The Research of the Strawberry Disease Identification Based on Image Processing and Pattern Recognition

Changqi Ouyang*, Daoliang Li, Jianlun Wang, Shuting Wang, and Yu Han

College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, P.R. China
ouyangchangqi@126.com

Abstract. In this paper, a synthesis segmentation algorithm is designed for the real-time online diseased strawberry images in greenhouse. First, preprocess images to eliminate the impact of uneven illumination through the “top-hat” transform, and remove noise interference by median filtering. After comprehensively applying the methods of gray morphology, logical operation, OTSU and mean shift segmentation, we can obtain the complete strawberry fruit area of the image. Normalize the extracted eigenvalues, and use eigenvectors of part of the samples for training the BP neural network and support vector machine, the remaining samples were tested in two kinds of disease strawberry recognition model. Results show that support vector machines have a higher recognition rate than the BP neural network.

Keywords: strawberry fruit, gray morphology, OTSU, mean shift segmentation, Pattern recognition.

1 Introduction

Strawberry fruit image segmentation is used widely in strawberry picking and fruit grading.In strawberry picking , the strawberry fruit is mature normal strawberry, it’s easy to segment the fruit image because the color of the fruit area is almost the same. In strawberry fruit grading , the strawberry fruit has already been picked, and it’s not difficult for image segmentation because the background is relatively consistent. The strawberry image segmentation algorithm discussed in this paper aims to offer the objects to the real-time online identification of the three common strawberry diseases. Therefore, this algorithm is for the image segmentation of the complete strawberry fruit without picking it. As the color distribution of diseased strawberry fruit is inconsistent, and fruit background is quite complicated, it needs to combine kinds of image processing methods to obtain the complete segmentation of the strawberry fruit according to the characteristics of the diseased fruit area.

Pattern recognition is an intelligent activity, including analysis and judgment. The analysis process is to determine the division of pattern class characteristics and

* Corresponding author.

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 69–77, 2013.

expression method; judgment is reflected in the characteristics of the unknown object, and its judgment belongs to a certain class. It aims to use a computer to achieve the ability to identify the class, which is the contrast of two different levels of identification ability. Pattern recognition system includes the sample acquisition, preprocessing, feature extraction, classification decisions and classifier design.

2 Material and Equipment

The objects studied in this paper are the strawberries of the three common diseases in greenhouse,as shown in Fig 1.These three diseases are powdery mildew, shrinkage and uneven ripening. The characteristic of powdery mildew is that the fruit surface is coarse and with a layer of white powder. The characteristic of shinkage is that the fruit atrophies and is deformed. The characteristic of uneven ripening is that the fruit's color is uneven, but its surface is smooth.

(a) powdery mildew (b) shrinkage (c) uneven ripening

Fig. 1. Three common diseases of strawberry

3 Image Preprocessing

As a result of noise, uneven illumination, and the transform of analog and digital signals of the CCD camera, the online images captured in the greenhouse under natural light are blurred in details ,such as its contours. Therefore, before image segmentation, the image preprocessing must be taken to eliminate the impact.

3.1 Image Denoising

To ensure that the edge of the strawberry fruit is not vague while image denoising, the median filtering is applied to remove image noise, in detail, each channel of the RGB image is removed noise through median filtering respectively, and then the denoised image can be obtained through the integration of the three denoised channels.

3.2 Uneven Illumination Effect Elimination

Strawberry images captured under natural light are greatly affected by uneven illumination. If this impact is not eliminated, the object may be segmented into disperse areas. In this paper, the “top-hat” transform is applied to eliminate the impact.

4 Image Segmentation

4.1 The Extraction and Transform of the Red and Green Area Templates

Figure 2 (a) shows the preprocessed strawberry image. The red area template can be got by 2R-G-B, as shown in Figure 2 (b), where R, G and B were the three channels of the color image in RGB space. Due to the uneven surface of the strawberry fruit, the fruit area contains many holes, which needs to be filled in order to get the closed template. Then, execute the OTSU operation, the initial fruit template can be obtained as shown in Figure 2 (c).

