Identification of Powdery Mildew and Stripe Rust in Wheat Using Hyperspectral Imaging

YAO Zhi-feng[1,2,3]，LEI Yu[1,2,3]，HE Dong-jian[1,2,3]*

1. College of Mechanical and Electronic Engineering，Northwest A&F University，Yangling 712100，China
2. Key Laboratory of Agricultural Internet of Things，Ministry of Agriculture and Rural Affairs，Yangling 712100，China
3. Shaanxi Key Laboratory of Agricultural Information Perception and Intelligent Service，Yangling 712100，China

Abstract Powdery mildew and stripe rust are two of the most prevalent and destructive wheat diseases causing severe decreases in wheat yield in China. It is necessary to quantitatively identify different diseases for spraying specific fungicides. In this study，a line-scanning hyperspectral imaging system (ImSpector V10E) was utilized to capture spectral and imagery information of wheat leaves infected by powdery mildew，stripe rust and normal leaves. Based on 320 hyperspectral images，strong spectral reflectivity responses were discovered at the bands of 550～680 nm in the wheat leaves infected with powdery mildew and stripe rust after the savitzky-golay (SG) smoothing method. To reduce the dimensionality of the spectral matrix，3，6 and 30 variables were extracted as sensitive wavelengths from full spectra for different diseases using X-loadings of principal component analysis (PCA)，successive projections algorithm (SPA)，and competitive adaptive reweighted sampling (CARS)，respectively. Least squares support vector machine (LS-SVM) and extreme leaning machine (ELM) were applied to build identification models using full spectra and sensitive wavelengths extracted by X-loadings of PCA，SPA and CARS to distinguish powdery mildew，stripe rust and normal leaves. The accuracy rates of all the models in the calibration set and test set were above 94.58%. Among these models，the ELM classification model combined with X-loadings of PCA had the best performance，with accurate identification rates of 99.18% on the calibration set and 100% on the test set. Moreover，this model was simple in structure with only three variables (560，680 and 758 nm). Meanwhile，the microstructure of three kinds of wheat leaves were also studied. Although the infection mechanisms of these two diseases were slightly different，they both destroyed the mesophyll cells，reduced chlorophyll content and photosynthesis markedly. The string of changes leaded to weakened light absorption but increased reflectivity in the visible light band. Thus，the results indicated the potential for the rapid and non-destructive detection of wheat diseases by hyperspectral imaging，which could help to develop online multispectral detection system for different kinds of plant diseases.

Keywords Powdery mildew；Stripe rust；Hyperspectral imaging；Sensitive wavelengths；Identification model

中图分类号：O433.4 **文献标识码**：A **DOI**：10.3964/j.issn.1000-0593(2019)03-0969-08

Introduction

Powdery mildew caused by *Blumeria graminis* and stripe rust induced by *Puccinia striiformis f. sp. tritici* are two of economically diseases that have compromised the safe production of wheat in China for a long time[1]. They each have a wide incidence range，strong infectivity and epidemicity，and

Received：2018-08-27；**accepted**：2019-01-06

Foundation item：Key Industrial Chain Projects of Shaanxi Province (2015KTZDNY01-06)，Key Science and Technology Program of Shaanxi Province (2017NY-104)，Science and Technology Plan Project of Yangling Demonstration Area (2015NY-10)

Biography：YAO Zhi-feng，(1984—)，female，doctoral candidate，College of Mechanical and Electronic Engineering，Northwest A&F University e-mail：yzf8466@163.com * Corresponding author e-mail：hdj168@nwsuaf.edu.cn

cause severe damage. What's more, they are frequently complicated by each other. The pathogenesis of the two wheat diseases are different, and different bactericides should be applied for prevention. Incorrect use of pesticides poses a hidden danger to food security, and pesticide residues can also pollute the soil and underground water.

Common methods for the diagnosis and detection of plant diseases include visual plant disease estimation by human raters, microscopic evaluation of morphology features to identify pathogens, as well as molecular, serological, and microbiological diagnostic techniques[2]. But these methods demand experienced individuals with well-developed skills in disease detection and are thus subject to human bias. Consequently, a convenient, accurateand simple method is urgently needed for qualitative identification of different kinds of diseases at wheat seedling stage.

