Agricultural Mechanization, Electrification and Automation (2022)
Clausius Scientific Press, Canada
DOI: 10.23977/hyde.2022.010101
ISSN 2616-1885 Vol. 1 Num. 1

Identification of Grapefruit Black Spot Based on Hyperspectral Imaging using Naïve-Bayes Classifier

Sitan Ye[1], Haiyong Weng [2,*]

[1]School of Engineering, Newcastle University, Newcastle, UK
[2]College of Mechanical and Electrical Engineering, Fujian Agriculture and Forestry University, Fuzhou, China
**corresponding author*

Keywords: Grapefruit black spot, Hyperspectral imaging, Classification, Random frog, Principle component analysis, Naïve-Bayes

Abstract: Citrus black spot (CBC) was considered as one of quarantine diseases to citrus production all over the world. Timely removal citrus infected by CBC can stop it propagation and reduce economic losses in the process factory at harvest time. Hyperspectral images were obtained from CBC-infected, healthy and melanose grapefruits, respectively. Principle component analysis (PCA) and random frog (RF) were performed to select optimal features for classification. Two supervised classifiers, Naïve-Bayes and K-nearest neighbour algorithm (KNN), were implemented to compare their classification performance. It found that five wavelengths (488, 532, 534, 682 and 684 nm) selected by PCA combining with Naïve-Bayes (PCA-NB) achieved a best classification accuracy of 100% while PCA-KNN, RF-KNN, and RF-NB reached 86.1%, 86.2% and 86.1%, respectively. The research has demonstrated that hyperspectral imaging was a potential technology for grapefruit black spot on-line detection.

1. Introduction

Citrus black spot (CBC), caused by the fungus *Phyllosticta citricarpa*, is considered as one of main serious quarantine diseases to citrus production in the worldwide. The fungus reproduces through sexual ascospores and becomes maturation driven by moisture and temperature [1]. Many citrus varieties including lemon, mandarin and grapefruit can be infected by *P. citricarpa*, which greatly reduced the fruit yield and appearance quality. In the European Union, the CBC-infected citrus is prohibited in the commercial trade to avoid *P. citricarpa*, propagation in citrus growing orchard [2]. Therefore, it is of great importance to detection CBC during the fruit process factory and orchards to stop it spread to other CBC free citrus-growing areas. CBC typical symptoms can be identified as red orange legion in circle shape with black margins in the early stage, and the disease spot becomes reddish brown to dark brown with edge uplift along with the time [3].

The management of CBC mainly relies on fungicide sprays during peak season of infection, which often begins from four to five months after fruit set [4] and postharvest control by controlling storage conditions [2]. Timely discarding CBC-infected fruit can effectively stop conidia and

ascospores of *P. citricarpa* disease epidemics. However, in the package house, other citrus defectives including melanose and canker could be often confused as CBC in the process. Therefore, CBC defectives detection could make a contribution to prevent this quarantine disease spreading to other places that free of it and reduce economic losses.

Fungicides-based is one of effective methods for CBC control but has limitations of pesticide residue and fungicide resistance [5]. Currently, spectral imaging technique has been widely used in plant disease and fruit defectives detection during harvest time because it is environment friendly, fast and non-destructive .Yu et al. developed a method to identify jujube crack defect feature using hyperspectral imaging in order to guarantee jujube quality for an increasingly demanding of products to consumers [6]. Principal component image (PC4) was selected to identify the location of crack and achieved a good classification. Qin et al. has designed a technology of real-time citrus canker detection based on a two-band spectral imaging system [7]. Two monochrome cameras combined with two bandpass filters at 730nm and 830nm, respectively, were used to capture narrowband images for classification which reached an overall accuracy of 95.3%. Cen et al. had conducted a cucumber chilling injury detection investigation and found that majority of the optimal wavelengths selected by three feature selection methods were all from spectral transmittance in the short-near infrared range [8]. Features selected by sequential forward selection (SFS) together with SVM achieved the best classification result. Li et al. had built a hyperspectral imaging system for common defectives (i.e. wind scarring, infestation, canker and copper burn) on oranges [9]. Principal component analysis (PCA) was implemented to select potential wavelengths for on-line multispectral imaging system. It found that six wavelengths for third principal component provided a better classification result and two-band ratio (R875/R691) had a good ability to discriminate stems from other defectives. Additionally, many application based on spectral imaging have been reported for fruit quality assessment, such as blueberries, apples and bananas [10-12].