Conduct the logical AND operation between the template complementation and Figure 2 (a) to remove the fruit area in Figure 2 (a) while keeping the image outside of the template unchanged. The result is shown in Figure 3 (a). The green area template can be obtained by 2G-R-B, as shown in Figure 3 (b). Fig 3 (a) can be removed the green zone in similar way, and backfilled with the red template, then the result can be got as shown in Fig 3 (c). Now, the green area around the fruit is all black, and the fruit region is almost all white. The purpose of this series of operations is to separate strawberry fruit region from other objects, and create a favorable condition for the subsequent image segmentation.

(a)Preprocessed Image (b)Red region (c)Initial fruit template

Fig. 2. Initial fruit template

(a) Image erased part of the fruit (b) Green region template (c) Image backfilled the red area

Fig. 3. The separation between the fruit region and other objects

4.2 Strawberry Fruit Area Color Clustering

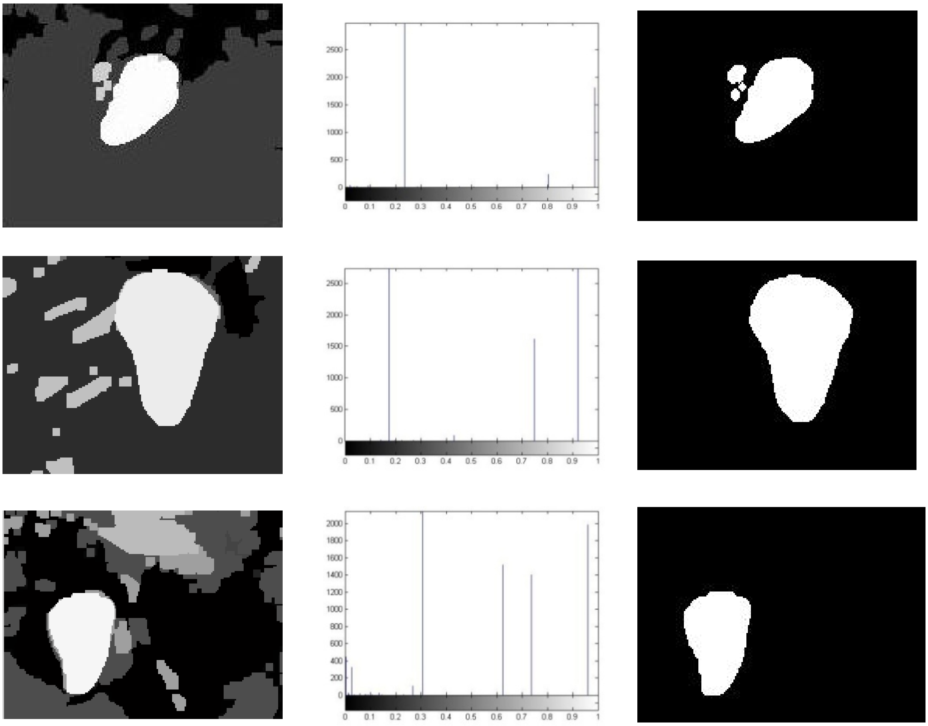
Because K-means cluster and fuzzy C-means cluster both need to set the cluster number initially, which is unknown for the image containing complicated background, it is infeasible to apply these two methods to segment the fruit image.

The mean shift algorithm is an adaptive probability density gradient ascent method of searching for the peak. It is as follows:1)Select the window size and initial position ;2)Calculate the gravity center of the window ;3)Move the center of the window to the center of gravity ;4)Repeat Step2 and Step 3 until the center of the window "convergence", that is, the moving distance of the window must be less than a certain threshold.Each pixel in the image can be set as the initial point, after performed the mean shift algorithm respectively, the initial points converged to the same point are considered as the same class. Therefore, the mean shift algorithm can be applied to image cluster segmentation.