Hyperspectral imaging is an emerging technique, which combines conventional imaging and spectroscopy to acquire both spatial and spectral information from the detected target in the same time. Because of the combined features, it has been successfully used for agricultural product quality evaluation[3], food defect inspection[4], plant disease identification[5] and so on. Most scholars have carried out experimental studies on the common disease types in crops such as wheat[6], barley[7], rice[8], sugar beet[9] and oilseed rape[10], attempting to clarify the spectral response features of disease and search for optimal models to improve the exactness of disease recognition.

With regard to wheat diseases, some scholars have studied the spectral response features of stripe rust and powdery mildew. For instance, a study conducted by Zhao et al.[11] showed that the spectral reflectance of wheat leaves infected by stripe rust was positively correlated with disease severity at the bands of 550～680 and 750～1 300 nm. Zhang et al.[12] discovered a strong spectral response at 520～720 nm of wheat powdery mildew. It is noteworthy that the characteristic bands of wheat stripe dust and powdery mildew are greatly overlapped in some spectral waves, but at this present time, to the best of our knowledge, there are few reports on the detection of these two kinds of wheat diseases. Liang et al.[13] selected the sensitive bands of powdery mildew at 519, 643, 696, 764, 795, and 813 nm, and the sensitive bands of stripe rust at 494, 630, 637, 698, 755, and 805 nm, to establish the discriminant model for recognizing powdery mildew and stripe rust, with the accuracy of 92%. However, this method involves many bands, which is not convenient for the development and realization of a multispectral system for wheat diseases. Therefore, further intensive study is required.

The objective of this study is to discriminate powdery mildew and stripe rust in wheat through a series of operations, including spectral data collection, dimensionality reduction, related characteristic band extraction, and identification model establishment based on hyperspectral imaging technology. This optimized identification model can provide a new method for the rapid and accurate diagnosis of powdery mildew and stripe rust in wheat at the leaf level.

1 Materials and Methods

1.1 Sample Preparation

Wheat Mingxian 169 was grown in plastic pots (7 cm×7 cm×8 cm) at a density of 10～15 strains in each pot, in a rust-free growth chamer [(16±3) ℃, 16 h light/8 h darkness] in the State Key Laboratory of Crop Stress Biology for Arid Areas and College of Plant Protection, Northwest A&F University, China. When the seedling had grown to two or three leaves, the first wheat seedling leaf was gently rubbed with a clean moistened finger to remove the waxy layer from the leaf surface. *Puccinia striiformis* and *Blumeria graminis* on naturally infected leaves (from living tissues) were collected and evenly smeared on the first wheat seedling leaf with a brush, and the inoculation zones were labeled with a marker. Each pathogen was inoculated on ten pots, while another ten pots without inoculation were used as healthy controls. The inoculated plants were covered with wet polyethylene bags to maintain 100% relative humidity and were stored for 24 h at 10 ℃ in a dark chamber. Immediately after incubation, plants were transferred to a clean growth chamber set to a diurnal cycle of 16 ℃ for the 16-h light period and 13 ℃ for the 8-h dark period[14]. On day 15, various scabs were observed on the surface of the inoculated wheat leaves. 80 strip rust leaves, 100 powdery mildew leaves, and 140 healthy leaves were collected.

1.2 Acquisition and Calibration of Hyperspectral Images

The hyperspectral images of wheat leaves were captured by a line-scanning hyperspectral imaging systemin reflectance mode. The hyperspectral imaging system was constituted by a visible/infrared imaging spectrometer (ImSpector V10E, Spectral Imaging, Finland), an area array CCD camera with 320×256-pixel resolution (XEVA2616, XenICs, Belgium), an illumination system with four 150-W quartz tungsten halogen lamps adjusted at an angle of 45° to illuminate the camera's field of view, a mobile platform operated by a stepper motor, a camera obscurer, and a computer with data acquisition and preprocessing software (Spectral-Cube data acquisition V10 software). The hyperspectral imaging system was set up in the laboratory with the room temperature of (28±1) ℃ and relative humidity of 50%. The spectral region was 375～1 017 nm, with a total of 256 bands. To obtain clear and undistorted hyperspectral images, the distance between samples

and lens was 65 cm, and the speed of the conveyor was 14 mm • s^{-1}, and the exposure time was 5 ms during the image acquisition.

