THE OPTIMAL CHOICE OF THE CHARACTERISTIC WAVELENGTHS IN SPECTRAL IMAGING FOR CUCUMBER DOWNY MILDEW

/

黄瓜霜霉病光谱图像特征波长优化选取

As. Prof. PhD.Eng. Geng C.X.[1)], Ms. Stud. Eng. Wang P.[1)], As. Prof. PhD.Eng. Qin P.L.[2)], Eng. Zhang W.B.[3)], As. Prof. PhD.Eng. Wang P.B.[*1)]

[1)]Robotics and Microsystems Centre，Soochow University，Suzhou/China;
[2)] Wisdom Agriculture College, Suzhou Polytechnic Institute of Agriculture, Suzhou/China;
[3)]Suzhou Agricultural Machinery Technology Promotion Station, Suzhou/China;
Tel:+8618662299887; E-mail: pbwang@suda.edu.cn

Keywords: *cucumber Downy Mildew; spectral image; characteristic wavelength; sensitive waveband*

ABSTRACT

Aiming at the spectral image acquisition of cucumber downy mildew, the study proposed the optimal selection method of spectral imaging characteristic wavelengths combining the mean-variance analysis and analysis of variance. In this study, the variation degree and range of reflection spectrum was quantitatively measured for both healthy leaves and disease spots in cucumber plants. The analysis showed that the three sensitive wavebands (493 nm - 533 nm, 561 nm - 653 nm, and 689 nm - 698 nm) could properly reflect the inter class difference in spectral reflectance between the dorsal part of lesion and healthy leaves. When the selected combination of spectrum was used as the reference for disease spectral imaging acquisition, it showed that the choice of sensitive wavebands can better reflect the difference in the characteristics between lesion and healthy leaves. It has laid the foundation for obtaining the effective spectral images of disease and simplifying the algorithm for image identification.

摘要

针对黄瓜霜霉病的光谱图像获取，提出了基于均值方差法和方差差异分析相结合的光谱图像特征波长优化选取方法，定量衡量了黄瓜作物健康叶片和病斑的反射光谱差异程度及差异范围，提取了 493 nm - 533 nm、561 nm - 653 nm、689 nm - 698 nm 三个敏感波段，分析表明上述波段能较好反映病斑与健康叶片正面的类间光谱反射差异，并将选定的光谱组合作为病害光谱图像获取的参考依据，表明敏感波段的选择能较好地反映病斑与叶片间的特性差异，为获取病害有效光谱图像和简化图像识别算法奠定了基础。

INTRODUCTION

The characteristics of the reflectance spectrum of any plant vary under the influence of disease and biological stress; however, there are rules for the variation (*Manjunath et al., 2014; Shi et al., 2016; Bandaru et al., 2016*). There have been many studies for detecting plant diseases and biological stress from the perspective of spectral analysis and spectral imaging at home and abroad. For instance, (*Wei et al., 2014*) used three wavelengths (518, 711 and 980 nm) to differentiate the maturity in persimmons into four stages: immature, early mature, fully mature and over mature; and the accuracy can be as high as 95.3%. (*Sirisomboon et al., 2013*) used near-infrared spectroscopy (NIRS) to detect rice mould infections and found that the error rate could be reduced to 18.723%, and the deviation to 4.613%. (*Mahlein et al., 2013*) studied the waveband between 450 and 950nm and developed a specific spectral disease index (SDIS) for crop disease detections, and the result showed that the index is highly accurate and sensitive. (*Min et al., 2005*) found the wavelengths related to chlorophyll are 450, 550, 660 and 719 nm, and those related to protein are 2054 and 2180 nm. (*Kong et al., 2013*) used the spectroscopic analysis together with image processing and stoichiometric methods, established a least squares support vector machine (LSSVM) model for the early diagnosis of grey mould in stems of tomatoes. The result suggested that using the hyperspectral imaging technology to diagnose the grey mould disease of tomato stems at early stage is feasible. (Jiang et al., 2013) used three sensitive wavelengths (532, 610 and 700nm) to measure chlorophyll content in tomato leaves and acquired outstanding results.

