Journal of Food Engineering 179 (2016) 11–18

Contents lists available at ScienceDirect

Journal of Food Engineering

journal homepage: www.elsevier.com/locate/jfoodeng

Hyperspectral imaging analysis for ripeness evaluation of strawberry with support vector machine

Chu zhang [a], Chentong Guo [b], Fei Liu [a], Wenwen Kong [a], Yong He [a, *], Binggan Lou [b]

[a] College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou, 310058, China
[b] College of Agriculture and Biotechnology, Zhejiang University, Hangzhou, 310058, China

ARTICLE INFO

Article history:
Received 19 April 2015
Received in revised form
5 January 2016
Accepted 6 January 2016
Available online 22 January 2016

Keywords:
Hyperspectral imaging
Strawberry
Ripeness classification
Texture features
Data fusion

ABSTRACT

A hyperspectral imaging system covering two spectral ranges (380–1030 nm and 874–1734 nm) was applied to evaluate strawberry ripeness. The spectral data were extracted from hyperspectral images of ripe, mid-ripe and unripe strawberries. The optimal wavelengths were obtained from spectra of 441.1–1013.97 and 941.46–1578.13 nm by loadings of principal component analysis (PCA). Pattern texture features (correlation, contrast, entropy and homogeneity) were extracted from the images at optimal wavelengths. Support vector machine (SVM) was used to build classification models on full spectral data, optimal wavelengths, texture features and the combined dataset of optimal wavelengths and texture features, respectively. SVM models using combined datasets performed best among all datasets. SVM models using datasets from hyperspectral images at 441.1–1013.97 nm performed better with classification accuracy over 85%. The overall results indicated that hyperspectral imaging could be used for strawberry ripeness evaluation, and data fusion combining spectral information and spatial information showed advantages in strawberry ripeness evaluation.

1. Introduction

Fruit ripeness, the key factor for determining the optimal harvest time for fruits, is crucial for fruit growers. Fruit ripeness is a quite complex issue. Traditionally, fruit ripeness is identified by human experience or laboratory based detection of quality parameters (color, texture, chemical constituents, etc.) or their combinations. Usenik et al. (2014) used firmness, soluble solids content (SSC) and color information and taste to evaluate plum ripeness. Fuentes de Mendoza et al. (2013) used fatty acids, triglycerides and sterols profile of olive oil to evaluate olive ripeness. López Camelo and Gómez (2004) used color features to evaluate tomato ripeness. Shinya et al. (2013) studied the fruit mass, soluble solids content (SSC), ground skin color features, the spectral absorbance difference at 670 nm and 720 nm index (I_{AD}), fruit/flesh firmness and uniaxial compression strength related to peach ripeness. Azodanlou et al. (2004) used total volatile compounds, total acidity, total sugar content (degrees Brix) and fruit firmness to evaluate strawberry ripeness. These methods could acquire satisfactory accuracy, but suffer the problems of being time consuming, high cost and the rigorous requirement of operation skills. Advanced techniques, including machine vision, spectroscopy and spectral imaging, have been applied to evaluate the degree of fruit ripeness in a fast, non-destructive, low-cost and convenient way.

Machine vision technique captures the image of the samples, using external features extracted from the images like color, shape, size and texture features combined with the measured reference values of different chemical and physical properties to build models to predict the chemical and physical properties of unknown samples. Changes of external features of fruits during ripening make it feasible to evaluate fruit ripeness with machine vision technique. Machine vision has been proved efficient in fruit ripeness evaluation, and fruit chemical constituents could also be predicted by machine vision. Fadilah et al. (2012) used color features from images captured by a RGB camera to evaluate oil palm ripeness. Tan et al. (2010) used a RGB camera to evaluate oil palm ripeness and oil content. Ji et al. (2013) studied the banana ripeness by a digital imaging method with color feature analysis. Mendoza and Aguilera (2004) studied banana ripeness by computer vision system with color feature analysis, texture analysis and chemical analysis. Velez-Rivera et al. (2014a) used a computer vision system to evaluate mango ripeness with color features, TA, SSC, firmness and ripening index (RPI).