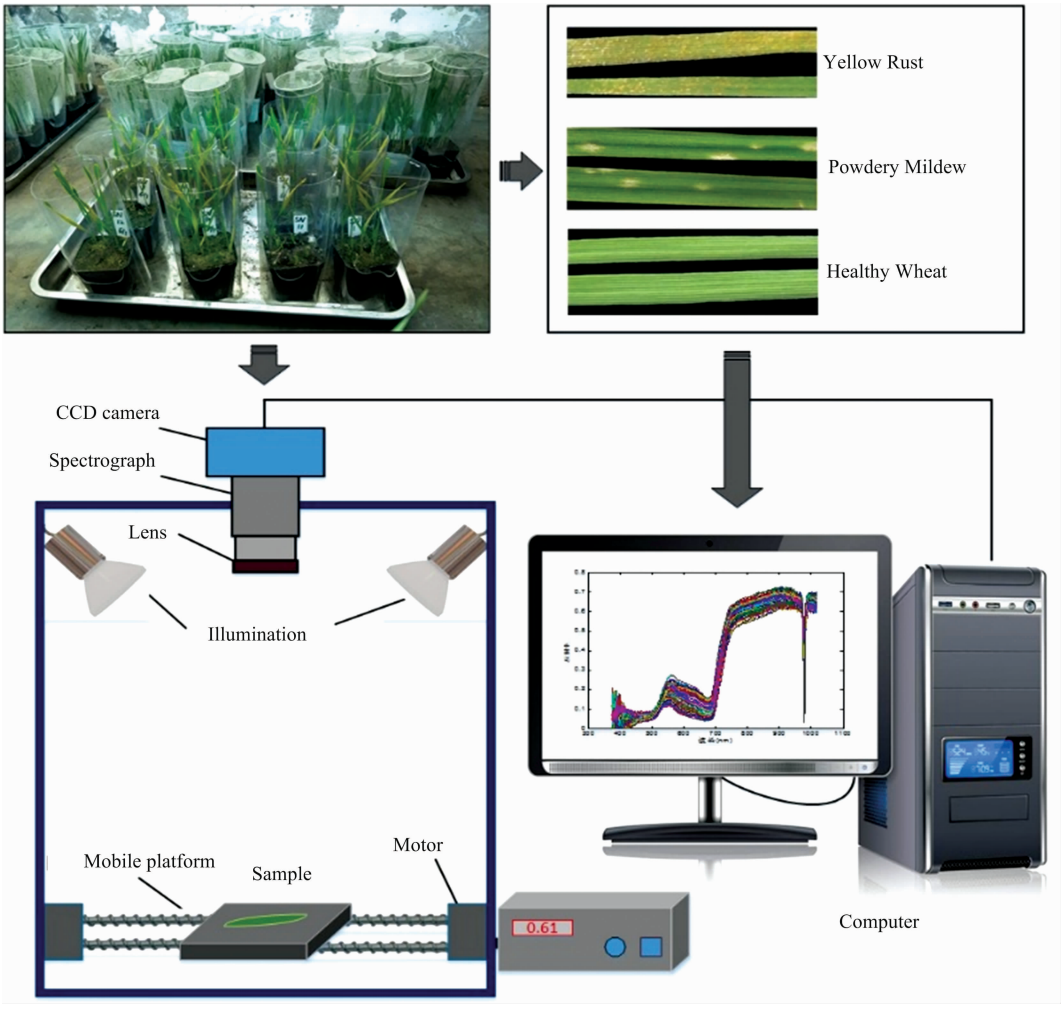

Fig. 1　Photographs of wheat seedlings (upper) and schematic diagram of hyperspectral imaging system (lower)

1.3　Spectral Acquisition and Preprocessing

The spectral reflectance of collected samples were collected from hyperspectral images using ENVI+IDL 5.1 software (ITT Visual Information Solutions, Boulder, CO, USA). Rectangles of 5×4 pixels were selected on inoculation zones of the corresponding leaf as the regions of interest (ROI). The spectra of each pixel in the ROI were averaged, and this spectrum was considered the spectral reflectance of the sample. According to this procedure, a total of 320 mean reflectance spectra of three kinds of wheat samples were obtained and imported into the MATLAB 2016a software (MathWorks, Natick, MA, USA) for further data analysis.

The common processing methods of SG smoothing, multiple scatter correction, standard normal variate transformation, first derivative, and second derivative were tested in this study. SG smoothing was found to eliminate spectral noise while maintaining the spectral characteristics. Therefore, SG smoothing was selected for data treatment throughout the study.

