Proceedings of the 2014 International Conference on
Advanced Mechatronic Systems, Kumamoto, Japan, August 10-12, 2014

Image Segmentation Algorithm for Disease Detection of Wheat Leaves

Xiaojing Niu
College of Information Engineering
Northwest A & F University
Yangling, Shaanxi, China

Shihui Guo
National Centre for Computer Animation
Media School,Bournemouth University
lBournemouth,UK

Meili Wang*
College of Information Engineering
Northwest A & F University
Yangling, Shaanxi, China

Hongming Zhang
College of Information Engineering
Northwest A & F University
Yangling, Shaanxi, China

Xianqiang Chen
College of Information Engineering
Northwest A & F University
Yangling, Shaanxi, China

Dongjian He
Mechanical and Electronic Engineering
Northwest A & F University
Yangling, Shaanxi, China

Abstract: Wheat diseases are harmful to wheat production, but there are few segmentation algorithms that can effectively identify common diseases of wheat leaves. This paper proposes an automatic and efficient solution with K-means clustering. Firstly, the colour image is transformed to Lab colour space from RGB. Clustering is then done by taking the absolute difference between each pixel and the clustering centre in Lab colour space. Unlike traditional methods, our method does not need manual setting for threshold value and is not affected by the selected channel. Our results show that the segmentation accuracy rates for three common diseases (powdery mildew, leaf rust and stripe rust) is more than 90%, which proves the efficacy of our method.

Keywords: wheat disease; K-means clustering algorithm; segmentation algorithm

I. INTRODUCTION

Crop disease is harmful to agricultural production. This is traditionally diagnosed by experts or pathogenic identification, which is time-consuming and delays the optimal treatment time. Computer vision can conduct real-time diagnosis for diseases accurately, rapidly and effectively. By analyzing the common types and characteristics of wheat diseases, it has been found that wheat disease mainly concentrates on the leaves [1] and can be identified by computer vision techniques.

A typical image recognition process includes image pre-processing, segmentation, feature extraction and pattern recognition [2]. Image segmentation is one of the key steps, and the precision of the segmentation directly influences the reliability of feature extraction and the accuracy of recognition. Currently, segmentation algorithms for common crop diseases on leaves include threshold methods, edge detection [3], clustering methods [4,5,6,7] and segmentation methods based on statistical pattern recognition [8] or neural networks [9]. At present, there are image segmentations for rice [9, 10] and maize leaf diseases [11] as well as multi-graph-based segmentation for wheat powdery mildew and stripe rust diseases [12]. These methods aim to segment one or two kinds of crop diseases from an image. However, because of wheat diseases on leaves are diversified, there is not a unified image segmentation method for some common wheat diseases .

Colour images provide more information than gray images, and the segmentation of a colour image is quite important for the recognition of diseases [13,14]. Li et al.[14] applied the k-means clustering algorithm to segment grape leaf diseases. In this paper, we introduced the k-means clustering algorithm to automatically separate diseased regions from non-diseased regions of wheat leafs. The image is converted from RGB into Lab colour space and the pixels are clustered by the differences of nearby pixel colour information. The experimental result shows that our method can successfully accomplish the segmentation task on wheat leaf images of three common diseases (powdery mildew, leaf rust and stripe rust).

The paper is structured as follows: Section 2 states the source of experimental images and simply explains the pre-processing procedure. Section 3 discusses K-means clustering segmentation algorithm in detail. Section 4 compares and analyzes the segmentation results and Section 5 concludes the paper.

II. IMAGE COLLECTION AND PRE-PROCESSING

The images used in this study consist of images of powdery mildew, leaf rust and stripe rust that were gathered from the literature [15] and the Internet. Figure 1 lists the

This work was made possible in part by the National High-tech Research and Development Program (863 Program：2013AA10230402) , the "Five-twelfth" National Key Technology R&D Program (2012BAH30F01, 2012BAH29B04), Major Project of Chinese National Natural Science Foundation (41301283, 41301507)and the fundamental research funds (ZD2012018,QN2013051, 2013BSJJ105, ZD2012018) of Northwest A&F University.

normal leaf (Figure 1(a))and the leaf with diseases (Figure 1(b)(c)(d)) .

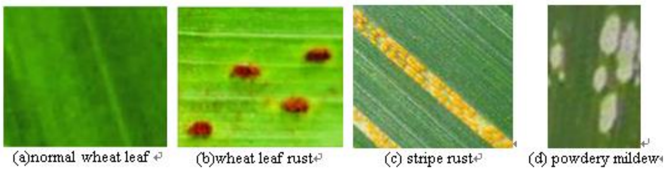


Figure 1 Normal wheat leaf and leafs with diseases

Image pre-processing [2] includes filtering. The purpose of filtering is to remove noise and improve the stability of the disease characteristics. In this study, median filtering is adopted. The value of a point was replaced by the median value of its surrounding points, the process can remove isolated noise points.