* Corresponding author.
E-mail address: yhe@zju.edu.cn (Y. He).

http://dx.doi.org/10.1016/j.jfoodeng.2016.01.002
0260-8774/

However, machine vision only captures the external information of the targets, uses the external information to evaluate fruit external and internal quality. Spectroscopy technique could obtain internal quality information as a fast and non-destructive method. Changes of internal quality of fruits during ripening make it feasible to use spectroscopy technique to evaluate ripeness of fruits. Studies have been reported of using spectroscopy technique to evaluate fruit ripeness and fruit quality. Larrain et al. (2008) used a portable near-infrared spectroscopy (NIRS) instrument to determine sugar (Brix), pH, and anthocyanin concentration related to ripeness of grape. Giovenzana et al. (2014) used a handheld optical system to determine total soluble solids and polyphenols of grape to evaluate grape ripeness. Tarkosova and Copikova (2000) used NIRS to determine carbohydrate content of banana to evaluate banana ripeness. Qin et al. (2012) used spatially offset Raman spectroscopy to evaluate tomato ripeness.

Both machine vision and spectroscopy could evaluate ripeness without measuring the chemical or physical parameters, they use the predefined ripeness stages as the reference ripeness data to build models and extract the information contributing to fruit ripeness. However, the two techniques focus on different aspects of the samples. Spectral imaging technique is the integration of machine vision technique and spectroscopy technique, acquiring both spatial information and spectral information simultaneously. Hyperspectral imaging (HSI) is one of the common forms of the spectral imaging technique.

HSI has been successfully used to determine fruit quality, fruit ripeness and fruit damage. Wei et al. (2014) used spectral data and texture features from hyperspectral images to evaluate ripeness of Astringent persimmon. Tallada et al. (2006) used hyperspectral imaging system to determine strawberry firmness related to ripeness. Lleo et al. (2011) used spectra indices extracted from hyperspectral images to evaluate peach ripeness. Velez Rivera et al. (2014b) used hyperspectral imaging to detect mango mechanical damages. Each pixel within hyperspectral image contains a spectrum which shows the spectral property of the pixel. Generally, the average spectrum of all pixels of the pre-defined sample region (also called region of interest, ROI) can be used to build regression or discriminant models, and the spectrum of each pixel can be used for prediction to acquire chemical, physical or category information of the pixel by the established models to form a prediction map (also called distribution map).

The main objective of this study is to evaluate strawberry ripeness by HSI at different spectral ranges, using both spectral and spatial information. The specific objectives of this study are to: 1) explore the potential of two different spectral ranges to evaluate strawberry ripeness; 2) establish SVM models to identify strawberry ripeness based on full spectral data and selected optimal wavelengths; 3) extract and apply texture information to identify strawberry ripeness; 4) combine spectral data and texture features to identify strawberry ripeness.

2. Materials and method

2.1. Strawberry samples

Strawberry samples (cv. Hongyan) were collected from a local strawberry park in Hangzhou, Zhejiang province, China. Sixty ripe strawberries, 60 mid-ripe strawberries and 60 unripe strawberries were collected by local experienced strawberry growers, all samples were healthy without bruises. The samples were taken to the laboratory for hyperspectral image acquisition after collection and cleaning.

2.2. Hyperspectral image acquisition

2.2.1. Hyperspectral imaging system

A laboratory based hyperspectral imaging system was used in this study (Fig. 2). The system acquired hyperspectral images by 2 different camera systems covering two spectral ranges (380–1030 nm with 512 bands and 874–1734 nm with 256 bands). The former is acquired by an imaging spectrograph (ImSpector V10E; Spectral Imaging Ltd., Oulu, Finland), a 672 × 512 CCD camera (C8484-05, Hamamatsu, Hamamatsu City, Japan) with a camera lens (OLES23; Specim, Spectral Imaging Ltd., Oulu, Finland), and the latter is acquired by an imaging spectrograph (ImSpector N17E; Spectral Imaging Ltd., Oulu, Finland), a 320 × 256 CCD camera (Xeva 992; Xenics Infrared Solutions, Leuven, Belgium) with a camera lens (OLES22; Specim, Spectral Imaging Ltd.). Above all, the system contained two 150 W tungsten halogen lamps (Fiber-Lite DC950 Illuminator; Dolan Jenner Industries Inc, Boxborough, MA, USA) placed in two sides of the camera symmetrically at a 45° angle for illumination, a conveyer belt driven by a stepper motor (Isuzu Optics Corp, Taiwan, China). The system was placed in a dark room, and controlled by a computer.