The kennard-stone(KS) algorithm was adopted to divide 80 powdery mildew samples, 100 wheat stripe rust samples and 140 normal wheat samples at a ratio of 3∶1 in this study. Afterward, the respective division results were summarized to form calibration sets with 240 samples and test sets with 80 samples.

1.4　Characteristic Wavelengths Selection and Classification Models

Full spectra contain all spectral information with large data files, information redundancy, and multiple collinear variables, which could add to the complexity of identification model, reduce the computation speed, and affect the model accuracy. Therefore, it is necessary to reduce the dimensionality of the spectral data to improve the correct identification rate. In this study, PCA, SPA and CARS[15] were used as optimization methods to select the characteristic wavelengths.

Classification approaches aim to divide the data into a

number of distinct classes. LS-SVM and ELM[16] algorithm-were applied to build classification models to distinguish powdery mildew, stripe rust and normal leaves in this study. And the model performances were evaluated by accuracy rate of identification. Typically, a higher correct identification rate was associated with better model precision. In addition, the number of input variables in the model was used to evaluate whether the model was simple or not, and fewer input variables in the model suggested a simpler model.

2 Results and Discussion

2.1 Spectral Features of Wheat Leaves

The mean spectral reflectivity curves and standard deviation of stripe rust, powdery mildew, and normal wheat samples were shown in Figure 2. It could be seen that the spectra of inoculated and unvaccinated wheat seedlings had similar change trends with wavelength. However, the mean reflectivity of inoculated wheat leaves was higher than that of the healthy samples over the band range of 400～1 000 nm. In particular, there was a strong spectral reflectivity response in the wheat leaves infected with powdery mildew and stripe rust at the band of 550～680 nm, and the greatest reflectance difference could reach 0.15 compared with healthy leaves. In addition, the leaf reflectivity manifested small fluctuations in the band at 750～900 nm, which was largely consistent with the result trend observed by Yuan et al[17]. It was previously reported that pathogen infection led to structural destruction of chloroplasts, loss of cell water, and weakened the spectral absorption of pigment and water. Thereby, the spectral reflectivity of the infected leaves increased at the visible band and the short-wave infrared band.

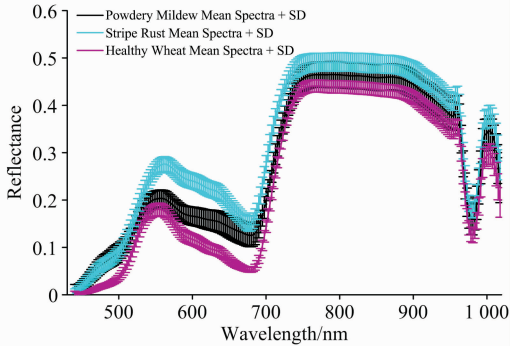

Fig. 2 Mean spectral reflectance curves [with standard deviation (SD)] of wheat leaves with powdery mildew, stripe rust and healthy samples covering the range of 375～1 017 nm

2.2 Selection of Effective Wavelengths

2.2.1 Selection of Effective Wavelengths by PCA

PCA was employed to transform the full wavebands (256 wavebands) into several principal components (PCs). Loadings of first PCs were applied for qualitatively identifying the optimal wavelengths that were responsible for the specific features. The first three PCs explained 97% original variations and their score plots were displayed in Figure 3(a). It indicated that three groups of samples provided an apparent clustering and could be distinguished clearly in the score plots. The X-loadings of PC1 to PC3, which revealed the importance of the analyzed variables, were shown in Figure 3(b). The X-loadings values reflected the correlation coefficients of the principal components at different wavelengths. The wavelengths corresponding to the wave peaks or troughs were considered as the characteristic wavelengths. The first three X-loadings plots of PCA indicated that the reflectance at three wavelengths (560, 680 and 758 nm) with the relatively large loading coefficients had the greatest discriminatory effect on healthy and infected wheat.