To our knowledge, there are few researches about grapefruit black spot detection with a non-destructive method. This article reports on the use hyperspectral imaging technology to defect CBC on grapefruit at postharvest time. Therefore, the object of this investigation was to develop suitable feature selection methods and classification algorithms for grapefruit black spot detection. The specific objects were to:(1) obtain hyperspectral images of grapefruit samples with CBC and other similar peel defectives; (2) select optimal wavelengths for classification by using different feature selection methods and image processing algorithm for CBC lesions detection; (3) develop discrimination models based on optimal wavelengths and compare the performance of the classifiers.

2. Materials and Methods

2.1. Grapefruit Samples Collection

Grapefruit samples were picked up from Zhangzhou city, Fujian province, China. in October, 2021. Grapefruit with black spot, melanose and normal peel conditions were collected in this investigation by expert-based inspection. A total of 113 grapefruits (52 black spot, 24 melanose and 37 normal) with good shape and similar color were selected for detection. Hyperspectral images of each sample were acquired as soon as possible after the samples detaching from trees.

2.2. Hyperspectral Images Acquisition System

A push-broom hyperspectral images acquisition system was employed in this research (Figure 1). The system was consisted of a charge coupled device (CCD) camera with 672×512 pixels resolution (C8484-05, Hamamatsu Photonics, Hamamatsu City, Japan), a spectrograph with 2.8 nm

resolution in the range of 379-1023 nm (ImSpector V10E, Spectral Imaging Ltd., Oulu, Finland), two line lights (Fiber-Lite DC950, Dolan Jenner Industries Inc., Boxborough, MA), a mobile platform (IRCP0076, Isuzu Optics Crop, Taiwan), a dark box and a computer installed with V10E software which can modify several parameters to achieve the best output performance. In order to obtain right shape and clear hyperspectral images, a set of parameters must be adjusted. Grapefruit samples were placed on the mobile platform with the moving speed of 1.82mm/s under 90ms exposure time light illumination condition. The distance between lens and sample was 30cm and the two line lights were placed on both sides of lens above the sample to reduce shadow. Each component was fixed in dark box to avoid any impact on hyperspectral images collection. A total of 113 hyperspectral images were obtained in sequence.