The unique reflectance characteristics at different wavelengths in a plant's spectrum can be effectively transformed into the expression of target information and detection information of spectral imaging via analyses (*Sun et al., 2016; Ma et al., 2016; Story et al., 2010*).

For efficient spectral imaging of plant diseases in complex environments, it requires analyses of their spectral properties and selection of characteristic wavebands (*Ye et al., 2016; Polat et al., 2014*). The aim of this study was to use spectral analysis methods to determine the optimal spectral wavebands that could diagnose the downy mildew disease, to increase the grey scale contrast of the corresponding spectral image object and background, and to provide a useful basis for the subsequent image processing.

MATERIAL AND METHODS

- **Experiment Materials**

The cucumber for experiment was the cultivar "Zhongke 958", grown with a trailing method. Spectral data collection was conducted in the solar greenhouse at Shang-Zhuang Experimental Station of China Agricultural University in November 2015. Owing to the need for supportive materials in trailing cultivation, such as bamboo poles, nylon slings and clamps, there were a considerable amount of background information in the spectral imaging acquisition results. In order to study the cucumber downy mildew and the spectral characteristics in cucumbers, leaves and complex background, 11 types of samples, including the dorsal and ventral sides of leaves with lesions caused by downy mildew, the dorsal and ventral sides of the healthy leaves, cucumbers, dried leaves, bamboo poles, green nylon slings, black clamps, dry soil and wet soil, were selected as the main objects for spectral data collection.

- **Choice of Instrument**

The spectral reflectance data was collected using the FieldSpec Pro VNIR handheld spectrometer (Analytical Spectral Devices, Inc., USA). The spectrometer measures wavelengths ranging between 325 nm and 1075 nm, and the spectral resolution is at 1 nm. Prior to data collection, the standard whiteboard was used to calibrate and eliminate systematic errors. After the calibration, the spectral reflectance was measured in the 11 types of samples. The sensor probe was vertically pinned and the sampling height was 20 mm. The measured object was at the same level as the standard white board in measurement. 20 probe points were randomly selected on each sample, with each point being tested for 5 times. The mean of the 5 measurements was then used as the final measurement for each testing point. The raw spectral reflectance data of each measured point was acquired at the waveband between 325 nm and 1075 nm. The measurements were saved in Microsoft EXCEL format and further analyzed using Matlab.

- **Data processing method**
- **Spectral reflectance analysis**

A highly consistent spectra and an identical trend of reflectance was observed in the dorsal and ventral sides of lesion and healthy leaves. In the meantime, the distribution of healthy and lesion parts was spatially intersected. Here, a mean-variance analysis was used for an emphasis on analysing the volatility of the spectra for the dorsal and ventral sides of lesion and healthy leaves. The method was described as follows:

By defining the estimated value for the mean reflectivity of sample x is:

$$\hat{\mu}_x = \frac{1}{n}\sum_{m=1}^{n} x_m \quad (1)$$

where $\hat{\mu}_x$ - is the estimated value for the mean reflectivity of sample x; n - number of sampling points;

X_m - the reflectivity of the sampled wavelength for the m-th sampling point of sample x.

One could prove the estimate $\hat{\mu}_x$ is unbiased. Thus:

$$E(\hat{\mu}_x) = \mu_x \quad (2)$$

Where μ_x- the true value of the mean reflectivity of sample x.

When the samples are uncorrelated and the estimates are consistent, then:

$$\lim_{n\to\infty} E\mu_x - \hat{\mu}_x = 0 \quad (3)$$

When the true value of the mean reflectivity of a given sample is known, its variance can be estimated by the following equation:

$$\delta_x^2 = \frac{1}{n}\sum_{m=1}^{n}(x_m - \mu_x)^2 \quad (4)$$

Where δ_x^2 - the variance of sample x.

Here, the estimate is unbiased and consistent. However, if the true value of the mean is unknown, then $\hat{\mu}_x$ should be used to replace μ_x in the estimation, and the estimation would be biased.

