating Fungal Disease severity in Wheat. Biosystem Engineeing 95(1), 9–20 (2005)
12. Zhang, Y.C., Mao, H.P., Hu, B., Li, M.X.: Features Selection of Cotton leaves Image based on Fuzzy Feature Selection Techniques. In: International Conference on Wavelet Analysis and Pattern Recognition, pp. 124–129 (2007)
13. Zhou, L., Tong, X. J.: Application of Two-Dimension Wavelet Transform in Image process of Pests in Stored Grain. Seventh International Conference on Machine Learning and Cybernetics, pp. 2790-2793 (2008)
14. Cho, J., Choi, J., Qiao, M., Ji, C., Kim, H., Uhm, K., Chon, T.: Automatic identification of whiteflies, aphids and thrips in greenhouse based on image analysis. International Journal of Mathematics and Computers in Simulation 1(1), 46–53 (2007)
15. Gonzalez, R.C., Woods, R.E., Eddins, S.L.: Digital Image Processing using Matlab. Pearson Education (2007)
16. Elter, M., Held, C., Wittenberg, T.: Contour tracing for segmentation of mammographic masses. Physics in Medicine and Biology 55, 5299–5315 (2010)
17. Yuan, X., Situ, N., Zouridakis, G.: A narrow band graph partitioning method for skin lesion segmentation. Pattern Recognition 42, 1017–1028 (2009)