2.3. Image acquisition and correction

To acquire clear and non-deformable images containing the full samples, the height between the lens and the sample, the moving speed of the conveyer belt and the exposure time of the camera should be adjusted. For images at 380–1030 nm, the height, the moving speed, and the exposure time were set as 360 mm, 2.05 mm/s and 0.05 s, respectively. For images at 874–1734 nm, the height, the moving speed, and the exposure time were set as 360 mm, 25 mm/s and 5 ms, respectively. The samples were placed in the conveyer belt, and were scanned line by line along the Y-axis with the sample moving along the X-axis at a certain speed.

The white reference image and dark reference image should be acquired under the same experimental conditions of hyperspectral image acquisition for image correction. The dark reference image was captured by turning off the light source and covering the camera lens completely by its opaque cap, and the white reference image was acquired using a piece of white Teflon, and the correction was conducted by the following equation:

$$I_c = \frac{I_{raw} - I_{dark}}{I_{white} - I_{dark}} \quad (1)$$

where I_c was the corrected image, I_{raw} was the raw image, I_{white} was the white reference image, I_{dark} was the dark reference image. The resolution of the acquired images at 380–1030 nm was 92 PPI (pixels per inch) and the resolution of the acquired images at 874–1734 nm was 32 PPI.

2.3.1. Spectral data extraction

For hyperspectral images at 380–1030 nm and 874–1734 nm, ROI should be predefined to extract spectral information. In this study, ROI of each sample was defined as the entire fruit sample region of each strawberry, and the spectral information of the ROI was extracted. Average spectrum of all pixels within the ROI was used to represent the sample.

2.4. Texture feature extraction

The texture features of strawberry hyperspectral images could represent the features of strawberry, like color, roughness, intensities, firmness, chewiness and their arrangements, which could directly or indirectly relate to fruit quality or ripeness. Texture

features have been used to evaluate fruit ripeness. Gray level co-occurrence matrix (GLCM) (Haralick et al., 1973), a widely used method for texture feature extraction, is used to extract texture features in this study. Four popular texture features extracted by GLCM was studied in this study including correlation, contrast, entropy and homogeneity (Mendoza and Aguilera, 2004). Correlation measures the linear dependence of neighborhood grayscale in an image. Contrast measures the local variations and reflects the image definition. Entropy measures the image information. Homogeneity measures the closeness of the distribution of elements of an image.

Hyperspectral image contains gray-scale images at continuous wavelengths, and texture features could be extracted from each gray-scale image, resulting in large amount of data, which is difficult to process. In this study, texture features of direction of 0°, 45°, 90° and 135°and distance of 1 were extracted from the gray-scale images at selected optimal wavelengths, and the average values of the 4 directions were used.

3. Data analysis

3.1. Support vector machine

To accurately identify strawberry ripeness, classification models were built based on the spectral data and the texture features. Support vector machine (SVM) was a supervised classification method with good generalization ability, dealing with both linear and nonlinear data efficiently (Devos et al., 2009; Langeron et al., 2007; Li et al., 2014). SVM had been proved as a reliable and efficient method in spectral data. To perform SVM, kernel functions should be determined, and parameters of SVM model should also be determined. In this study, the kernel function of SVM was chosen as radial basis function (RBF), and the parameters of SVM were optimized by a grid search procedure and 5 fold cross validation. The performances of SVM models were evaluated by classification accuracy.

3.2. Optimal wavelengths selection

Dealing with the immense amount of data generated by HSI is the main challenge for HSI application. The spectral bands of hyperspectral images are highly correlated, resulting in co-linearity and redundant information. The analysis of the immense amount of data is a heavy computation task, which requires high performance hardware and software. The noises and useless information in the data are the factors affecting the data analysis results, and how to reduce the influence of the noises and useless information is the concerned issue in hyperspectral image analysis. However, optimal wavelengths selection, aiming to select only a few wavelengths which carry the most of useful information with minimum co-linearity and redundancy from full spectra, is believed to reduce data dimensionality and solve the mentioned problems in dealing with hyperspectral image data. In this study, principal component analysis (PCA) was used to select optimal wavelengths. Loadings of the wavelengths in each principal component (PC) indicated the contributions of the wavelengths, the wavelengths with higher absolute loading value were with greater contributions, and the peaks and valleys of loading plot (indicating the high loading values) could be selected as optimal wavelengths (Serranti et al., 2012; Rodriguez-Pulido et al., 2013; Kamruzzaman et al., 2011). Generally, the first few PCs could explain most of the variance, and loadings of these PCs could be used to select optimal wavelengths.