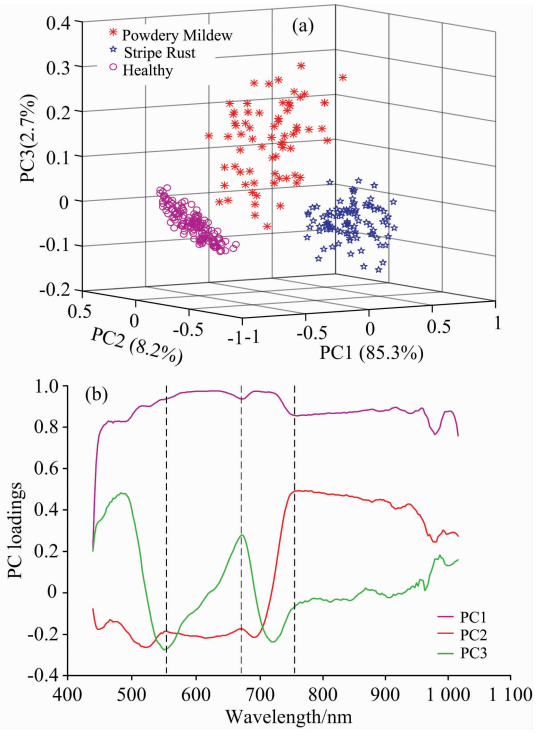

Fig. 3 Characteristic variables selected by PCA

(a): Score projection of the first three principal components; (b): X-loadings plot of the first three principal components on full wavelength region

2.2.2 Selection of Effective Wavelengths by SPA

The number of characteristic wavelengths to be selected was set from 1 to 30, and the optimal characteristic wavelengths were selected based on the root mean square error of calibration (RMSEC). As shown in Figure 4(a), the value of RMSEC decreased rapidly when the number of wavelengths increased from 1 to 5, then decreased slowly with the increase of number of variables. Since too much variables will increase

the complexity of model and slow down the computation speed, the number of variables is often selected when the RMSEC value is the minimum or has a relatively small decrease with increasing numbers of variables. In this study, the determined number of wavelengths was 6. The wavelengths selected finally were marked with solid shapes in Figure 4(b): 481, 516, 562, 675, 789 and 838 nm. The selected wavelengths were used as the inputs for the discrimination models.

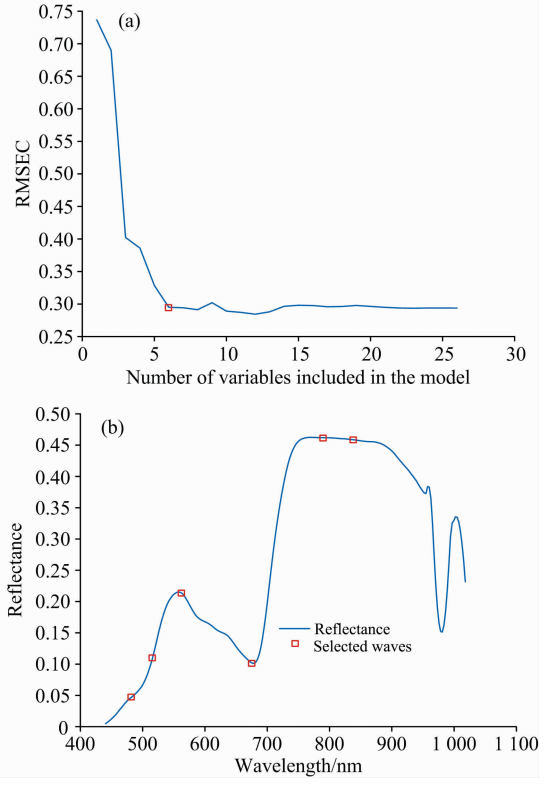

Fig. 4　Characteristic variables selected by SPA

(a): Variation of RMSEC with increasing number of variables in SPA; (b): 6 characteristic wavelengths selected by SPA

2.2.3　Selection of Effective Wavelengths by CARS

The results for variable selection for wheat diseases using the CARS method are shown in Figure 5. Under the action of the exponential decay function, the rate of decrease of variable number became slower with the increase in running times, indicating the stages of roughing and handpicking of the CARS algorithm for key variable selection. The RMSECV value first decreased and then increased, with the smallest RMSECV value at the sampling frequency of 22, which suggested that information variables irrelevant to wheat disease information were eliminated in steps 1～22. As shown in Figure 5, the selected spectral variable subset was optimal at the sampling frequency of 22, which contained 30 spectral variables.

