



# IMAGE SEGMENTATION METHOD FOR COTTON MITE DISEASE BASED ON COLOR FEATURES AND AREA THRESHOLDING

<sup>1</sup>DIAO ZHIHUA, <sup>2</sup>WANG HUAN, <sup>3</sup>SONG YINMAO, <sup>4</sup>WANG YUNPENG

<sup>1</sup>lecturer, Department of Electrical Engineering, ZZULI, Zhengzhou, CHINA

E-mail: [wanghuan19880123@163.com](mailto:wanghuan19880123@163.com)

## ABSTRACT

Cotton mite disease is a common disease. Taking the cotton mite disease as the research subject, a segmentation method based on color features and area thresholding is proposed under the complex background. The proposed algorithm is comprised of three main steps. First, to extract the analogous disease spots (disease spots and stems) from green plants. Second, some special characteristics are detected in gray histogram, afterwards convert the segmented images into 8 bit gray-scale images based on single thresholding. Finally, compare the disease spots' area with stems' and then segmented binary images by using area thresholding. The experimental results show that this algorithm is of effective in segmenting cotton disease spots; the average correct extraction rate of the algorithm can reach 94.79%.

**Keywords:** *Image Segmentation; Cotton Mite Disease; Color Features; Area Thresholding*

## 1. INTRODUCTION

In China, cotton, as the most significant economic crop, plays a decisive role in changing our Agricultural Production Structure. So how to improve the cotton production had become the hot issue in our country. In recent years, in order to prevent a disease called cotton bollworm, large-scale insect-resistant transgenic cotton had been widely planted. But the other disease, cotton spider (cotton mite disease), was breeding in cotton field very quickly which affected the yield and quality severely. So, preventing this disease from cotton had brooked no delay. In the current domestic agricultural production, many experts used personal experiences or pathological knowledge to analyze what those diseases' types are and how those diseases' severity are, but, these identified methods have many disadvantages, such as strong subjectivity, slow speeding, high error rate and low instantaneity. Therefore, we need some new ways to overcome these weaknesses, and rapid development of computer vision technology supplied its possibilities.

At present, many domestic and foreign scholars have devoted to the researches of pest diseases identification based on computer vision technology [1-11]. Perez changed the color into the gray by  $(G-R) / (G+R)$  [12]. Mao Hanping proposed an adaptive segmentation method of crop disease

based on fuzzy C-mean clustering algorithm (FCM). The optimal cluster number and the degree of fuzziness of FCM were chosen through cluster validity and experiments respectively. The results showed that this method was satisfactory to separate disease parts from normal part of leaves [13]. Blasco J, Aleixos N proposed an improved algorithm, which could be sub-divided into three steps that are described as follows: selected seeds, region growing and split-and-merge. Afterwards, the results showed that the segmentation accuracy could reach 95% [14]. But, these above-mentioned study were still under simple background (Firstly, the study object would be extracted from the background directly by the personal ways, and then we deal with the target by the computer software). The researches under complex background (opposite to the simple background) were rare. In this article, the study object is cotton mite disease and the main study contents would include: 1) the segmentation of analogous disease spots and green plants under background; 2) the segmentation of cotton mite disease spots and stems. Aiming at these two big problems, we present an image segmentation algorithm based on color features and area thresholding, which could lay a foundation for automatic identification of cotton mite disease.

2. MATERIAIS AND METHODS

During the June and July in 2011, all the cotton mite disease images were picked out by Nikon D90 SLR digital camera under 12.3 million pixels with a resolution 2592\*1723, when we shoot them, the lens should show a certain angle with cotton diseased leaves for adequate light. The distance between the lens and the earth was about 1.5 meters. In order to following analysis and processes, we converted BMP images to JPG and stored them in the computer. The algorithms in this paper are implemented by C++ language under Microsoft Visual C++ 6.0, and the platform is a personal computer with an Pentium(R) 2.93 GHz CPU, 752MB RAM. The Flowchart of this segmentation algorithm is illustrated in Fig.1, which includes three main steps:

- 1) The segmentation based on Excess green feature 2G-R-B is performed to distinguish analogous disease spot area and non-disease spot area under complex background;
- 2) According to some special characteristics which are displayed in gray histogram, we convert the segmented images into 8 bit gray-scale images based on single thresholding;
- 3) Compare the disease spots' area with stems' to segment the disease spots from the analogous disease spot area by area thresholding.