3.3. Software

Spectral data and texture features extraction were conducted on ENVI 4.6 (ITT, Visual Information Solutions, Boulder, USA), and SVM models were built on Matlab R 2009b (The Math Works, Natick, USA), PCA was conducted on Unscrambler® 10.1 (CAMO AS, Oslo, Norway).

4. Results and discussion

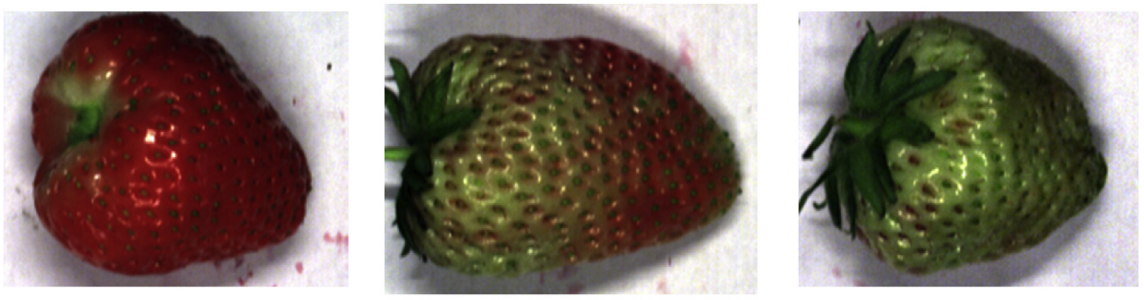
4.1. Spectral profiles

Average spectra of the ripe, mid-ripe and unripe strawberries at the two spectral ranges were shown in Fig. 3. Because of the obvious noises in the head and the end of the spectra acquired by two different imaging systems, only spectra at 441.1–1013.97 nm and 941.46–1578.13 nm preprocessed by Savitzky-Golay smoothing (polynomial degree 2 with 7 points) were used.

As shown in Fig. 3 (a), the great differences were observed between 500 nm and 700 nm. The differences near 535 nm indicated the differences of the content of anthocyanin in different ripeness strawberry samples (Lleo et al., 2011). The differences near 675 nm were mainly caused by the different content of Chlorophyll (Tallada et al., 2006). As presented in Fig. 1, the color of the strawberry samples at different ripeness stage were different. In Fig. 3 (b) the average spectrum of each ripeness stage in the range of 941.46–1578.13 nm showed similar spectral curves and slightly differences of reflectance value.

4.2. Principal component analysis

Principal component analysis (PCA) was used for qualitative analysis of differentiating strawberry ripeness stages. PCA was



(a) ripe strawberry (b) mid-ripe strawberry (c) Unripe strawberry

Fig. 1. RGB images of ripe, mid-ripe and unripe strawberry samples.

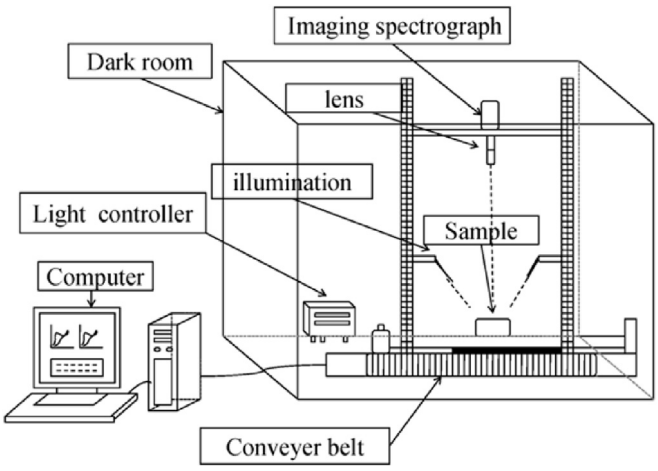


Fig. 2. Hyperspectral imaging system.

performed on the preprocessed spectra at the two spectral ranges. For spectra at 441.1–1013.97 nm, PC1 and PC2 explained 66.34% and 21.25% of variance, respectively. Fig. 4(a) shows the 2D scores scatter plot of PC1 and PC2. It could be observed that the ripe samples and the unripe samples were well grouped and separated clearly, the mid-ripe samples could also be observed a clear separation from the ripe samples, and there were overlaps between the mid-ripe samples and the unripe samples. For spectra at 941.46–1578.13 nm, PC1 and PC2 explained 92.25% and 5.55% of variance, respectively. The 2D scores scatter plot (Fig. 4)(b) showed that there were no evident separation of the three kinds of the samples, and overlaps were observed of the samples of 3 ripeness stages. The results of 2D PCA scores scatter plot indicated that identifying strawberry ripeness stages using images at 441.1–1013.97 nm would be more promising than using images at 941.46–1578.13 nm. However, discriminant models needed to be investigated for quantitative analysis of identifying strawberry ripeness.