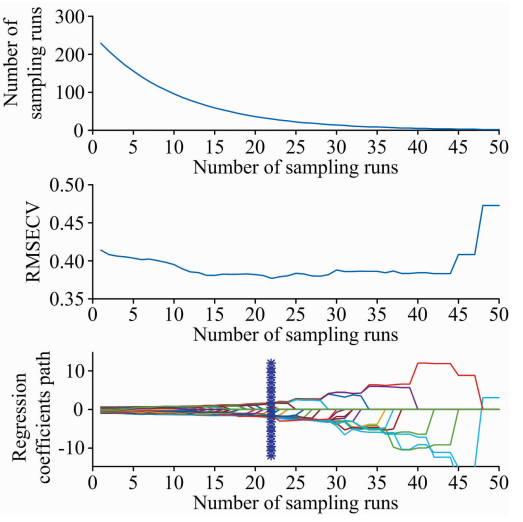

Fig. 5　Characteristic variables selected by CARS

2.3　Comparison and Analysis of the Identification Models

Spectral data of the samples were divided into the calibration set and the test set based on the KS method. Meanwhile, the full spectra, the effective wavelengths selected by PCA, SPA and CARS, were treated as the input variables of the discrimination models. The discrimination models to identify different diseases were constructed by ELM and SVM, and the models were verified using the test set samples. The identification rates of the three kinds of wheat leaves in the calibration set and test set by the proposed model were shown in Table 1. The correct identification rates of the eight models in the calibration set and test set were above 94.58%. It could be seen that the model based on the full spectra (FS) did not display superiority, indicating that the full spectra not only contained useful information but also contained noise. The 30 spectral bands selected by CARS could comprehensively reflect the effective information in the original spectra to distinguish the powdery mildew, stripe rust, and healthy leaves. However, there were only six input variables in the model constructed by SPA and three input variables in the model of X-loadings of PCA, which were only 2.3% and 1.1% of the FS input variable numbers, thus greatly simplifying the recognition model.

2.4　Microstructure Images of Mesophyll Tissue Image

Wheat disease leads to external morphological and internal physiological changes of the leaves. In this paper, the images of cell microstructure of the wheat samples were obtained with a laboratory electron microscope (Olympus-BX53, Olympus Corp., Tokyo, Japan). Figure 6 depicted the microstructure images of mesophyll tissue of normal leaves and leaves affected by stripe rust or powdery mildew. It could be seen that the healthy ones had a complete structure of cross-cut sections, a compact structure and plentiful epidermal

cells, and the chloroplasts were attached around the cell wall [Figure 6(c)]. The powdery mildew pathogen covered the leaf surface to form the powdery mildew layer, then entered the leaf tissue through a primary germ tube, destroyed the histological structure of the mesophyll cell, deprived the leaf of nutrients, and reduced the water and chlorophyll contents [Figure 6(a)]. Urediniospores landed on the wheat surface, entered the leaf tissue and accumulated to create a large number of summer spore banks or spore beds, after which the histological structure of the mesophyll cell was destroyed and the chlorophyll content reduced [Figure 6(b)].

Table 1 Results of identification method for classifying different wheat diseases

Classification method	Variable selection methods	Number of variables	Calibration sets detected (undetected)				Test sets detected (undetected)			
			Powdery mildew (60)	Stripe rust (70)	Healthy wheat (110)	Accuracy /%	Powdery mildew (20)	Stripe rust (30)	Healthy wheat (30)	Accuracy /%
SVM	FS	256	59 (1)	70 (0)	110 (0)	99.58	20 (0)	30 (0)	30 (0)	100
	CARS	30	58 (2)	70 (0)	110 (0)	99.17	19 (1)	30 (0)	30 (0)	98.75
	SPA	6	52 (8)	69 (1)	110 (0)	96.25	18 (2)	30 (0)	30 (0)	97.50
	PCA	3	47 (13)	70 (0)	110 (0)	94.58	18 (2)	30 (0)	30 (0)	97.50
ELM	FS	256	54 (6)	69 (1)	110 (0)	97.92	19 (1)	30 (0)	30 (0)	98.75
	CARS	30	56 (4)	70 (0)	110 (0)	98.33	19 (1)	30 (0)	30 (0)	98.75
	SPA	6	58 (2)	70 (0)	110 (0)	99.17	20 (0)	30 (0)	30 (0)	100
	PCA	3	59 (1)	70 (0)	110 (0)	99.58	20 (0)	30 (0)	30 (0)	100