The detailed algorithms can be discussed in section 3.

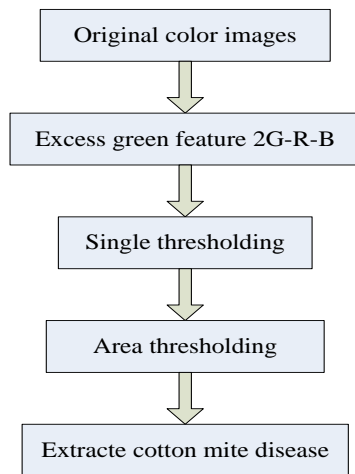


Fig.1 Flowchart Of Segmentation Algorithm For Cotton Mite-Disease Image

3. THE SEGMENTATION OF COTTON MITE DISEASE IMAGE

3.1 The Segmentation Of Analogous Disease Spot Area And Non-Dusease Spot Area Under Complex Background.

In the cotton mite disease images, normal cotton leaves and weeds belong to green plant categories. Nevertheless, cotton mite disease and stems belong to non-green plant categories. According to the color features, the original color images could be divided into two parts, one is the analogous disease spot area and the other is non-disease area. The analogous disease spot area includes mite disease spots and stems, the remaining regions are non-disease area. As mite disease and stems are very similar in color, so it is very difficult to extract the disease spots from the analogous disease spots by the means of color features. Consequently, first, we segmented the analogous mite disease area from the original images by color features, and then extracted the disease spots from the analogous mite area by area thresholding.

Until now, many different color characteristic factors, such as, 2G-R-B、2R-B-G、1.4G-R-B、G/R+B、H etc [15-24] can be used to separate the green plants and non-green plants especially in the Image Processing Technology. According to the researches and analysis of original cotton mite disease image, we can find green plants have some particular properties about color space: For the same pixel in 3-D color space, the pixel values in G-D color space are always higher than in R-D and B-D space. So, we extract analogous disease spots from original images by Excess green feature 2G-R-B. The concrete segmentation algorithm is shown as below: Assuming the coordinate (x,y) is the pixel point, R(x,y), G(x,y), B(x,y) respectively represent the gray values for 3-D color components. In order to later processing, we could set the gray values of analogous disease spot area into white(255), and set the non-disease spot area's into gray(0-255). The Excess green feature (EXG) is given by the expression:

$$EXG(x,y) = \begin{cases} 2G(x,y) - R(x,y) - B(x,y) \\ 255, & 2G(x,y) - R(x,y) - B(x,y) > 255 \\ 0, & 2G(x,y) - R(x,y) - B(x,y) < 0 \end{cases} \quad (1)$$

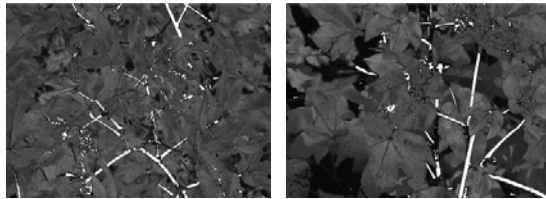
Two original color images are shown in Fig.2. The segmented images after EXG processing are shown in Fig.3. We can find that the white regions are analogous disease spots; on the contrary, Gray's are non-disease spot area. The experiment results

showed that non-green color areas could be extracted from green plants effectively by EXG.



a. b.

Fig.2 Original color images of cotton mite-disease

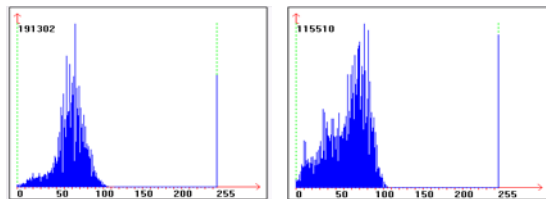


a. b.

Fig.3 The Segmentation Results Based On

### 3.2 The Binaryzation Of Analogous Disease Spot Area And Non-Disease Spot Area.

After the processing, we can get some different gray-scale images (Fig.3a-b). In order to distinguish the differences between analogous disease spot area and non-disease spot area in gray characteristics, some experimental results in the Fig.3 in gray histograms were appended in Fig. 4(a-b).



a. b.