Nevertheless, an unbiased variance can be estimated with the following formula:

$$\delta_x^2 = \frac{1}{n-1}\sum_{m=1}^{n}(x_m - \mu_x)^2 \quad (5)$$

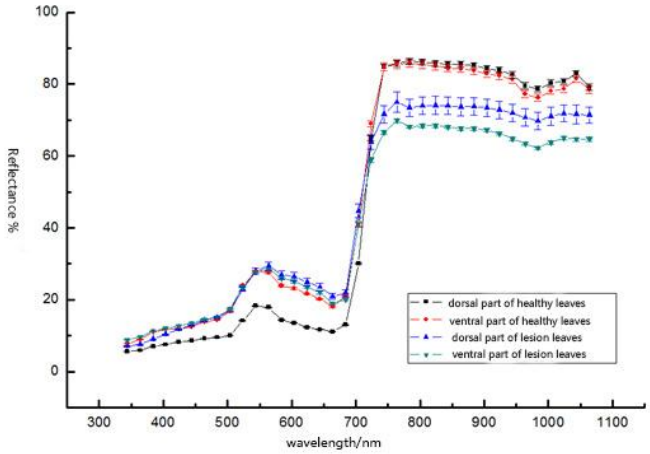

Fig. 1 - The spectral fluctuate characteristic of double sides of the disease and the healthy leaves

When the Savitzky-Golay smoothing (SG) (*Li et al., 2015*) was used to eliminate random noises, as shown in Fig. 1, the fluctuation of sample data was smaller and the sample variance was also smaller; thus, the reflectance curves of the four types of sample could be distinguished with ease. In the visible band, the maximum spectral reflectance for the dorsal and ventral sides of healthy leaves was found at 760 nm (94.35% and 93.03%, respectively). The maximum reflectance for the dorsal and ventral sides of lesion leaves was at 763 nm (75.51% and 70.30%, respectively). In near infrared band, the minimum spectral reflectance at the near infrared band for the dorsal and ventral sides of healthy leaves was at 988 nm (77.77% and 75.22%, respectively). For the dorsal and ventral sides of lesion leaves, the minimum reflectance was at 974 nm (68.91% and 61.58%, respectively). The maximum spectral reflectance for the dorsal part of the lesion leaves was 75.22% at 1068 nm, and for the ventral part was 68.96% at 823 nm. The result from the qualitative analysis showed that the differentiation was more obvious at near infrared band than at the visible light band, although the spectral reflectance was less distinguishable between them according to the extreme values of their spectral reflectance.

- **Selection on the characteristic sensitive wavebands**

Based on the results from the qualitative analysis, this study used a wavelength selection method based on the analysis of variance to quantitatively analyse the level and range of differences of the reflectance spectra in healthy and lesion leaves in cucumbers, and to confirm the ideal characteristic sensitive wavebands that may diagnose the cucumber downy mildew (*Camargo et al., 2008; Yao et al., 2013*). Here, M was defined as the number of categories. For clarification, Formula (1) was reformatted as Formula (6):

$$\bar{x}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} x_{ij} \quad (6)$$

where $\bar{x}_j$ - the mean reflectivity of the spectra in *j* category; n_j - the number of samples in *j* category;

x_{ij} - the reflectivity of the spectra of the *i*-th sample in *j* category.

The mean reflectivity for all samples can be expressed by:

$$\bar{x} = \frac{1}{Mn_j}\sum_{i=1}^{M}\sum_{j=1}^{n_j} x_{ij} \quad (7)$$

where $\bar{x}$ - the mean reflectivity for all samples; *M* - the number of categories.

The error sum of squares S_E represents the variation in spectral reflectance of different categories and is expressed by Formula (8). As shown in Fig. 2, the error sum of squares S_E for the spectrum of the four samples was of relatively high spectral reflectance at the bands of 500 nm and above, which was the most obvious on the dorsal part of lesion leaves.

$$S_E = \sum_{i=1}^{M} \sum_{j=1}^{n_j} (x_{ij} - \bar{x}_j)^2 \tag{8}$$

where S_E - the error sum of squares.

