A Study of Image Processing on Identifying Cucumber Disease

Yong Wei, Ruokui Chang, Yuanhong Wang, Hua Liu, Yanhong Du, Jianfeng Xu, Ling Yang

A Study of Image Processing on Identifying Cucumber Disease

Yong Wei [1], Ruokui Chang [1*], Hua Liu [1],Yanhong Du[1] , Jianfeng Xu [1]
[1]Department of Electromechanical Engineering, Tianjin Agricultural University, Tianjin, P.R. China 300384
[*]Corresponding author, Address: Department of Horticulture, Tianjin Agricultural University, Tianjin, P.R .China 300384 , weiytj@sohu.com

Abstract. Plant disease has been a major constraining factor in the production of cucumber，the traditional diagnostic methods usually take a long time, and the control period is often missed. We take computer image processing as a method, preprocessing the images of more than 100 sheets of collected samples of cucumber leaves, using the region growing method to extract scab area of leaves to get three feature parameters of shape, color and texture. And then, through the establishment of BP neural network pattern, the model identification accuracy of cucumber leaf disease can reach 80%. The experiment shows that by using this method, the diseases of cucumber leaves can be identified more quickly and accurately. And the feature extraction and automatic diagnosis of cucumber leaf disease can be achieved.

Keywords: cucumber disease; texture feature; feature extraction

1. Introduction

As an important component of plant pathology, leaf disease is increasingly valued by botanists and pathologists. Computer vision technology provides new research ideas and methods for rapid automatic identification and diagnosis of plant disease. As early as 1999, Yuataka Sasaki studied the automatic diagnostics of cucumber anthracnose from the perspective of spectral reflectance properties and shape characteristics[1].Wang Shuangxi of Shanxi Agricultural University has taken the cucumber disease leaf in greenhouses as an example to carry out in-depth research on image segmentation, image enhancement and feature extraction. In this study, we regard the common cucumber angular leaf spot, powdery mildew and downy mildew as the main research objects, exploring the way of achieving rapid and accurate diagnosis of diseases of cucumber through extracting three characteristic values of shape, texture and color. It will provide technical support for the safe production of cucumber.

2. Materials and Methods

The experimental images are collected from cucumber greenhouses of Dangcheng village, Xinkou town, Xiqing District of Tianjin. To ensure leaf image are of comparability, after making the leaves which are infected by different diseases clean and dry, we take some photos from the same angle and height under the same environmental conditions. Digital camera of Olympus FE5030 is used in the experiment, with the pixel of 14 million. The processor is Intel Core i3, 2GB memory, and the captured image is 1820 * 960 pixels.

It took about two hours to deal with image processing of one picture because of the high pixel. So we use screenshot tool to reduce it to 750 * 650 pixels. This can greatly save the time of single image processing, increase availability and meet the requirements of the pixels in the follow-up experiments.

In the experiment, we use the matlab image processing tool; comprehensively use the knowledge of image processing, plant pathology, color science and pattern recognition to diagnosis cucumber angular leaf spot, powdery mildew and downy mildew.

We can get scabs on the leaves and extract three eigenvalues of shape, texture and color by processing the known disease of cucumber leaf image. It will be sent into the neural network. Finally, by using the trained neural network we can identify unknown disease in cucumber leaves.

3. Image Processing of Cucumber Leaves

3.1. Image Processing of Cucumber Leaves

This section includes image preprocessing, image binarization, morphological image enhancement, the extraction of image background with region growing method, the binary image of Scab by using image subtract operation to obtain color images of the scab. As is shown in **Fig.1.**