A. *Maintaining the Integrity of the Specifications*

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

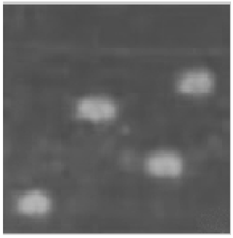
III. K-MEANS CLUSTERING SEGMENTATION ALGORITHM

Traditional segmentation algorithm requires users to manually set parameters for different kinds of wheat diseases and the segmentation accuracy for a specific disease is significantly determined by the selected channel in RGB space[15]. We propose a K-means clustering algorithm in the Lab colour space for wheat disease segmentation.

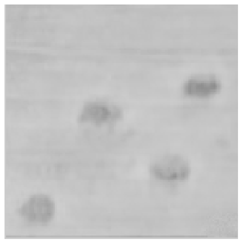
A. *Colour space conversion*

The Lab colour model is composed of three elements: Luminosity (L) and two colour-related parameters. L represents lightness and ranges from 0 to 100. Parameter a represents colours ranging from magenta to green, b represents colours ranging from yellow to blue and both take values from -128 to +127.

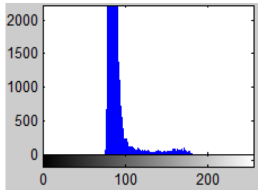
The Lab colour model has several advantages, such as device independence and a wide colour gamut. Moreover, it can compensate for the uneven colour distribution of the RGB colour model, which contains excessive transition colours from blue to green, but lacks yellow shades and transition colours from green to red. Please refer to [15] for the conversion from RGB to Lab colour space. Figure 2 shows a and b component images and their corresponding histograms for Lab images of common wheat leaf diseases (taken leaf rust as an example). It can be seen that the disease images have obvious differences within a and b components and that the histograms of a and b components are complementary. These findings will be used in our next step to segment the image into the background and disease.



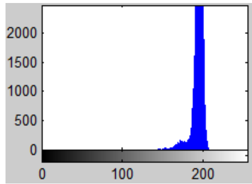
(a) **a** component



(b) **b** component



(a1)histograms of **a** component



(b1)histograms of **b** component

Figure 2 Images of the **a** and **b** components and their corresponding histograms for wheat leaf rust

B. *Implementation of K-means clustering algorithm in wheat leaf disease segmentation*

K-means clustering groups n objects into k clusters while ensuring high similarity within the cluster. The similarity is calculated according to the average value of objects in the cluster [16]. Section 3.1 analyzed that how the Lab a and b colour components possess significant differences in wheat images. Moreover, there exist obvious differences between the disease and the background of the images. As a result, we can segment the image into background and disease two clusters.

Pictures used to test the algorithm are 250 × 250 RGB colour images. After converting them into Lab colour space, the K-means clustering algorithm segments the a or b component images. The segmentation steps are as follows:

① The image data are linearly serialized according to the 2D positions of pixel points, $pixel(i), i \in Z$, where Z is a non-negative integer. Let $A = \{pixel(i), i \in Z\}$, and the set of i is denoted by P.

② Two data point positions (pos1, pos2) are randomly selected from the serialized data to be the initial clustering centres. The selected data points should meet the following conditions:

$$\begin{cases} pos1 = random(), pos1 \in P, pixel(pos1) \in A, \\ pos2 = random(), pos2 \in P, pixel(pos2) \in A \end{cases} \quad (1)$$

where $pos1 \neq pos2$ and $\text{pixel}(pos1) \neq pixel(pos2)$. The clusters are defined by setting pos1 and pos2 as clustering centres C1 and C2, respectively.

③ Data points of the image are segmented into clusters according to the selected cluster centres. The current pixel point pixel(i) is segmented into the cluster with the smallest colour distance from the cluster centre to the pixel:

$$\begin{cases} D_1 = \left|Dis(pixel(i) - pixel(pos1))\right| \\ D_2 = \left|Dis(pixel(i) - pixel(pos2))\right| \end{cases}, \quad pixel(i) \in \begin{cases} C_1, D_1 \le D_2 \\ C_2, D_1 > D_2 \end{cases} \tag{2}$$

The position of the pixel point when placing it into the specified cluster is recorded. In other words, all pixel points in C1 and C2 have their own specific position information.