Based on the above discussion, the mean shift segmentation algorithm is used in image segmentation, which make pixels of the strawberry fruit area cluster into the same class in color space, namely strawberry fruit regional pixels are all of the same value.

4.3 Target Area Extraction

After the operation of converting the image into grayscale and filling holses, execute the “open” operation on the image with circular structural element of a radius of 2 to remove the adhesion between the target and background. The experiments showed that: when the split threshold is 0.8, the grayscale threshold segmentation can obtain good results as shown in Figure 4; and then extract the largest connected area of the image to obtain strawberry fruit template. Combine this template and the preprocessed image by logical AND operation, then the image of the complete strawberry fruit area can be obtained.



(a) Gray image (b) Histogram of (a) (c) threshold segmentation image

Fig. 4. Threshold segmentation image

5 Strawberry Fruit Image Segmentation Algorithm

After the above discussion on the image preprocessing and image segmentation method, here are the strawberry fruit image segmentation algorithm steps, including image reading, image preprocessing, image segmentation（shown in Fig 5）.

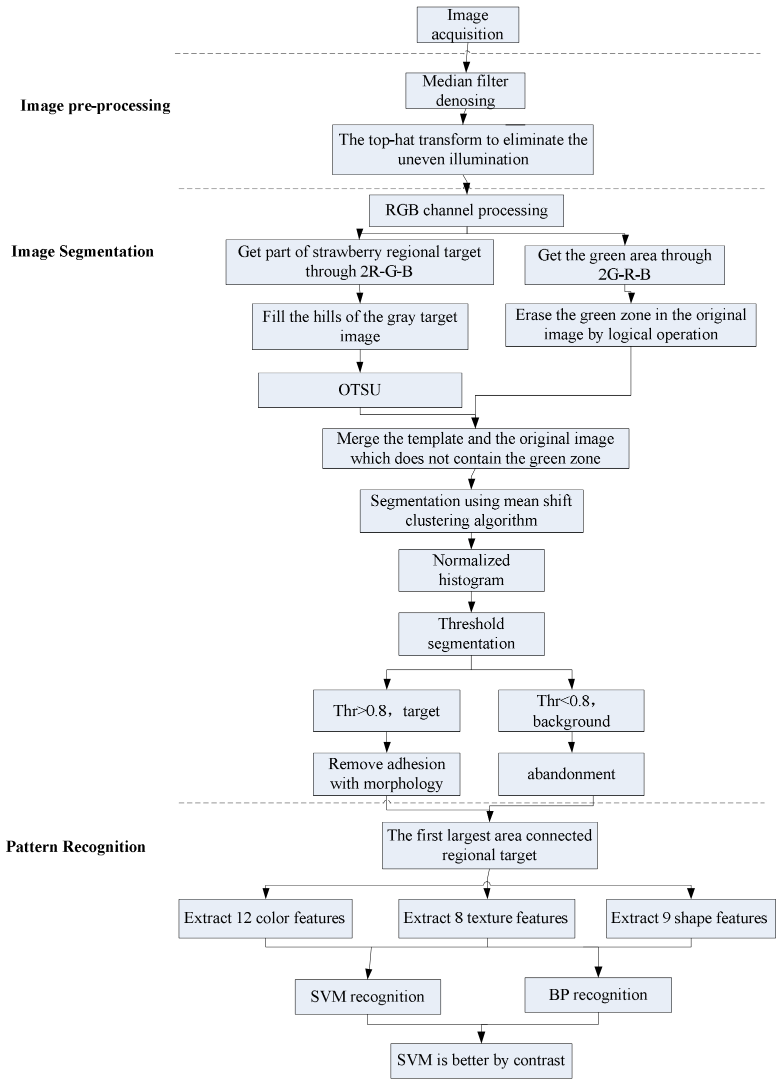


Fig. 5. The flow chart of the segmentation algorithm

6 Segmentation Results

Figure 6, Figure 7, Figure 8, and Figure 9 are the segmentation results of the above strawberry image segmentation algorithm, which corresponds to normal strawberry, powdery mildew of strawberry, shrinkage of fruit disease strawberries and uneven ripening disease strawberries respectively. Each subgraph corresponds to the result after the corresponding step of the above segmentation algorithm. The algorithm can get effective segmentation results under natural illumination.