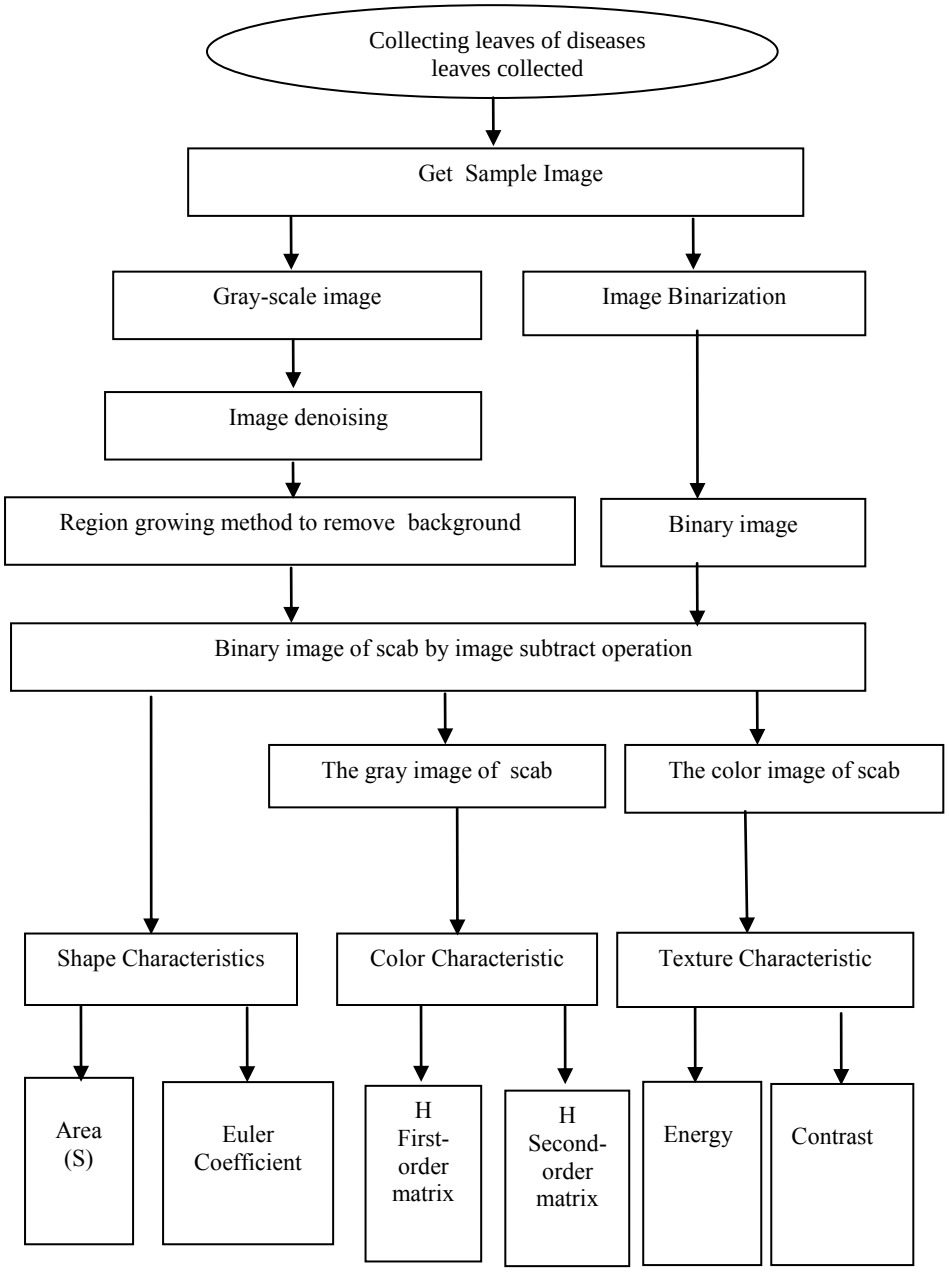
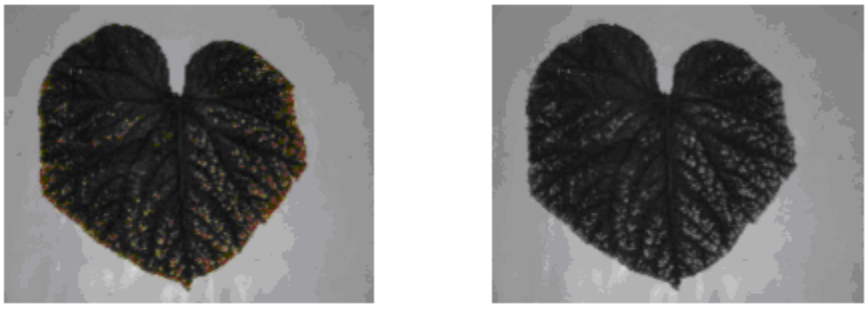