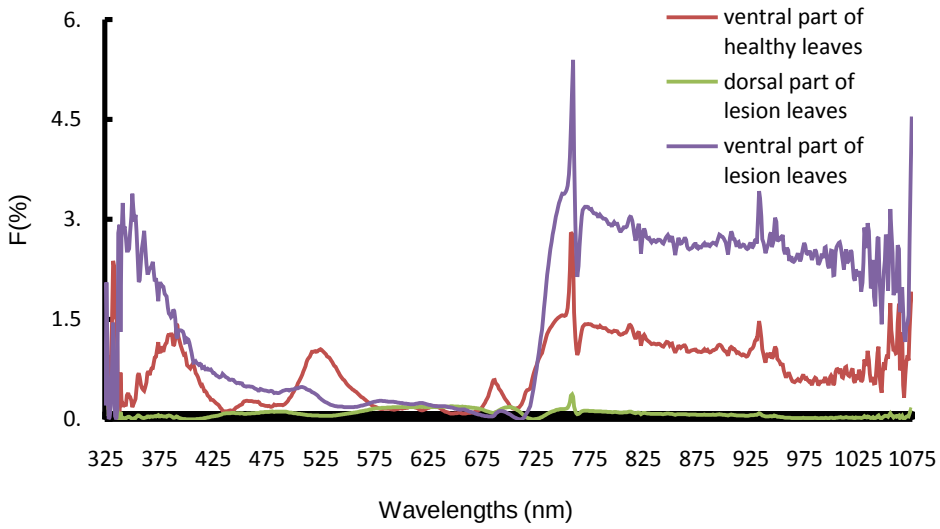

Fig. 2 - The error sum of squares SE of reflectance spectrum for the double sides of healthy cucumber leaf and disease area

Here, S_A was defined as the efficacy sum of squares, representing the difference of spectral reflectance among different categories, expressed by Formula (9).

The curves of the efficacy sum of squares S_A were shown in Fig. 3, where a consistency of the reflectance trend among the four samples can be seen.

$$S_A = \sum^{M} \sum^{n_j} (\bar{x}_j - \bar{x})^2 \tag{9}$$

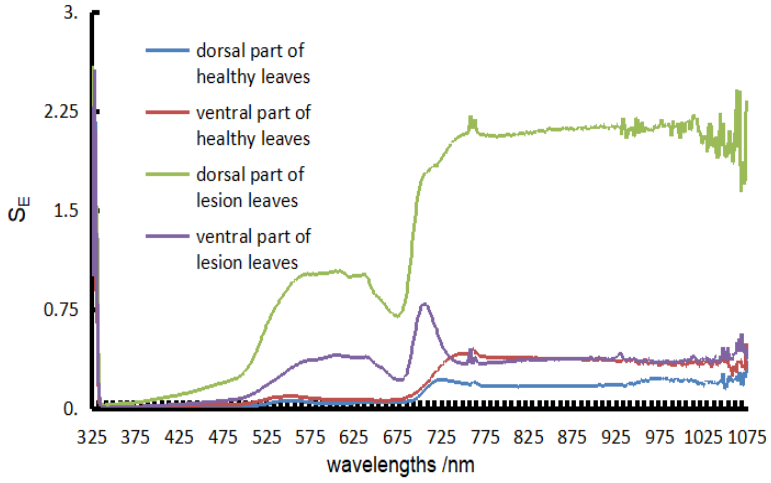

Fig. 3 - The efficacy sum of squares SA of reflectance spectrum for the double sides of healthy cucumber leaf and disease area