④ After the clustering segmentation is completed, the pixel points of component a or b are segmented into two clusters. According to the existing clusters, new clustering centres can then be calculated:

$$pos_i' = \frac{1}{n}\sum_{j=1}^{n} pixel(j), i = 1 \; or \; 2, n = size(C_i), pixel(j) \in C_i \tag{3}$$

⑤ If the new clustering centre meets the convergence condition $\left|pos_i - pos_{i-1}\right| < \varepsilon, i \ge 2$, then the clustering process is finished; otherwise the algorithm returns to Step ③ and repeats until the convergence condition is satisfied.

⑥ Binary classification is conducted for component a or b in the original image according to the results of the clustering segmentation:

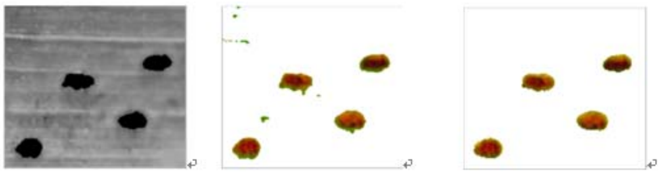
$$pixel(i) = \begin{cases} 0, pixel(i) \in C_1 \\ 1, pixel(i) \in C_2 \end{cases}, \quad pixel(i) \in A \tag{4}$$

⑦ The diseased regions of the original colour image are segmented according to the binary information given in Step ⑥ for each pixel position

IV. COMPARISON AND ANALYSIS OF SEGMENTATION RESULTS

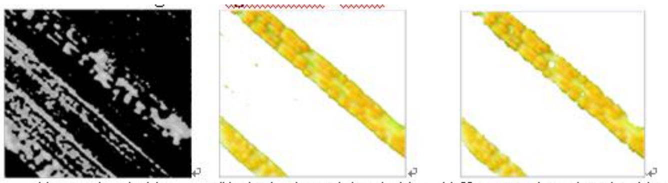
A. Comparison of segmentation results

We implemented and compared the results of grey level segmentation, single channel threshold segmentation and the Lab-based K-means clustering algorithm for wheat leaf rust, stripe rust and powdery mildew. Figures 3-5 present the accordingly results respectively.



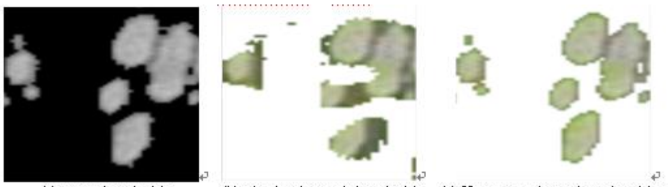
(a) grey threshold (b) single-channel threshold (c) K-means clustering algorithm

Figure 3 Segmentation results of wheat leaf rust



(a) grey threshold (b) single-channel threshold (c) K-means clustering algorithm

Figure 4 Segmentation results of wheat stripe rust



(a) grey threshold (b) single-channel threshold (c) K-means clustering algorithm

Figure 5 Segmentation results of wheat powdery mildew

The segmentation results show that regions of disease and background cannot be distinguished well via grey threshold segmentation, as over-segmentation often occurs. As for single-channel threshold segmentation, different parameters have to be selected and set according to the different wheat leaf disease images and noise remains after segmentation. The Lab-based K-means clustering segmentation algorithm achieves better results than other methods.

B. Analyses on experimental results

We select 90 images of wheat disease (30 each for leaf rust, stripe rust and powdery mildew) and apply our algorithm to each image.

Table 1 lists partial experiment results.

TABLE I. PARTIAL TEST RESULTS OF COMMON WHEAT LEAF SCAB IMAGE SEGMENTATION

Image name	Disease type	Spots number	Spots number of segmentation results	Segmentation rate/ %	Description of segmentation results
leaf0. jpg	Leaf rust	4	4	100	The edge is clear
leaf1. jpg	Leaf rust	80	77	96	Some edges have noise
leaf2. jpg	Leaf rust	55	54	98	Some edges have noise
leaf3. jpg	Leaf rust	59	55	93	Some spots are not fully separated
leng0 .jpg	Stripe rust	2	2	100	The stripe is clear, and the edge has some non-diseased regions
leng1 .jpg	Stripe rust	2	2	100	Some background is not separated

leng2.jpg	Stripe rust	2	2	100	Some background is not separated
leng3.jpg	Stripe rust	2	2	100	Some background is not separated
white 0.jpg	Powdery mildew	6	6	100	The edge has some non-diseased regions
white 1.jpg	Powdery mildew	11	10	91	The edge has some non-diseased regions
white 2.jpg	Powdery mildew	41	40	98	There are some background spots
white 3.jpg	Powdery mildew	16	16	100	Some spots are not fully separated
Image name	Disease type	Spots number	Spots number of segmentation results	Segmentation rate/ %	Description of segmentation results