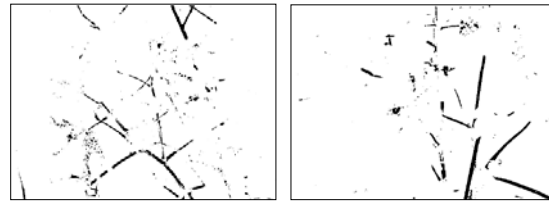
Fig.4 Gray histogram

From the Fig.4, We can find that there have double peaks belong to two different gray-scale areas in the gray histograms: 1) the non-disease spot area whose gray values are less than 120; 2) the disease spot area's are 255. According to these features, the targets would be extracted from the background by thresholding. Because of its intuitive properties, simplicity of implementation, computational speed, steady performance and high efficiency, Thresholding becomes a traditional image segmentation algorithm and is also widely used in image segmentation. In addition, opting the value of  $T$  is the most critical factor. For extracting the analogous disease spot area and displaying black, this paper using single thresholding to segment image after the processing. Suppose

$f(x, y)$  is the input,  $g(x, y)$  is the output,  $T$  is a threshold, the segmented image,  $g(x, y)$  is given by:

$$g(x, y) = \begin{cases} 255, & f(x, y) < T \\ 0, & f(x, y) > T \end{cases} \quad (2)$$

Select a right threshold,  $T$ , and then results are shown in Fig.5.



a. b.

Fig.5 The segmentation result based on single thresholding

In Fig.5, we could find that these black pixels are analogous disease spots areas which include mite disease and stems, and the white background are non-disease spots' areas. Experiments show that, using the single thresholding could segment the analogous disease spots' areas very well.

### 3.3 The Segmentation Of Mite Disease Spots And Stems Based On Area Thresholding.

According to the researches on many cotton mite disease image samples, we could find that the diseased parts are generally on the leaves and the stems exist independently. In the Fig.5, the disease spots and stems distribute every corner of picture disorderly and irregularly after the segmentation based on  $2G - R - B$  and single thresholding. Through the further amplification to binary image, the segmented images are constituted by a large number of independent 8-connected region, and the connected region of stems are larger than cotton mite disease spots'. According to this feature, the image can be grouped into two categories: the one is small connected region composed by cotton disease spots; the other is large connected region composed by cotton stems. On the basis of differences in areas (pixel numbers) between the two regions, we select the right threshold to eliminate stem regions by using area thresholding, moreover, the remaining regions are cotton mite disease area. Suppose every 8-connected region is  $S$ , area thresholding is  $nBlackArea$ . The specific area threshold equation is defined as:

$$S = \begin{cases} 255, & S > nBlackArea \\ 0, & S < nBlackArea \end{cases} \quad (3)$$