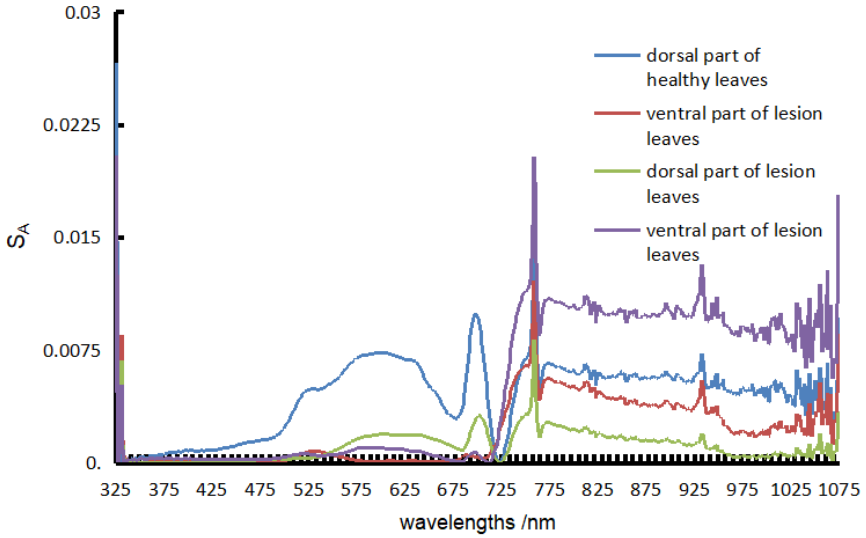

Fig.4 - The curve of F value of reflectance spectrum for the double sides of healthy cucumber leaf and disease area

Hence, for a given wavelength, the distinctiveness of the reflectance spectrum (F) in a sample can be shown as:

$$F = \frac{S_A}{S_E} \times 100\% \tag{10}$$

Where F is the distinctiveness of the reflectance spectrum.

The F value represents a relative value for the classification error and the random error.

The greater value means the greater difference in spectral reflectance between the categories.

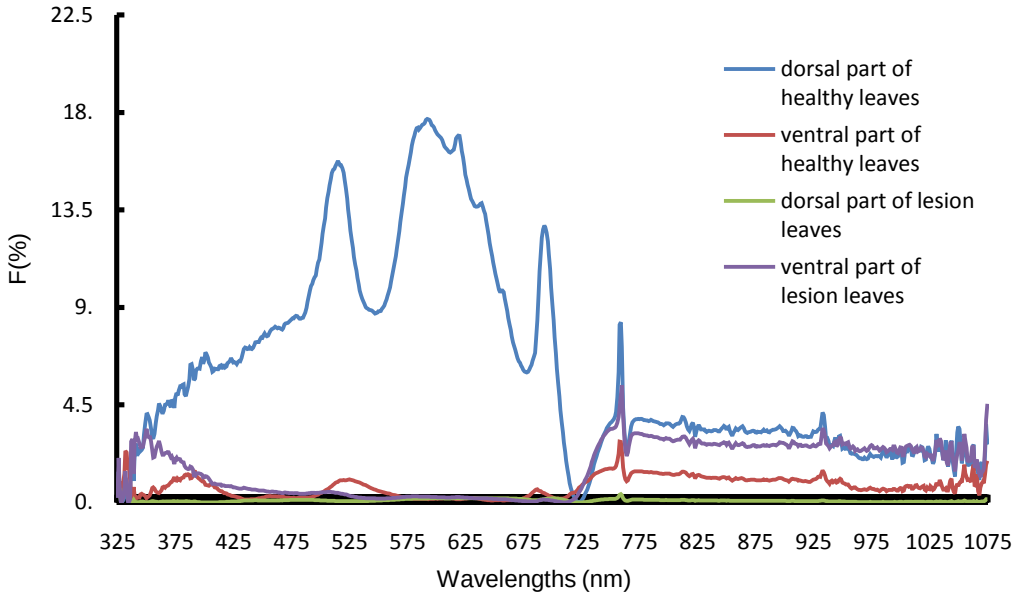

Fig. 5 - The curve of F value of reflectance spectrum for the back side of healthy cucumber leaf and the double sides of disease area