As shown in Table 1, in images with few scabs and clear edges, scab regions can be completely separated. If the image has noise and crossed-spot regions, the accuracy of segmentation decreases. The results show that the segmentation result of the Lab-based K-means clustering algorithm achieves satisfactory accuracy for common wheat leaf disease images. All results are greater than 90%, and the segmentation rate for stripe rust is 100%.

V. CONCLUSION

This paper proposed an unsupervised, automatic K-means clustering algorithm to perform the segmentation task of wheat leaf scab images based on the Lab colour space. Different from existing work, our work does not require manual setting for threshold value. The experiments on wheat leaf images with three common diseases proved the satisfactory performance of our method in terms of accuracy, efficiency and automation. We believe that the increased accuracy in segmentation will assist in following processes (such as feature extraction and identification) for automatic disease detection and analysis of wheat leaf.

References

[1] Cook, R James and Veseth, Roger J, 1991.Wheat health management. APS Press St. Paul, MN

[2] Gonzalez, Rafael C and Woods, Richard E and Eddins, Steven L,.2009. Digital image processing using MATLAB. Gatesmark Publishing Knoxville.

[3] Mori F, Mori T. Region segmentation and object extraction based on virtual edge and global features[C]//Computer Vision-ACCV 2012 Workshops. Springer Berlin Heidelberg, 2013, pp.182-193.

[4] Hu J, Li D, Chen G, et al. Image Segmentation Method for Crop Nutrient Deficiency Based on Fuzzy C-Means Clustering Algorithm[J]. Intelligent Automation & Soft Computing, vol. 18, no. 8, 2012, pp. 1145-1155.

[5] Ji Z, Xia Y, Chen Q, et al. Fuzzy c-means clustering with weighted image patch for image segmentation[J]. Applied Soft Computing, vol.12, no. 6, 2012, pp. 1659-1667.

[6] He Y, Yousuff Hussaini M, Ma J, et al. A new fuzzy c-means method with total variation regularization for segmentation of images with noisy and incomplete data[J]. Pattern Recognition, vol. 45, no. 9, 2012, pp. 3463-3471.

[7] Mignotte M. A de-texturing and spatially constrained K-means approach for image segmentation[J]. Pattern Recognition Letters, vol. 32, no. 2, 2011, pp.,359-367.

[8] TIAN Y, LI C. Color image segmentation method based on statistical pattern recognition for plant disease diagnose [J][J]. Journal of Jilin University of Technology (Natural Science Edition), vol. 2, 2004, pp. 028.

[9] Zhou Y, Wang Y, Yao Q. Segmentation of rice disease spots based on improved BPNN[C]//Image Analysis and Signal Processing (IASP), 2010 International Conference on. IEEE, 2010, pp. 575-578.

[10] Liu, Li Bo. Research on the Segmentation Method of Rice. Applied Mechanics and Materials, no.220 .2012, pp. 1339-1344.

[11] Zhao Y X, Wang K R, Bai Z Y, et al. Research of maize leaf disease identifying system based image recognition[J]. Scientia Agricultura Sinica, vol. 40, no. 4, 2007, pp. 698-703.

[12] Hu X, Li B, Xi L. Multi-instance graph approach to wheat leaf disease segmentation[J]. Transactions of the Chinese Society of Agricultural Engineering, vol. 28, no. 13, 2012, pp. 154-159.

[13] Xiaohong H, Bingjun L, Fang L. Image segmentation based on graph theory in multi-color space for maize leaf disease[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), vol. 44, no.2, 2013, pp. 177-181.

[14] Li G, Ma Z, Huang C, et al. Segmentation of color images of grape diseases using K_means clustering algorithm[J]. Transactions of the Chinese Society of Agricultural Engineering, vol. 26, no. 14, 2010, pp.32-37.

[15] LIU Lianzhong, ZHANG Wu, ZHU Cheng. Image Recognition of Wheat Diseases Based on Improved Color Feature [J]. Journal of Anhui Agricultural Sciences, vol.40, no.26, 2012 , pp. 12877-12879.

[16] Tatiraju S, Mehta A. Image Segmentation using k-means clustering, EM and Normalized Cuts[J]. Department of EECS, 2008, pp. 1-7.