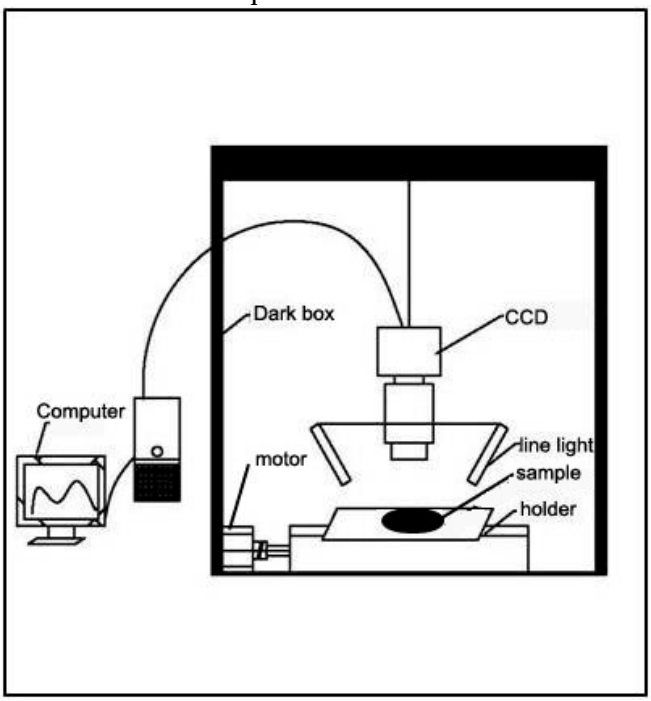

Figure 1: System for acquisition of grapefruit hyperspectral images

2.3. Data Analysis

Since the shape of most grapefruits were ellipse and had large curve radius, the distance between samples to lens varied greatly which causing a distortion in shape of samples and heterogeneity of light distribution. In order to obtain a higher signal to noise ratio, the spatial resolution of original hyperspectral images (672×512 pixels) was resized as 400×400 pixels after discarding the boundary pixels. In addition, the wavelength in the range of 380-409 nm was also given up for a reason of containing much noise. Eventually, 486 wavelengths from 410 to 1023 nm were used for further analysis. Raw hyperspectral images were then calibrated by using a Teflon disk (99.9% reflectance) image and dark image (0% reflectance). The relative images were computed as each waveband image subtraction of dark image and then division the difference between Teflon disk images and dark images [13]. The region of interest (ROI) with about 25×25 pixels was extracted from black spot, melanose and normal tissues, respectively, based on the ENVI software platform. All hyperspectral data processing procedure was carried out in ENVI 4.6 (ITT visual information solutions, Boulder, CO, USAs), Matlab 2010a (MathWorks, Inc., Natick, MA, USA) and 'The Unscramber X 10.1' (CAMO PROCESS AS, Oslo, USA).

Mean spectrum from ROI of each grapefruit was computed and then was used for optimal wavelengths selection. PCA was applied on mean spectrum from ROI of each grapefruit to obtain principal component loadings. The loadings of first four principal components were used to identify important wavelengths and a higher absolute loading value indicting a more important waveband [14]. In addition, Random frog (RF) was also performed to select optimal wavelengths for discriminating different peel conditions. It outputs the probability of each selected variable. A higher probability of being chosen variable indicates a more important waveband. RF is an effective variable selection algorithm from high-dimension that can gain the most relevant variables for classification [15, 16].

For a classification task, two supervised classifiers of K-nearest neighbour algorithm (KNN) and Naïve-Bayes were used to compare their classification performance. KNN is a non-parametric learning algorithm that chooses k neighbours from training data set using a distance function and then classify into a class based on largest voting rules. In machine learning, Naïve-Bayes classifier is probabilistic classifier based on Bayes' theory with independence assumptions between the features [17]. All samples were divided into training and testing set by Kennard-Stone algorithm and a confusion matrix was used to show the classification results.

3. Results and Discussion

3.1. Mean Spectra Reflectance of Grapefruit Samples

Mean spectra extracted from region of interest (ROI) of hyperspectral images was calculated for the representation of each disease spectra. Typical mean spectra of selected ROIs over the range of 410-1023 nm wavelength from the three peel conditions of black spot, melanose and normal grapefruits was shown in Figure 2. In general, the spectra of three peel conditions showed a similar trend in the Vis/NIR range. The reflectance of black spot over NIR range was higher than normal and melanose samples in most case. The difference between normal and melanose did not exhibit stronger contrast than that of black spot peel condition. Features of local minima were observed around 675 nm and 960 nm due to the absorption of chlorophyll and water, respectively. Red edge position (REP) over the range of 680-750 nm in black spot grapefruit showed a slight left shift compared with normal and melanose grapefruit. The absorption around 675 nm in black spot was the lowest, but the normal grapefruit was largest. The different spectra reflectance around 680-750 nm was relation with different chlorophyll content in the ROI position on the sample surface tissues. This can be ascribed as that the fungus destroyed the grapefruit surface chlorophyll for reproduction. The mean reflectance around 960 nm (O-H 2nd overtone) in black spot peel condition was highest among the three peel condition. A decrease of water content in black spot and melanose grapefruit might also be able to be referred as a result of fungi-host interaction.