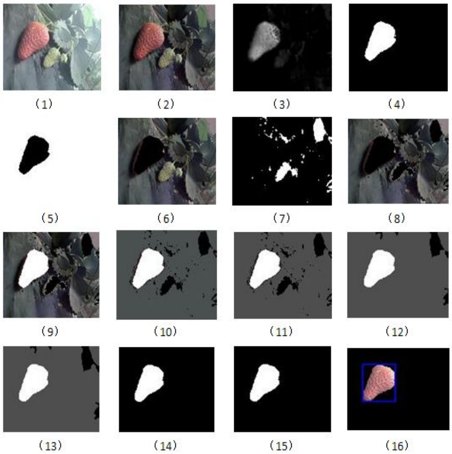


Fig. 6. Segmentation results of the normal strawberry

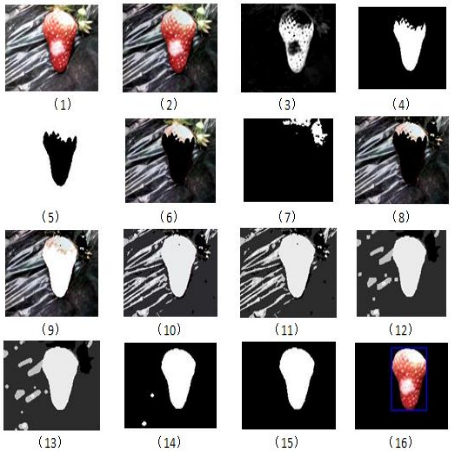


Fig. 7. Segmentation results of the powdery mildew strawberry fruit

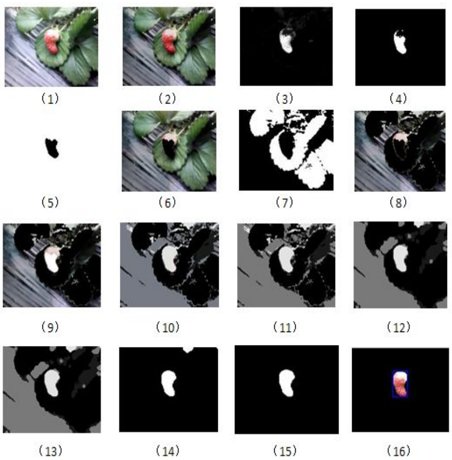


Fig. 8. Segmentation results of the strawberry fruit with the shrink disease

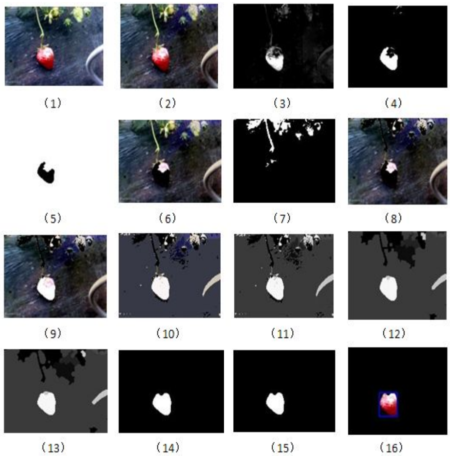


Fig. 9. Segmentation results of the strawberry fruit withthe uneven ripening disease

7 Experimental Schemes

7.1 Data Selection

a)Data access

The color, texture and shape feature were extracted from four kinds of strawberry respectively, including 12 kinds of color characteristic parameters,8 kinds of texture parameters, nine kinds of shape characteristic parameters, so each sample corresponds to a set of characteristic data of a total of 12 +8 +9 = 29 kinds of characteristic parameters. The sample size of each type of disease strawberry and normal strawberry is 20. Set category labels are as follows: normal strawberry labels 1, powdery mildew strawberry labels 2,fruit shrink disease strawberry labels 3, and the uneven ripening disease strawberry labels 4.The two main equations are:

mean $\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$

standard deviation $\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$

b)Division of the data

Establishing training and testing are two essential processes while using BP neural network and SVM to build the classifier. So data should be divided into two parts, one for training and one for the test. 3/4 of the data are used for training, 1/4 of the data are used for testing. Select 60 sets of data randomly as a training data set, the remaining 20 sets of data as a test data set from the data set.