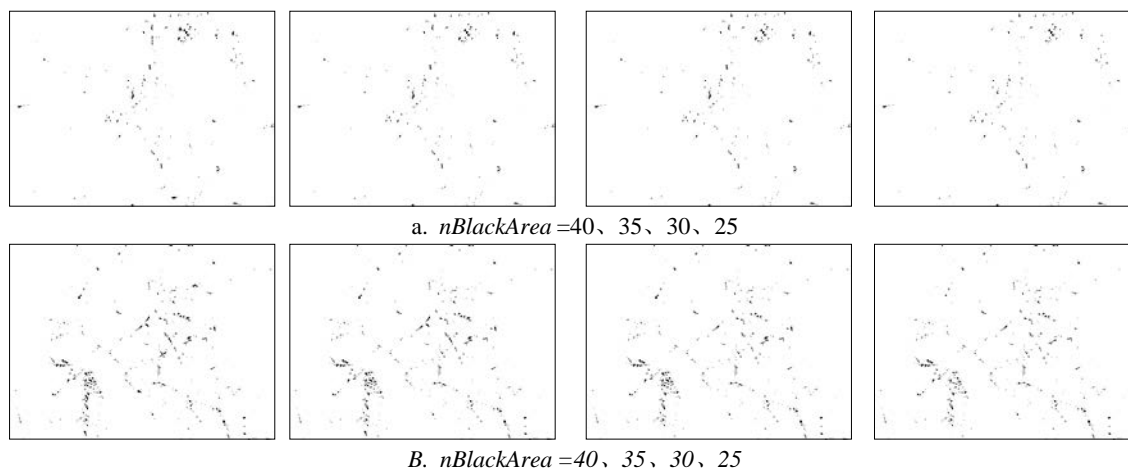


Fig.6 Comparison Of The Segmentation Results With Different nBlackArea Values

In order to better eliminate the stems, we select the nBlackArea 40、35、30、25 separately in this experiment. The two images' results are shown in Fig.6 after area thresholding. We can clearly find that when the nBlackArea is 40 or 35, all of the disease spots are retained, but some small stems still exist. When the nBlackArea is 25, although all the cotton stems are eliminated, some larger disease spots are reduced or removed completely; however, when the nBlackArea is 30, the cotton disease spots can not be only retained, but the stems are eliminated effectively. In order to evaluate the effectiveness of this area thresholding method, this paper referred to the George's evaluation method [25], and compared actual disease area with extracted disease area according to the different nBlackArea. The experimental results are respectively in Table 1.

Table1 The Experiment Results Of Area Thresholding

Num	nBlackArea	actual disease area	extracted disease area	error extraction rate(%)(e <sub>r</sub> )
a	40	22196	23328	5.1
	35	22196	23261	4.8
	30	22196	21153	4.7
	25	22196	21086	5.0
b	40	16834	17760	5.5
	35	16834	17726	5.3
	30	16834	15992	5.0
	25	16834	15959	5.2

In the Table1, error extraction rate is defined as:

$$e_r = \frac{|\text{extracted disease area} - \text{actual disease area}|}{\text{actual disease area}} \times 100\% \quad (4)$$

In Table 1, data shows that the values of actual disease area are solid to each image. Because the area of stems or disease spots which are larger than nBlackArea are segmented, therefore, when the values of nBlackArea of reduced continuously, extracted disease area decreased correspondingly. Suppose the largest 8-connected region is in the analogous disease spot area, and the smallest is. We could put forward some conclusions as follows: 1) nBlackArea>Smax, no pixels in these binary images could be removed from analogous disease spot area; 2) nBlackArea≤Smax, stems and even disease spots whose values are between and will be eliminated; 3) nBlackArea>Smax, all the pixels in the images will be removed completely. Therefore, selecting a appropriate nBlackArea is the key point. A further anlysis show that when nBlackArea>35, the extracted disease area is larger than the actual disease area and the error extraction rate is on the increasing. This situation indicates that some small stems are not been eliminated. When nBlackArea<30, the extracted disease area is smaller than the actual disease area and the error extraction rate is also on the increasing. It means that some disease spots and stems are segmented. As a result, the optimal nBlackArea is 30. When nBlackArea is 30, the error extraction rate is respectively 4.7% and 5.0%.

#### 4. EXPERIMENT AND RESULT

In order to test the validity and generality of the proposed method, 30 color images of cotton mite-

disease are picked out to segment by proposed algorithm, a split plot with repeated measures in the error extraction rate was used as the experimental design.

*Table2 The Experiment Results Of Segmentation Algorithm*

Num	error extraction rate (%)	Num	error extraction rate (%)
1	5.1	16	4.8
2	4.9	17	5.3
3	5.2	18	4.7
4	4.7	19	5.1
5	4.6	20	4.9
6	5.1	21	5.0
7	5.0	22	5.2
8	4.5	23	5.0
9	4.8	24	4.6
10	4.3	25	4.5
11	5.0	26	4.9
12	4.9	27	4.4
13	4.7	28	4.6
14	4.8	29	5.1
15	5.1	30	5.2

From the statistics, it is known that the average error extraction rate of 30 color images of cotton mite-disease is 5.21%, that is to say, the average correct extraction rate could reach 94.79%.

## 5. CONCLUSIONS AND DISSUSIONS

According to the properties of cotton disease image, this paper proposed an image segmentation algorithms based on color features and area thresholding. This proposed algorithm combined 2G-R-B、single threshold and area thresholding to get disease spots very greatly. Some advantages were listed as following:

Compared with some researches [26-27]based on simple background, this study which is on the basis of complex background is more practical.

According to the color features of disease spots, this segmentation algorithm which could change the color images of cotton mite-disease into binary is on the combination of EXG and single thresholding, this proposed algorithm overcame EXG+Autothreshold 's [28-29]defects, for instance: which has illumination problems; Relative error rate is high; computational complexity etc.