The raw spectral data was analyzed using the analysis of variance, and the curves for the F value variation among the reflection spectra of both ventral and dorsal parts of healthy and lesion leaves were shown in Fig. 4. By and large, the greatest difference was found in the spectral reflection in the dorsal part of healthy leaves, the ventral part of healthy leaves, and both dorsal and ventral parts of lesion leaves. In particular, the most obvious difference in spectral reflectance was within the visible light band; the curve for the F value variation showed a certain degree of characteristics and distinctiveness. The most significant

differentiation was seen at the waveband of 378 nm - 705 nm, with a minimum F value of 5.17%. The peak value was at 592 nm and the F value was at its maximum, 17.70%. The wavebands with the F value >10 were: 493 nm - 533 nm, 561 nm - 653 nm, and 689nm - 698 nm; the analyses suggested these wavebands can better reflect the difference of spectral reflection between the lesion and healthy leaves, and they can be regarded as the sensitive wavebands for the raw spectral data. However, as shown in Fig. 5, the characteristics of spectral reflectance in the ventral part of healthy leaves and the lesion leaves was unobvious in those above three wavebands

REASULTS

Spectrum Capture

The curves for the reflection spectra of all the 11 types of samples are shown in Fig. 6. The spectral reflectance of each sample was calculated as the mean of the spectral reflectance of all the pixels in the samples, overcoming the limitation of using partial observation to represent the whole picture in the traditional spectrum method. Using the measured spectral reflectance of the 11 samples and based on the general experience in division of wavebands in spectroscopy, the selected spectral detection region was firstly divided into two empirical sub-bands: the visible spectrum (380 - 780 nm) and the near-infrared spectrum (780 - 1025 nm). The rhythm of the spectral curves was remarkable (Fig. 6); a typical green-plant spectral curve could be found in every cucumber organ; that is, a strong reflection peak of chlorophyll around 550 nm, a trough around 680 nm, that is mainly caused by strong absorption of chlorophyll, and a high reflection plateau at 700 - 900 nm. The reflectance curves for non-plant samples were significantly different, indicating a great disparity.