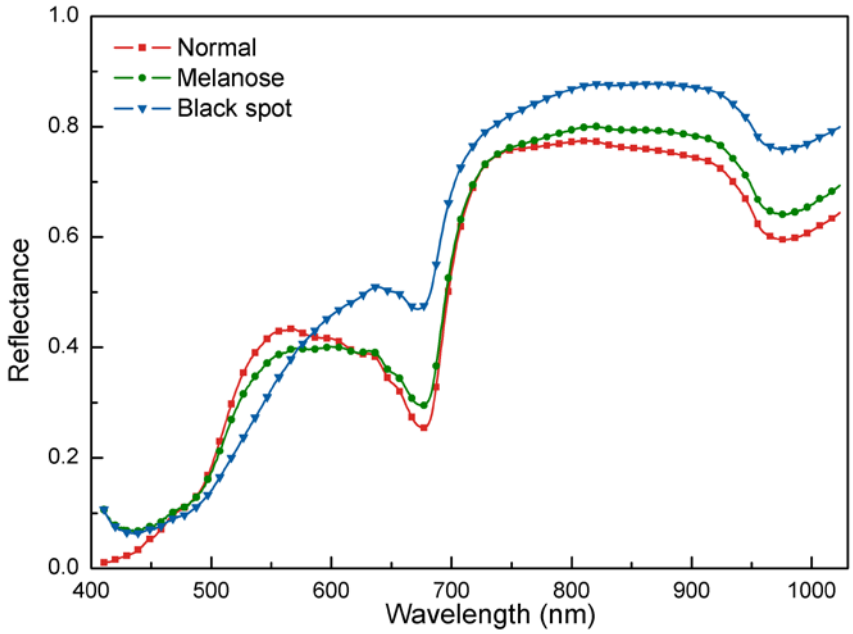

Figure 2: Mean spectrum of the normal, melanose and black spot grapefruit samples

3.2. Features Selection and Classification Performance Based on Hyperspectral Images

After the application of PCA, the new variables were orthogonal and independent mutually which can contribute to build a robust model and improve calculate rate. Figure 3 displays the total accumulative contribution rate of variance from the first four PCs (PC1, PC2, PC3 and PC4) of all grapefruit samples. PC1, PC2, PC3 and PC4 explained 61%, 27.2%, 9.3% and 2.3% variance, respectively. The first four PCs have explained 99.8% variables of raw hyperspectral images information from 113 grapefruit samples.

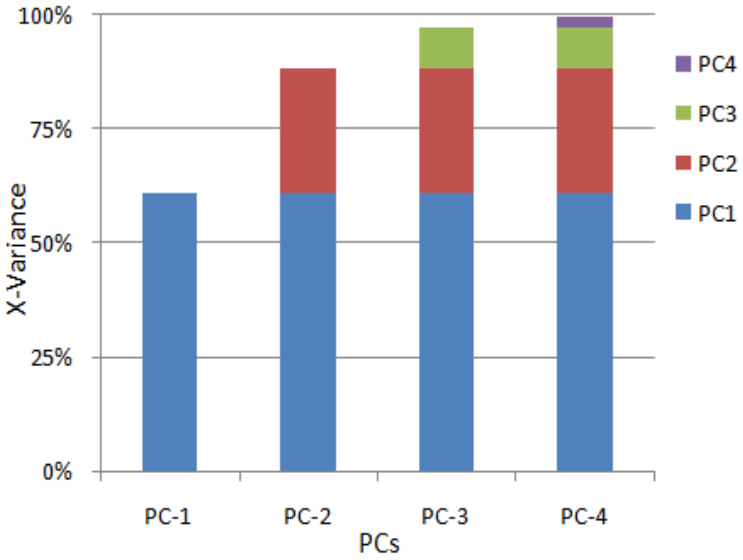

Figure 3: Accumulative contribution rate of variance from the first four PCs of all grapefruit samples