Though the comparison of differences between disease spots and stems in size, we could eliminate

the stems by area thresholding. Recently, an increasingly number of segmentation algorithms based on plant disease had been proposed. Ren Yanna proposed a probabilistic model based on pixel rice lesion segmentation. By adaptive lesions segmentation algorithm, Gauss mixture model was used to describe the color distribution of each pixel[30], However, these segmentation algorithms based on probability model have some defects: big works on calculation, high time complexity, and it is not suitable to deal with high-resolution and big amount pixels images; Fuzzy Clustering Analysis technology was applied in segmentation of disease image in the research of Ma Xu [31], But, classical fuzzy clustering algorithms also have some drawbacks, such as slow convergence speed, sensitive to initialization data and the segmented regions' number must be confirmed. Lv Fuxiang conducted a new algorithm based on Mathematical Morphology and spline interpolation, developed a high-resolution and non-destructive measuring system[32], But Mathematical Morphology have some weakness: a) because the algorithms were two-dimensional convolution operations, computational speed is slow; b) The selection of structuring elements didn't have a unified standard; c) The theory of color morphology was not matured. Liu Shuwen studied the expression method for fuzzy subordination and the mode of fuzzy back-propagation artificial neural network for 26 kinds of common grape diseases[33], Nevertheless, neural network system is not flawless, some deficiencies are as following: there are no methods to make easy explanation of the neural network till now. In case of over-learning, there must be some appropriate methods to judge the neural network strictly. Unless the problem is rather simple, it would take much more time to train the neural network. The area thresholding overcome all the above those shortcomings of segmentation method basically, and also has a good real-time performance and robustness.

From the above, this proposed algorithm can better complete segmentation of color images of cotton mite-disease; average correct extraction rate could reach 94.79%. It could lay a foundation for automatic identification of cotton mite disease.

## REFERENCES:

- [1] Chun-Chleh.Y, S.O. Prasher , J.A. Landry, K. Robert. A vegetation localization algorithm for precision farming [J]. Biosystems engineering, 2002, 81(2): 137-146.