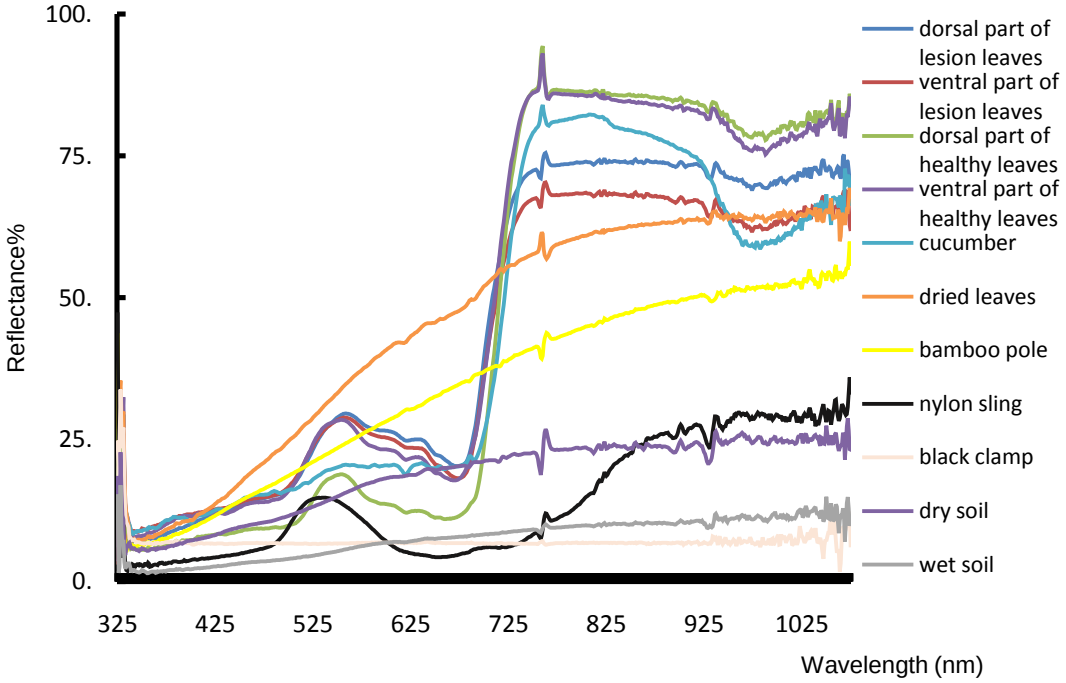

Fig. 6 - Spectral reflection curves of the 11 kinds of samples

Characteristic spectral imaging

Owing to the limitation of the greenhouse environment and the establishment of image acquisition system, the main object for the image detection was the dorsal part of the leaf. From the section "Selection on the characteristic sensitive wavebands", we can see that the wavebands with the F value >10 were: 493 nm - 533 nm, 561 nm - 653 nm, and 689nm - 698 nm; the analyses suggested these wavebands can better reflect the difference of spectral reflection between the lesion and healthy leaves, and they can be regarded as the sensitive wavebands for the raw spectral data. The above three main bands were selected as the main spectral band for the disease information acquisition (*Singh et al., 2012*).

The spectral imaging was constructed through a black and white camera with light filters, and the filter types were one with a central wavelength of 506 nm, half-bandwidth of 35 nm and a maximum transmittance of 90%; the other with a central wavelength of 625 nm, half-bandwidth of 20 nm and a maximum

transmittance of 80%; and the third with a central wavelength of 680nm, half-bandwidth of 20 nm and a maximum transmittance of 80%. The collected images were shown in Fig. 7.

The differences between lesion and healthy leaves were demonstrated clearly in the images and the use of grayscale images at characteristic wavebands as the raw images for disease extraction greatly simplified the algorithm for disease extraction and reduced computing amount, and thus improved the timeliness of image processing for the spray robot operation to make decisions on-site.

A.Spectral image at 506 nm B.Spectral image at 625 nm C.Spectral image at 680

Fig. 7 - The spectral image of cucumber leaves

CONCLUSIONS

- By establishing the optimal detection waveband for healthy and lesion cucumber leaves, this study qualitatively analysed the differentiation of both in the visible light and near infrared wavebands. The result showed that the difference was more distinct in the near-infrared band than in the visible light band, but less distinct in the spectral reflectance, from the perspective of the extremums of spectral reflectance.
- The F value for the distinctiveness of the reflectance spectrum in samples was calculated through the analysis of variance. The three wavebands with the F value >10 (i.e., 493-533 nm, 561-653 nm and 689-698 nm) were selected, and the result suggested these wavebands can properly reflect the difference in the spectral reflection between the lesion and healthy leaves, and they can be regarded as the sensitive wavebands for the raw spectral data.
- In the examination of the spectral imaging for the cucumber plants in the three wavebands via the method of using a black and white camera with light filters, the result showed that the selected sensitive bands can properly reflect the differences in the inner components and external features in the healthy and lesion cucumber leaves and highlight the obvious differences in spectral imaging, hence simplify the complexity of the algorithm for image processing, and improve the timeliness of the algorithm.

ACKNOWLEDGEMENT

The study was supported by two National Key Research and Development Program of China (No. 2016YFD0700600)(No. 2017YFD0701500), and 2017 high-end training program for professional leaders of teachers in higher vocational colleges in Jiangsu Province(NO. 2017GRFX044).

REFERENCES

[1] Bandaru V, Daughtry C.S., Codling E.E. et al., (2016), Evaluating leaf and canopy reflectance of stressed rice plants to monitor arsenic contamination, *International Journal of Environmental Research and Public Health*, Vol.13, pp.606-609, Pubmed, Beltsville/U.S.A.;

[2] Camargo A，Smith J.S., (2008), An image-processing based algorithm to automatically identify plant disease visual symptoms, *Biosystems Engineering*, Vol.102, pp.9-21, Academic Press INC Elsevier Science, San Diego/U.S.A.;

[3] Jiang W.J., Sun M, (2011), Research on predicting modeling for chlorophyll contents of greenhouse tomato leaves based on multi-spectral imaging, *Spectroscopy and Spectral Analysis*, Vol.31, pp.758-761, The Chinese Optical Society, Beijing/China;

[4] Kong W.W., Yu J.J., Liu F. et al., (2013), Early diagnosis of gray mold on tomato stalks based on hyperspectral data, *Spectroscopy and Spectral Analysis*, Vol.33, pp.733-736, The Chinese Optical Society, Beijing/China;