The loading scores curve of top four principal components obtained from PCA which explained 99.8% variables of original hyperspectral images data were shown in Figure 4. The local minima or maxima in principal component loading scores curve revealed as important wavelengths. A local maxima in PC1 (684 nm), PC2 (536 nm), PC3 (682 nm) and PC4 (488 nm) loading scores curve was observed, respectively. A local minima in PC3 (532 nm) loading scores curve was also selected. Totally, 5 dominant wavelengths from 410-1023 nm wavelength were selected by PCA. In other words, these five wavelengths (488, 532, 534, 682 and 684 nm) can represent 486 bands data information, which reduced 99% redundant variables. The wavelengths selected by PCA were all in the range of 450-780 nm, where the reflectance of black spot, melanose and normal grapefruits may show a contrast to each other.

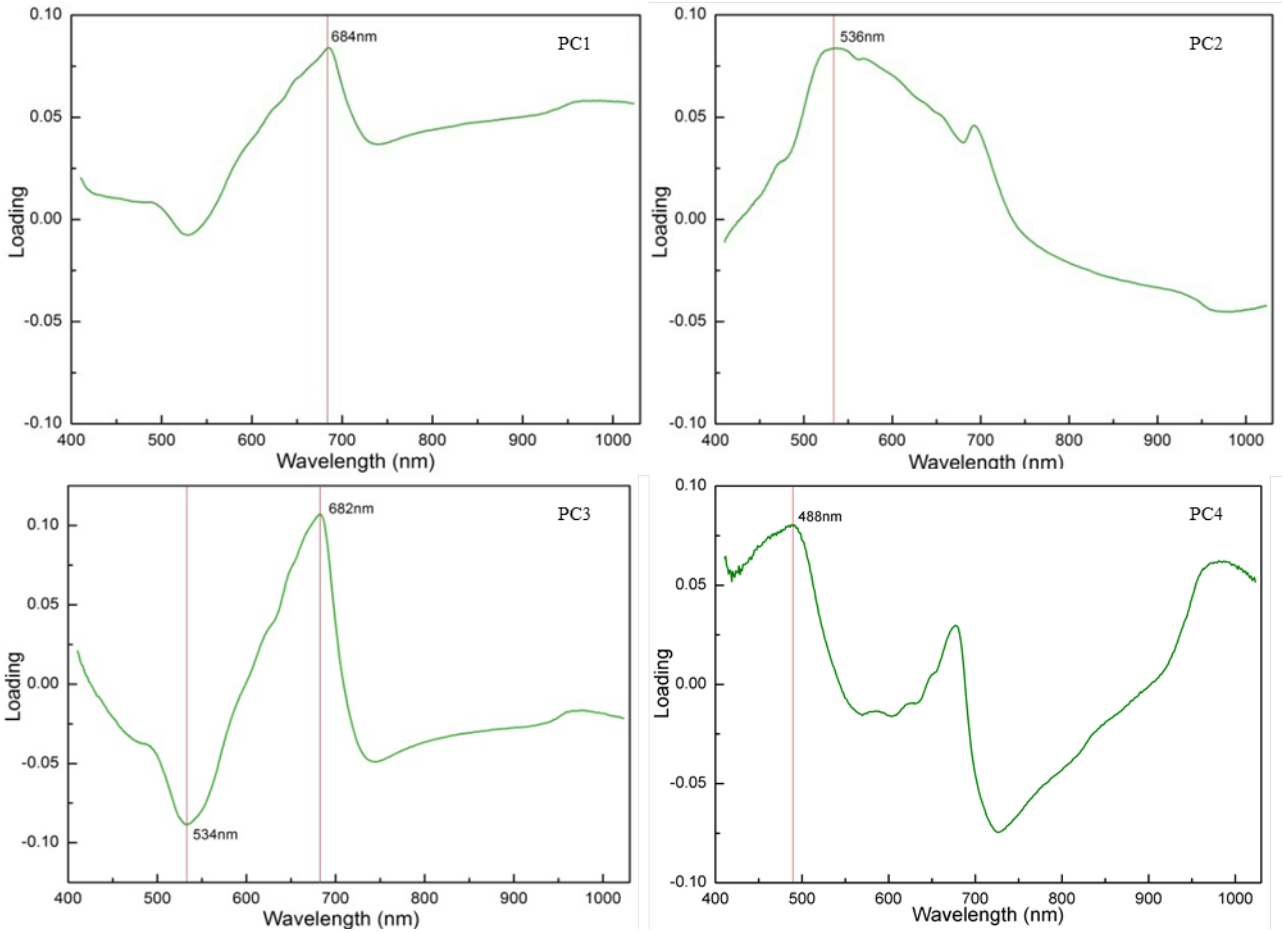

Figure 4: Loading scores curve of top four principal components obtained from PCA of hyperspectral images

The first six wavelengths selected by Random frog (RF) were 628, 635, 811, 900, 997 and 1013 nm, respectively when the threshold selected probability was 35% (Figure 5). Wavelength around 811 nm obtained the highest selected probability (89.5%) demonstrating the strongest relevance to discriminate normal grapefruits from black spot ones. as