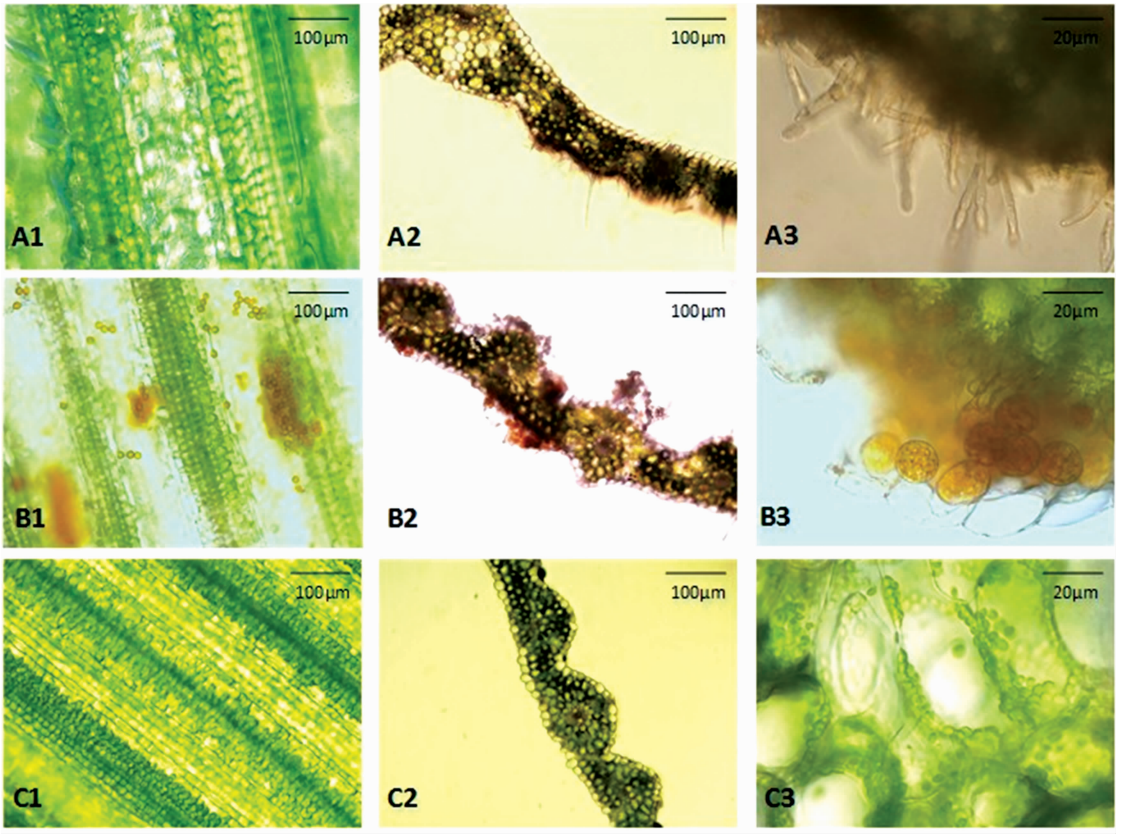

Fig. 6 Cell microstructure images of wheat leaf tissue infected with powdery mildew (a), stripe rust (b), and healthy wheat as control (c)

Although the infection mechanisms of these two diseases were slightly different, both destroyed the mesophyll cells, reduced chlorophyll content, intensified transpiration, reduced photosynthesis markedly. The string of changes led to

weakened light absorption but increased reflectivity in the visible light band. However, the optimal sensitive band for each disease varied according to the disease type, which was consistent with the observations of Mahlein[2]. Consequently, wheat disease identification could be quantitatively realized through monitoring the characteristic spectral changes of wheat, which could provide timely and reliable monitoring of wheat diseases.

3 Conclusions

A novel strategy for rapidly discrimination of stripe rust, powdery mildew and healthy wheat leaves was proposed using infrared spectroscopy imaging combined with multivariate analysis. To reduce the large dimensionality of the hyperspectral imaging, X-loading of PCA, SPA and CARS were employed to select the most suitable optimal wavelengths, based on which the LS-SVM and ELM discriminative model were established. Among all developed models, the PCA-ELM model achieved the best performance, with the accuracy of 99.18% on the calibration set and 100% on test set respectively. Moreover, this model was simple in structure with only three bands (1.1% of full spectra). The results indicated that it is possible to develop a multispectral system for the real-time detection of wheat diseases by hyperspectral imaging.