c)Data preprocessing

The physical meaning of the color, texture and shape characteristic value is different, not comparable, and their range changes, so there is a big numerical difference among them. Data normalization is necessary, or it will lead to the classifier learning slows down or even does not converge. In this study, each dimensional feature data were normalized preprocessed, used in a normalized mapping shown as follows:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}}$$

In which x is the corresponding eigenvalue of a certain characteristic properties,x_{max}, x_{min}shows the largest and the smallest eigenvalue of the properties. While classifier training is completed, the input data for the test needs to be normalized in order to enter the classifier.

7.2 BP Neural Network Training and Prediction

According to Kolmogorov theorem, using a three-layer BP neural network(29 × 59 × 4)for the classifier, in which the hidden layer activation function is the non-symmetric Sigmoid function: $f(x) = \frac{1}{1+e^{-x}}$, the output layer activation function is a linear function f (x) =x, and set a target error e = 0.01. Use the training data to train the BP neural network, to make the BP neural network can predict the nonlinear function output.

Use trained BP neural network to predict the test data, and use the predicted output and actual output of the BP neural network to analyze the classification ability of BP neural network.

7.3 SVM Training and Prediction

Implementation of SVM in this experiment uses the libsvm toolbox, the kernel function uses the radial basis kernel function. The training data is used to train SVM classifiers to get the SVM classification model.

Use the trained SVM classifier to make the label prediction of test data, and analyze the Classification ability of SVM with predicted SVM output and actual output.

8 Conclusion

Image segmentation of the strawberry fruit of the three common diseases consists of two main parts: the first part is image pre-processing, including the application of median filtering to remove the noise from image acquisition, and the application of top-hat transform to eliminate the effects of uneven illumination; the second part is image segmentation, using gray morphology, OTSU algorithm, the mean shift segmentation algorithm to segment the fruit from the image. These experimental results indicate that image processing algorithm can obtain effective segmentation results to the online images of the three common kinds of diseased strawberries and nomal ones under natural illumination.

This paper uses digital image processing, pattern recognition technology to extract the integrated fruit of disease strawberries. Based on parameters of color characteristic, texture features and shape features, use BP neural network and support vector machine to make classification and recognition for disease strawberry. The experimental results show that support vector machine has a higher recognition rate than the BP neural network so that support vector machine is used as the classifier for disease strawberry.

References

1. Zhou, T., Zhang, T., Yang, L., et al.: Comparison of two algorithms based on mathematical morphology for segmentaion of touching strawberry fruits. Transactions of the CSAE 23(9), 164–168 (2007); (in Chinese with English abstract)
2. Xie, Z., Zhang, T.: A new method of segmentation of strawberry image. Journal of China Agricultural University 11(1), 84–86 (2006); (in Chinese with English abstract)
3. Zhang, T., Zhou, T.: Strawberry harvesting robot. Journal of China Agricultural University 9(4), 65–68 (2004)
4. Xie, Z., Zhang, T., Zhao, J.: Ripened Strawberry Recognition Based on Hough Transform. Transactions of the Chinese Sociey for Agricultural Machinery 38(3), 106–109 (2007)
5. Yu, S., Liu, R.: OpenCV course. Beijing University of Aeronautics and Astronautics Press, Beijing (2009)
6. Zhang, Y.: Image Processing. Tsinghua University Press, Beijing (2008)
7. Zhang, Y.: Image Analysis. Tsinghua University Press, Beijing (2008)
8. Liu, H., Shen, J., Guo, S., et al.: Digital Image Processing Using Visual C++. China Machine Press, Beijing (2010)