[5] Li X.L., Yi S.L., Zheng Y.Q., et al., (2015), Analysis of phylogenetic relationships of main citrus germplasms based on FT-IR spectra of petals, *Intelligent Automation & Soft Computing*, Vol.21, pp.295-304, Taylor & Francis INC., Philadelphia/U.S.A.;

[6] Ma W, Wang X, Qi L.J., (2016), Study on spectrum recognition of leaf spot disease based on visible-near infrared spectroscopy, *Spectroscopy and Spectral Analysis*, Vol.36, pp.121-123, The Chinese Optical Society, Beijing/China;

[7] Mahlein A.K., Rumpf T, Welke P, et al., (2013), Development of spectral indices for detecting and identifying plant diseases, *Remote Sensing of Environment*, Vol.128, pp.21-30, Elsevier Inc., Malden/U.S.A.;

[8] Manjunath K.R., Kumar A, Meenakshi M. et al., (2014), Developing spectral library of major plant species of Western Himalayas using ground observations, *Journal of the Indian Society of Remote Sensing*, Vol.42, pp.201-216, Springer (India) Private Ltd., Ahmedabad/India;

[9] Min M, Lee W.S., (2005), Determination of significant wavelengths and prediction of nitrogen content for orange, *Transactions of the ASABE*, Vol.48(2), pp.455-461, American society of Agricultural and Biological Engineers, Michigan/U.S.A.;

[10] Polat I，Baysal O，Mercati F, (2014), Characterization of pseudoperonospora cubensis isolates from Europe and Asia using ISSR and SRAP molecular markers, *European Journal of Plant Pathology*, Vol.139, pp.641-653, Springer, Dordrecht/Netherlands;

[11] Shi T.Z., Wang J.J., Chen Y.Y., et al., (2016), Improving the prediction of arsenic contents in agricultural soils by combining the reflectance spectroscopy of soils and rice plants, *International Journal of Applied Earth Observations and Geoinformation*, Vol.52, pp.95-103, Elsevier B.V., Wageningen/Netherlands;

[12] Singh C.B., Jayas D.S., Paliwal J, et al., (2012), Fungal damage detection in wheat using short-wave near-infrared hyperspectral and digital colour imaging, *International Journal of Food Properties*, Vol.15, pp.11-24, Taylor & Francis INC., Philadelphia/U.S.A.;

[13] Sirisomboon C.D., Putthang R, Sirisomboon P, (2013), Application of near infrared spectroscopy to detect aflatoxigenic fungal contamination in rice, *Food Control*, Vol.33, pp.207-214, Elsevier Science Ltd., London/U.K.;

[14] Story D, Kacira M, Kubota C. et al., (2010), Lettuce calcium deficiency detection with machine vision computed plant features in controlled environments, *Computers and Electronics in Agriculture*, Vol.74, pp.238-243, Elsevier Science B.V., Wellington/Netherlands;

[15] Sun J, Lu X. Z., Zhang X.D. et al., (2016), Identification of red bean variety with probabilistic GA-PNN based on hyperspectral imaging, *Transactions of the Chinese Society for Agricultural Machinery*, Vol.47, pp.215-211, Chinese Society for Agricultural machinery, Beijing/China;

[16] Wei X, Liu F, Qiu Z.J. et al., (2014), Ripeness Classification of Astringent Persimmon Using Hyperspectral Imaging Technique, *Food and Bioprocess Technology*, Vol.7, pp.1371-1380, Springer, Newyork/U.S.A.;

[17] Yao H.B., Hruska Z, Kincaid R, et al., (2013), Detecting maize inoculated with toxigenic and atoxigenic fungal strains with fluorescence hyperspectral imagery, *Biosystems Engineering*, Vol.115, pp.125-135, Academic Press INC Elsevier Science, San Diego/U.S.A.;

[18] Ye H.J., Lang R, Liu C.Q., et al., (2016), Recognition of cucumber downy mildew disease based on visual saliency map, *Transactions of the Chinese Society for Agricultural Machinery*, Vol.47(5), pp.270-274, Chinese Society for Agricultural machinery, Beijing/China.