- [2] G.E. Meyer, J.C. Nero, D.D. Jones, Timothy.W.Hindman. Intensified fuzzy clusters for classifying plant, soil, and residue regions of interest from color images. *Computers and electronics in agriculture* [J]. 2004, 42(3): 161-180.
- [3] D. M. Woebbecke, G.E. Meyer, K.V. Bargaen, D.A. Mortensen. Color indices for weed identification under various soil, residue and lighting conditions. [J]. *Transactions of the ASAE*, 1995, 38(1): 259-269.
- [4] B.L.Steward, L.E.Tian. Machine-Vision weed density estimation for real-time, outdoor lighting conditions [J]. *Transactions of the ASAE*, 1999, 42(6): 1897-1909.
- [5] Tian Y W, Li T L. Color image segmentation of cucumber disease leaf using support vector machines [J]. *Chinese Journal of Scientific Instrument*, 2007, 28(43):461-463.
- [6] Yuan Y, Li M, Liang Q. Segmentation method for crop disease leaf images with complex background. [J]. *Transactions of the CSAE*, 2011, 27(2): 208-212.
- [7] A. Camargo, J.S. Smith. Image pattern classification for the identification of disease causing agents [J]. *Computers and Electronics in Agriculture*, 2009, 66 (2): 121-125.
- [8] Zhang Y C, Mao H P, Hu B. Separate algorithm for overlapping spots in crop disease image [J]. *Transactions of the Chinese Society of Agricultural Machinery*, 2008, 9(2): 112-115.
- [9] Tang L., L.E Tian, B. L. Steward, et al. Texture-based weed classification using Gabor wavelets and neural network for real-time selective herbicide applications[J]. *ASAE*, 1999, Paper No. 993036.
- [10] Xu G L, Zhang F L, S.G. Shah. Use of leaf color images to identify nitrogen and potassium deficient tomatoes [J]. *Pattern Recognition Letters*, 2011, 32(11): 1584-1590.
- [11] Lai J C, Tang X J, Li S K. Image segmentation method for sugarcane diseases based on color and shape features [J]. *Scientia Agricultural Sinica*, 2010, 43(7): 1363-1369.
- [12] A.J. Perez, F. Lopez, J.V. Benlloch, S. Christensen. Color and shape analysis techniques for weed detection in cereal fields [J]. *Computers and electronics in agriculture*, 2000, 25(3): 197-212.
- [13] Mao H P, Hu B, Zhang Y C. Segmentation of crop disease leaf images using fuzzy c-means clustering algorithm [J]. *Transactions of the CSAE*, 2008, 24(9): 136-140.
- [14] Blasco J, Aleixos N. Computer vision detection of peel defects in citrus by means of a region oriented segmentation algorithm [J]. *Journal of Food Engineering*, 2007, 81(3): 535-543.
- [15] Liu X C, Ai J Y, Tang J L. The segmentation of black rot lesion of cruciferous plant based on self-adaptive fuzzy threshold [J]. *Journal of image and Graphics*, 2009, 14 (7): 1334-1340.
- [16] Zhao J H, Luo X W, Zhou Z Y. Image segmentation method for sugarcane diseases based on color and shape features [J]. *Transactions of the Chinese Society of Agricultural Machinery*, 2008, 39(3): 100-103.
- [17] A. Camargo, J.S. Smith. An image-processing based algorithm to automatically identify plant disease visual symptoms [J]. *Biosystem Engineering*, 2009, 102 (1): 9-21.
- [18] Huang K Y. Application of artificial neural network for detecting Phalaenopsis seeding diseases using color and texture features [J]. *Computers and Electronics in Agriculture*, 2007, 57(1): 3-11.
- [19] T. Baum, U. Seiffert. HyphArea-Automated analysis of spatiotemporal fungal patterns [J]. *Journal of Plant Physiology*, 2011, 168 (1): 72-78.
- [20] J.I. Arribas, G.V.Sánchez-Ferrero, Gonzalo Ruiz-Ruiz. Leaf classification in sunflower crops by computer vision and neural networks [J]. *Computers and Electronics in Agriculture*, 2011, 78 (1): 9-18.
- [21] Xavier P. Burgos-Artizzu, Angela Ribeiro, Maria Guijarro. Real-time image processing for crop/weed discrimination in maize fields [J]. *Computers and Electronics in Agriculture*, 2010, 75(2): 337-346.
- [22] Lin X Y, Liu W Y, Chen X D. Spore image recognition of poplar disease [J]. *Chinese Journal of Scientific Instrument*, 2003, 24(4): 364-366.
- [23] Chen J J, Ji S W, Li J. Automatic measurement of danger degree of cotton insect pests using computer vision [J]. *Transactions of the CSAE*, 2001, 17(2): 157-160.
- [24] Aitkenhead M J, Dalgetty I A, Mullins C E, et al. Weed and crop discrimination using image analysis and artificial intelligence methods [J]. *Computers and Electronics in Agriculture*, 2003, 39(3): 157-171.



- [25] G.E. Meyer, J.C. Neto. Verification of color vegetation indices for automated crop imaging applications[J]. *Computer and Electronics in Agriculture*, 2008, 63(2): 282-293.
- [26] Chen Y H, Hu X G, Zhang C L. Algorithm for segmentation of insect pest images from wheat leaves based on machine vision [J]. *Transactions of the CSAE*, 2007, 23(12): 187-191.
- [27] P. Boissard, V. Martin, S. Moisan. A cognitive vision approach to early pest detection in greenhouse crops [J]. *Computers and Electronics in Agriculture*, 2008, 62(1): 81-93.
- [28] Zhang Z B, Luo X W, Hou F X. Segmentation algorithm based on color feature for green crop plants [J]. *Transactions of the CSAE*, 2011, 27(7): 183-189.
- [29] G.E. Meyer, J.C. Neto. Verification of color vegetation indices for automated imaging applications [J]. *Computers and Electronics in Agriculture*, 2008, 63(1): 282-293.
- [30] Ren Y N, Xi L, Wang Q. Design of rice lesion identification system based embedded [J]. *Computer Simulation*, 2011, 28(2): 341-344.
- [31] Ma X, Qi L, Zhang X C. Segmentation technology of *exserchilum turcicum* image based on fuzzy clustering analysis [J]. *Journal of Agricultural Mechanization Research*, 2008, 12: 24-26.
- [32] Lv F X, Miao Y B. Acquisition of melon morphologies through high-precision and non-destructive vision measurement [J]. *Journal of Agricultural Mechanization Research*, 2011,(7): 39-44.
- [33] Liu S W, Wang Q W, He D J. Grape disease diagnosis system based on fuzzy neural network [J]. *Transactions of the CSAE*, 2006, 22(9): 144-147.