Fig.1. Flow chart of image feature extraction

3.2. Image preprocessing

As a pre-processing of Image pre-processing feature extraction and image recognition, It is an important part of the process of disease identification. The purpose of preprocessing is to improve the image of intelligibility, It includes three links of image denoising, image background segmentation and binary image. Experiments show that median filtering is the best denoising, matlab implement median filter by calling function of medfilt . Deal with the effects is shown in **Fig.2.**



Original image Filtered gray scale

Fig.2. Partial results of pretreatment

We find that the applicability of using region growing method to extract background is the strongest by comparing region growing, quadtree segmentation and robert / prewitt operator in the experiment. By using this method, we can extract a clearer pre-treated gray background.

After several experiments, we choose a more suitable growing point, set the threshold condition as less than 0.5 and then get a more satisfying image of background segmentation, which is shown in **Fig.3**. In order to obtain a more satisfactory image, we have to select growing points or threshold repeatedly.

We can get the binarization of grayscale after pretreatment, making preparations for the following image subtract operation. The global thresholding and OTSU algorithm are used in the experiment, the threshold value is more accurate than that from the R, G, B, L-channel histogram of each component, and the error is smaller, the segmentation is more precise. It is considered as the best algorithm of the segmentation threshold selection, and it is easy to compute, not easily affected by the impact of brightness and contrast ratio of images. The result is shown in **Fig.4**.

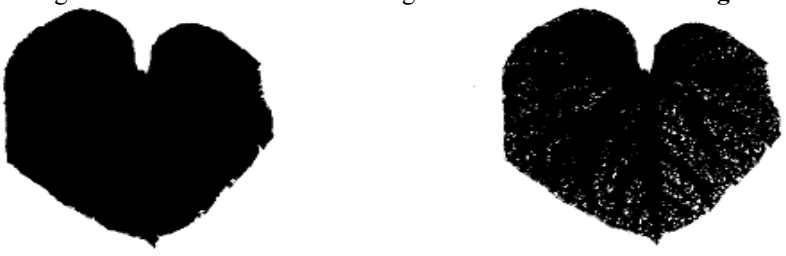


Fig.3. Background **Fig.4.** Binary image

The edge of binary image is not very satisfactory by using OUTS, if it used directly on image subtract operation, the flaws can be enlarged. Therefore, we should use morphological processing to deal with the binary image before the image subtract operation. There are four methods available: expansion, erosion, opening operation (erosion after the first expansion) and closing operation (after the first expansion of corrosion). If we compare the results of treatment (**Fig.5)**, we can see that the result of closing operation is more in line with the original image. Therefore, we can use closing operation to enhance the binary figure in experiments.

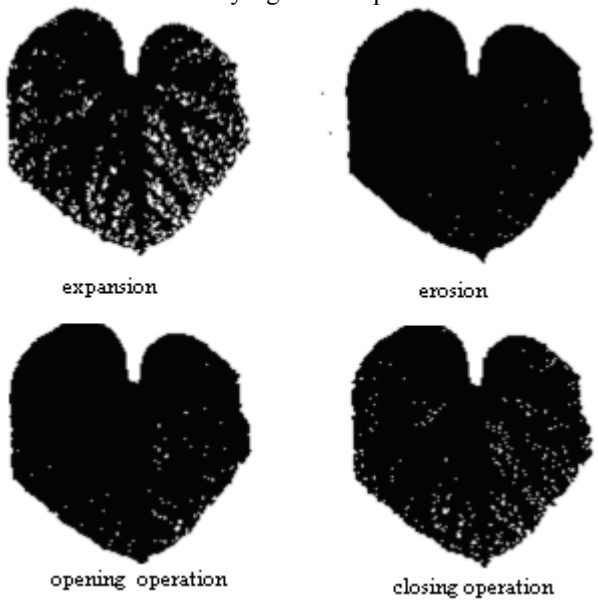


Fig.5. Results of morphological processing

We will get the binary image of scab with scab only if we make subtraction between the background image obtained in the experiment (**Fig.3**) and binary image. It is shown in **Fig.6**. According to the Binary image of scab, we can determine the pixel position of scab area, extract three-channel value of R, G, B from the original rgb map, then use cat function to gain color picture of scab with scab only, it is shown in **Fig.7**.