well as melanose ones. Four wavelengths (811, 900, 997 and 1013 nm) selected by RF were in the range of 780-1013 nm, where the mean spectra reflectance of black spot peel condition was largest among the three classes grapefruits. The other two selected wavelengths (628 nm and 635 nm) in the visible spectrum also revealed as local maxima of mean spectral reflectance of normal and black spot grapefruit. The

wavelengths selected by PCA and RF methods were not the same because of different selection criterion.

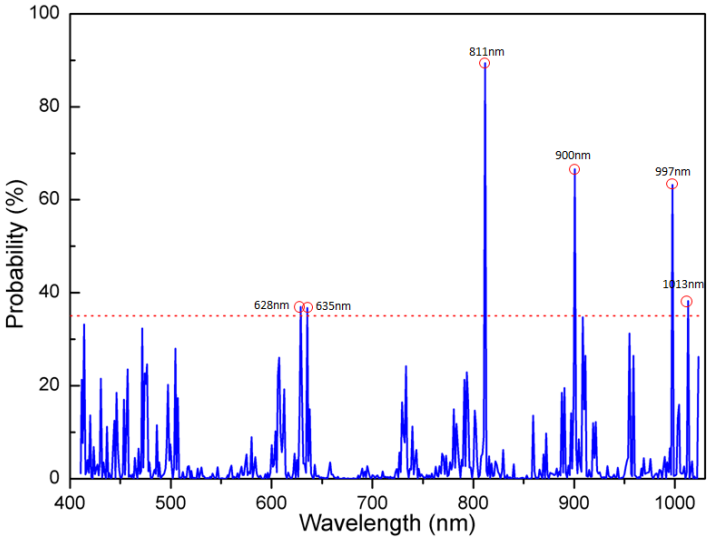

Figure 5: The selected probability of each variable using random frog (RF)

Table 1 has listed the classification results based on different combinations of classifiers and feature selection methods. Naïve-Bayes classifier with full wavelengths achieved an overall accuracy of 66.7%. Four melanose samples were misclassified into normal class. Black spot and normal samples were both classified into their own actual classes with 100% accuracy. K-nearest neighbour algorithm without any variables dimension reduction methods obtained 88% classification accuracy. One sample of each class was misclassified. The unsatisfactory classification results may due to that original hyperspectral images with 486 wavelengths may contain useless information which may have a negative impact on model prediction ability.