References

[1] Yang L, Zhang X, Zhang X, et al. PLoS ONE, 2017, 12(5).
[2] Mahlein A-K. Plant Disease, 2016, 100(2): 241.
[3] Sendin K, Williams P J, Manley M. Critical Reviews in Food Science and Nutrition, 2018, 58(4): 575.
[4] Li J, Rao X, Ying Y. Computers and Electronics in Agriculture, 2011, 78(1): 38.
[5] Liu Y, Xiao H, Sun X, et al. Transactions of the Chinese Society of Agricultural Engineering, 2018, 34(3): 180.
[6] Wang H, Qin F, Ruan L, et al. PLoS ONE, 2016, 11(4): e0154648.
[7] Senthilkumar T, Jayas D S, White N D G, et al. Biosystems Engineering, 2016, 147: 162.
[8] Cong S, Sun J, Mao H, et al. Journal of the Science of Food and Agriculture, 2018, 98(4): 1453.
[9] Mahlein A-K, Steiner U, Hillnhuetter C, et al. Plant Methods, 2012, 8.
[10] Zhao Y R, Yu K Q, Li X, et al. Scientific Reports, 2016, 6.
[11] Zhao J, Du S, Huang W, et al. 1st International Conference on Agro-Geoinformatics, Agro-Geoinformatics 2012. IEEE Computer Society, 2012: 5.
[12] Zhang J C, Pu R L, Wang J H, et al. Computers and Electronics in Agriculture, 2012, 85: 13.
[13] Liang D, Liu N, Zhang D, et al. Infrared and Laser Engineering, 2017, 46(1).
[14] Wang Z, Zhao J, Chen X, et al. Plant Disease, 2016, 100(1): 131.
[15] Li H D, Liang Y Z, Xu Q S, et al. Analytica Chimica Acta, 2009, 648(1): 77.
[16] Bian X H, Li S J, Fan M R, et al. Analytical Methods, 2016, 8(23): 4674.
[17] Yuan Lin, Zhang Jingcheng, Zhao Jinling, et al. Spectroscopy and Spectral Analysis, 2013, 33(6): 1608.

基于高光谱成像的小麦白粉病与条锈病识别

姚志凤[1,2,3]，雷　雨[1,2,3]，何东健[1,2,3]*

1. 西北农林科技大学机械与电子工程学院，陕西 杨凌　712100
2. 农业农村部农业物联网重点实验室，陕西 杨凌　712100
3. 陕西省农业信息感知与智能服务重点实验室，陕西 杨凌　712100

摘　要　小麦白粉病和条锈病是我国两种最普遍、最具破坏性的小麦病害，且田间常常混合发生。由于病源和发病机理不同，有必要对这两种病害进行准确区分和识别，以采取不同的防治措施。基于 ImSpector V10E 高光谱成像系统采集的条锈菌侵染叶片、白粉菌侵染叶片和健康叶片(共计 320 个)在 375～1 017 nm 范围内的高光谱图像，利用高斯平滑等预处理方法得到三种小麦叶片的平均光谱曲线，发现小麦白粉病和条锈病的敏感波段均集中在 550～680 nm 的色素强吸收位置，且趋势基本一致。针对两种病害的响应波段交叉重叠的问题，通过主成分分析-载荷法(PCA)、连续投影算法(SPA)和竞争性自适应重加权算法(CARS)对小麦叶片的光谱信息进行有效降维，分别优选出 3、6、30 个敏感波段和特征波长；在此基础上，采用最小二乘-支持向量机和极限学习机两种分类算法分别基于全波段、PCA、SPA 和 CARS 的优选波段，建立白粉病、条锈病和健康叶片的判别模型。结果表明，8 种模型的准确识别率均在 94.58％以上。其中，主成分分析-载荷法结合极限学习机模型最优，训练集与校正集的正确识别率分别为 99.18％和 100％，且结构简单，仅含有三个变量(占全波段的 1.1％)。最后，通过对小麦白粉病、条锈病以及健康叶片的显微结构分析，发现病菌入侵叶片，破坏细胞结构，导致叶绿素含量减少，光合作用效能降低，进而使得小麦在可见光波段光吸收程度减弱，反射率增大。可见，利用作物的高光谱图像信息能够准确地识别不同类型的小麦病害，为研发作物病害在线识别的多光谱系统提供重要的理论依据。

关键词　白粉病；条锈病；高光谱成像；特征波长；判别模型

(收稿日期：2018-08-27，修订日期：2019-01-06)

* 通讯联系人