Fig.6. Binary image of scab

Fig.7. RGB image of scab

3.3. The extraction of feature parameters

In order to create a system to identify different types of diseases of cucumber leaves, we must determine what characteristics of the objects to be diagnosed have, and then it will result in description parameters. Through the previous experiments, the images only containing scabs have been obtained. Then we will begin the study of the characteristics of scabs to find out the parameters that can describe these features. Typically, the blade can be divided into three characteristic values: the shape features, texture features and color features.

The shape characteristics selected through experiments is the Euler number of disease spots area. We use bwarea function to extract the disease spots area of binary image (**Fig.6**); the Euler number of image is the measurement which is the number of empty holes within the region of hole.We can gain the Euler number of disease spots through by Bweuler function.

Grain character is a kind of visual features which is not dependent on the color or brightness to reflect the homogeneous phenomenon in images; it is a kind of image features which reflect the properties of the spatial distribution of pixel gray level. There are two commonly used methods of texture description: statistical method and spectral method. Statistical method describes the texture that is based on the characteristics of the image histogram. Fourier spectral method is based on the description of the texture spectrum method.

We use gray level co-occurrence matrix (GLCM) to describe the texture, it belongs to statistical methods.Co-occurrence matrix describes the spatial structure of the characteristics and correlation of pixels which is based on gray value of image. In the experiment, we construct the GLCM of image which is based on joint probability density function of the location of two pixels.GLCM is the second-order statistical measure of gray-scale image. The structural parameters of Co-occurrence matrix is determined that d take 1, θ take 0 degrees. In matlab, use out = cooccurrence (I, 0, 1, 1) to be GLCM of grayscale images; Co-occurrence is the sub-function of extracted GLCM Functions. We can obtain texture characteristic parameters by extracting energy and contrast ratio of GLCM (**Fig.2**). Energy is a measure of image texture; it reflects the image gray level distribution and texture coarseness. Coarse texture, large energy; texture thin, energy is small. It is shown in formula 1.M, N for the GLCM rows and columns.

$$\text{Energy}= \sum_{i=1,j=1}^{M,N} out_{i,j}{}^{2} \tag{1}$$

Contrast (moment of inertia) is average of the image difference of all pixel gray-scale treatment of | i-j |. Gray level difference is the contrast of gray pixels of the more higher the value. It is the clarity of the image, the texture of the deep grooves. It is shown in formula 2.

$$\text{Contrast}= \sum_{i=1,j=1}^{M,N} (i-j)^{2} out_{i,j} \tag{2}$$

Colors are the main features for people to pay attention to and memorize images. Experiments show that the leaves in the H channel of HSV color space are the color feature parameter under the first moment and second moment. We can extract color feature parameters of disease spots, namely, the first moment and second moment of H color.

Among them, the tone of the first-order matrix and second-order matrix equation is shown as follows(3),(4).

$$M1=\frac{1}{N}\sum_{i=1,j=1}^{M,N} H_{i,j} \tag{3}$$

$$M2=\left[\frac{1}{N}\sum_{i=1,j=1}^{M,N}\left(H_{i,j}-M1\right)^2\right]^{1/2} \tag{4}$$

4. Cucumber Leaf Disease Identification System

BP neural network is mainly used in function approximation, pattern recognition, classification and data compression. It is a one-way transmission of multi-layer feedforward network. It consists of input layer, hidden layer (may have multi-layer) and output layer, each layer is formed by different nodes. For the cucumber disease diagnosis problem, the symptoms can be seen as input to the output of disease name (type number) of the nonlinear mapping problem. We chose input parameters of the network from the extracted six characteristic parameters of cucumber leaves image, which are, scab size S, EULER of scab area, first moment M1 of H color channel, second moment M2 of H color channel, gray degree of co-occurrence matrix of energy and contrast. The output vector of neural network is the type of cucumber disease. From the foregoing, the neural network input nodes n = 6, the output nodes for the m = 1. The structure of neural network model is shown in **Fig.8**.