Table 1: Classification results of black spot, melanose and normal grapefruit with confusion matrix

Full wavelengths	Naïve-Bayes				KNN			
	Black spot	Melanose	Normal	Accuracy (%)	Black spot	Melanose	Normal	Accuracy (%)
Black spot	19	0	0	100	18	0	1	94.7
Melanose	0	0	4	0	0	3	1	75.0
Normal	0	0	18	100	0	1	17	94.4
Overall Accuracy (%)				66.7				88.0
RF	Naïve-Bayes				KNN			
	Black spot	Melanose	Normal	Accuracy (%)	Black spot	Melanose	Normal	Accuracy (%)
Black spot	19	0	0	100	18	0	1	94.7
Melanose	0	3	1	75.0	0	3	1	75.0
Normal	0	3	15	83.3	0	2	16	88.9
Overall Accuracy (%)				86.1				86.2
PCA	Naïve-Bayes				KNN			
	Black spot	Melanose	Normal	Accuracy (%)	Black spot	Melanose	Normal	Accuracy (%)
Black spot	19	0	0	100	19	0	0	100

Melanose	0	4	0	100	0	3	1	75
Normal	0	0	18	100	0	3	15	83.3
Overall Accuracy (%)				100				86.1

PCA: principal component analysis; RF: random frog. Row data represents actual class, columns data represents prediction class.

The six wavelengths selected by random frog (RF) method had a higher overall accuracy (86.1%) than full wavelengths when Naïve-Bayes was used as a classifier. The individual classification accuracy of black spot, melanose and normal grapefruits was 100%, 75% and 83.3%, respectively. The reason was that these six selected wavelengths may eliminate useless information and enhance the difference among different class samples. As a result, 75% of melanose samples were identified rightly compared with full wavelengths, which was slightly higher than that of full spectral- Naïve-Bayes.

Five wavelengths (488, 532, 534, 682 and 684 nm) selected by PCA for the model input were also evaluated based on discrimination performance. The Naïve-Bayes classifier reached an overall accuracy of 100% of no misclassification. However, the KNN achieved a poor overall classification accuracy (86.1%). One sample in melanose class was misclassified into normal class and three samples in normal class were misclassified into melanose class on the contrary. The best classification results using PCA-Naïve-Bayes may be that the features selected by PCA could more suitable to explain the difference among these three-class surface peel conditions of grapefruit. The classification results had demonstrated that PCA-Naïve-Bayes was a best-fitting model for detection of black spot on grapefruit because of best classification accuracy with minimum feature subset size.

4. Conclusions

To sum up, the technology of hyperspectral imaging combined with image processing was conducted for grapefruit black spot detection. The classification performance using optimal wavelengths selected by random frog (RF) and principal component analysis (PCA) combined with different classifiers was evaluated. In comparison with different combinations, it found that the optimal wavelengths selected by PCA achieved the best classification result of 100% when Naïve-Bayes algorithm was used as a classifier. The five optimal wavelengths selected by PCA had potential ability to form an online multispectral imaging system for grapefruit black spot identification. This research had provided a feasible method to identify grapefruit black spot disease. However, other type defects on grapefruit surface were not included which indicated that optimal wavelengths selected by PCA may be not suitable for other peel diseases. Therefore, further investigation will focus on collecting more samples with other peel diseases and developing a more feasible identification algorithm for detection of grapefruit black spot.

Acknowledgements

This work was supported by Hand-held agricultural environmental parameters measuring station (111ZC7024).

References

[1] Martínez, M.J., Conesa, D., López, Q.A. and Vicent, A. (2015) Climatic Distribution of Citrus Black Spot Caused by Phyllosticta Citricarpa. A Historical Analysis of Disease Spread in South Africa. European Journal of Plant Pathology, 143, 69-83.

[2] Lombardo, P., Guimaraens, A., Franco, J., Dellacassa, E. and Pérez, F.E. (2016) Effectiveness of Essential Oils for Postharvest Control of Phyllosticta Citricarpa (Citrus Black Spot) on Citrus Fruit. Postharvest Biology and Technology, 121, 1-8.

[3] Daegwan, K., Thomas, F., Burks, M.A. and Ritenour, Q.J. (2014) Citrus Black Spot Detection Using Hyperspectral Imaging. Int J Agric & Biol Eng, 7, 20-26.

[4] Silva, J.G.J., Scapin, M., Silva, F.P., Silva, A.R.P., Behlau, F. and Ramos, H.H. (2016) Spray Volume and Fungicide Rates for Citrus Black Spot Control Based on Tree Canopy Volume. Crop Protection, 85, 38-45.

[5] Possiede, Y.M., Gabardo, J., Kava, C.V., Galli, T.L.V., Azevedo, J.L., et al. (2009) Fungicide Resistance and Genetic Variability in Plant Pathogenic Strains of Guignardia Citricarpa. Brazilian Journal of Microbiology, 40, 308-313.

[6] Yu, K., Zhao, Y., Li, X., Shao, Y., Zhu, F., et al. (2014) Identification of Crack Features in Fresh Jujube Using Vis/NIR Hyperspectral Imaging Combined with Image Processing. Computers and Electronics in Agriculture, 103, 1-10.

[7] Qin, J., Burks, T.F., Zhao, X., Niphadkar, N. and Ritenour, M.A. (2012) Development of A Two-Band Spectral Imaging System for Real-Time Citrus Canker Detection. Journal of Food Engineering, 108, 87-93.

[8] Cen, H., Lu, R., Zhu, Q. and Mendoza, F. (2016) Nondestructive Detection of Chilling Injury in Cucumber Fruit Using Hyperspectral Imaging with Feature Selection and Supervised Classification. Postharvest Biology and Technology, 111, 352-361.

[9] Li, J., Rao, X. and Ying, Y. (2011) Detection of Common Defects on Oranges Using Hyperspectral Reflectance Imaging. Computers and Electronics in Agriculture, 78, 38-48.

[10] Leiva, V.G.A., Lu, R. and Aguilera, J.M. (2014) Assessment of Internal Quality of Blueberries Using Hyperspectral Transmittance and Reflectance Images with Whole Spectra or Selected Wavelengths. Innovative Food Science & Emerging Technologies, 24, 2-13.

[11] ElMasry, G., Wang, N. and Vigneault, C. (2009) Detecting Chilling Injury in Red Delicious Apple Using Hyperspectral Imaging and Neural Networks. Postharvest Biology and Technology, 52, 1-8.

[12]. Rajkumar, P., Wang N., Eimasry, G., Raghavan, G.S.V. and Gariepy, Y. (2012) Studies on Banana Fruit Quality and Maturity Stages Using Hyperspectral Imaging. Journal of Food Engineering, 108, 194-200.

[13] Cen, H., Lu, R., Ariana, D.P. and Mendoza, F. (2013) Hyperspectral Imaging-Based Classification and Wavelengths Selection for Internal Defect Detection of Pickling Cucumbers. Food and Bioprocess Technology, 7, 1689-1700.

[14] Torbick, N., Becker, B. (2009) Evaluating Principal Components Analysis for Identifying Optimal Bands Using Wetland Hyperspectral Measurements from The Great Lakes, USA. Remote Sensing, 1, 408-417.

[15] Li, H.D., Xu, Q.S. and Liang, Y.Z. (2012) Random Frog: An Efficient Reversible Jump Markov Chain Monte Carlo-Like Approach for Variable Selection with Applications to Gene Selection and Disease Classification. Anal Chim Acta, 740, 20-26.

[16] Yun, Y.H., Li, H.D., Wood, L.R., Fan, W., Wang, J.J., et al. (2013) An Efficient Method of Wavelength Interval Selection Based on Random Frog for Multivariate Spectral Calibration. Spectrochim Acta A Mol Biomol Spectrosc, 111, 31-36.

[17] Pant, B., Pant, K. and Pardasani, K.R. (2010) Naïve Bayes Classifier for Classification of Plantand Animal miRNA. International Journal of Computer Theory and Engineering, 2, 1793-8201.