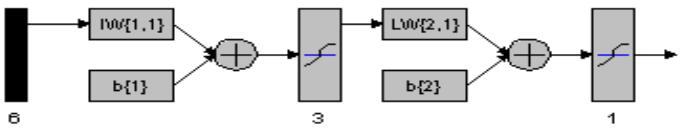


Fig.8. Neural network structure

The experiment verifies that if we choose three layers from the middle layers, the training error is 0.01, the training step is 10000, the input layer and hidden layer transfer function is tansig, the hidden layer and output layer transfer function is purline, network training function is trainlm, and when Network output error range is

0.01, we find a good trained neural network model to recognize successfully, the error is shown in **Fig. 9**.

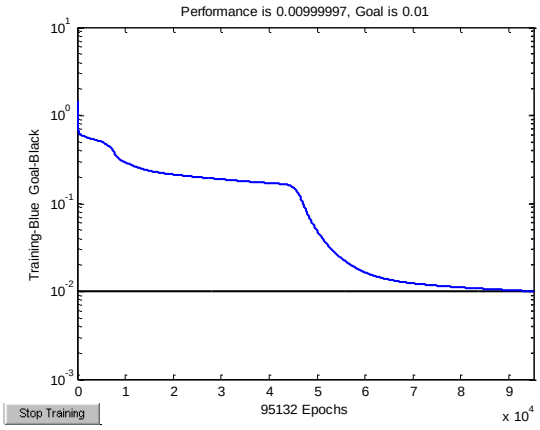


Fig.9. Neural network training error

Using the test samples to detect trained network, we discover that the correct rate of identification can reach 80%. It shows that the trained BP neural network can complete the basic disease recognition task. Due to fewer samples, recognition error rate was higher; Secondly, experimental results are vulnerable to the interference from other factors such as leaf images. For example, some of the leaves have been infected more than one virus. To improve the recognition rate caused by neural network requires the sample to be comparable experimental and statistical.

5. Conclusions

The experiments proved that the use of Matlab toolbox can be realized on cucumber leaf disease identification, Through the preprocessing operation of images, the experiment shows us that median filtering was an effective way to de-noising, the application of “Image Subtraction method” and “region growing” is an effective way to separate effective areas and extract image feature parameters of the region in a complex background. Experiment using the region growing method to extract leaf scab area, we established BP neural network model by obtaining the shape, color, texture characteristic parameters, the model identification accuracy of cucumber leaf disease can reach 80%.

According to the conclusions of this paper, there are some design limitations; it is proposed the following visions for the future:

(1) To achieve the degree of cucumber disease classification, at present, we can only identify the disease and we can not give the classification data of the extent of disease, and also, we should give the appropriate classification of treatment options to growers.

(2) The optimization algorithm and the processing speed should be improved. The current algorithm contains nested, loops and subroutine calls and so on, And with

high pixels of leaves photo, the processing time takes about 40 minutes. This will be improved in the future.

(3) Simplify the structure, expert systems should be made to achieve product-oriented.

ACKNOWLEDGEMENTS

Funding for this research was provided by Tianjin Agricultural University (P. R. China).

REFERENCES

1. Yuataka SASAKI, Tsuguo OKAMOTO, Kenji IMOU, TOR. Automatic Diagnosis of Plant Disease. Journal of JSAM.1999,61(2):119~126
2. Hamed Hamid Muhammed. Hyperspectralcrop reflectance data for characterizing and estimating fungal disease severity in wheat. J.Biosystems Engineering, 91(1),pp. 9-20(2005).
3. Jing Z., Shuangxi W., Xiaozhi D.: A study on method of extract of texture characteristic value in image processing for plant disease of greenhouse. J . Journal of Shenyang Agricultural University, 37 (3),pp.282 -285(2006).
4. Changxing G., Junxion Z.: Recognition and Features Extraction of Cucumber Downy Mildew Based on Color and Texture. J. Transactions of the Chinese Society for Agricultural Machienry, 42(3),pp.170-174.(2011).
5. Bingqi C., Xuemei G. Xiaohua L.: Image diagnosis algorithm of diseased wheat. J. Transactions of the Chinese Society for Agricultural Machienry, 40(12),pp.190-195(2009).
6. Hanping M.,Guili X , Pingping. L.: Diagnosis of nutrientdeficiency of tomato based on computervision.[J]. Transactions of the Chinese Society for Agricultural Machienry, 34(2),pp.73-75.(2003).