4.3. SVM models on full spectra

To build SVM models, the ripe, mid-ripe and unripe samples were assigned category values of 1, 2 and 3. The samples of each category were split into the calibration and the prediction set at the ratio of 2:1. 120 samples and 60 samples were in the calibration and the prediction set, respectively. The classification results of SVM models were shown in Table 1.

For the spectral range of 441.1–1013.97 nm, the discrimination accuracy was 99.17% for the calibration set and 91.67% for the prediction set. The ripe samples were successfully identified with no samples misclassified, while some of mid-ripe and unripe samples were misclassified. For the spectral range of 941.46–1578.13 nm, SVM model obtained relatively worse performances, with classification accuracy less than 90% in the prediction set. Some of the ripe, mid-ripe and unripe samples were misclassified into other ripeness stage.

The classification results of the two image acquisition systems were satisfactory, but SVM model for 941.46–1578.13 nm performed worse. The reason might be that the differences of the pigments between samples of different ripeness stages were more obvious, while the difference of the chemical constituents between closest ripeness stages were not significant. For the two different image acquisition systems, samples of closest ripeness stages (ripe and mid-ripe, mid-ripe and unripe) could be misclassified as each other, and samples of ripe and unripe stages were rarely misclassified as each other.

4.4. Optimal wavelengths selection

Hyperspectral image generates an immense amount of data, which was the main problem in hyperspectral image analysis. The use of average spectra of the samples could significantly reduce the data amount. However, the high dimensional full spectra suffered from the co-linearity and redundancy of the wavelengths, causing complex models and poor performances. Feature selection was studied to minimize the efforts caused by the problems and make the discriminant models more robust and more accurate. In order to reduce the high dimensionality of the extracted spectral data, optimal wavelengths were selected to use only a few wavelengths which carry the most important information to represent the full spectra.

For both spectra at 441.1–1013.97 nm and 941.46–1578.13 nm, the first 3 PCs explained most of the variance in both situations (PC1 = 66.34%, PC2 = 21.25% and PC3 = 6.92% for 441.1–1013.97 nm; PC1 = 92.25%, PC2 = 5.55%, and PC3 = 1.97% for 941.46–1578.13 nm), and loadings of the first 3 PCs were applied to select optimal wavelengths. Peaks and valleys which had the highest absolute loading values were selected as the optimal wavelengths of each PC. Fig. 5(a) and (b) showed the loading plots and the corresponding selected optimal wavelengths of the first 3 PCs for spectra at 441.1–1013.97 nm and 941.46–1578.13 nm 6 optimal wavelengths were selected in Fig. 5(a). Fig. 5(b) showed

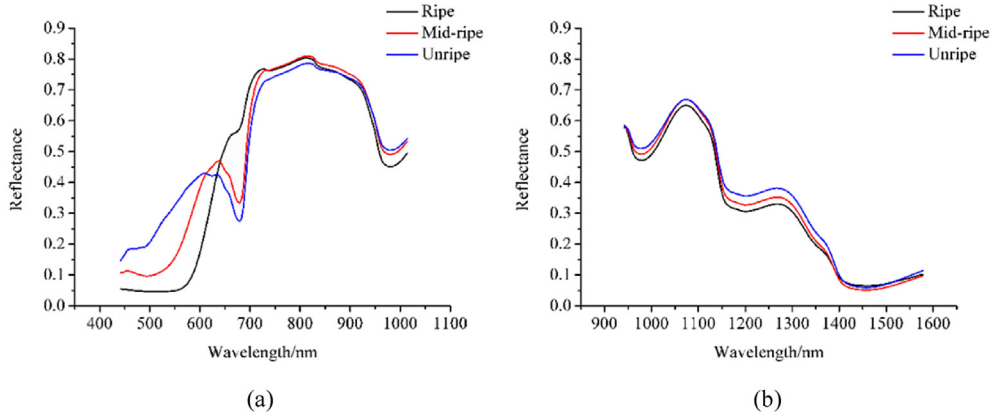


Fig. 3. Average spectrum of each ripeness stage of 441.1–1013.97 nm (a) and 941.46–1578.13 nm (b).

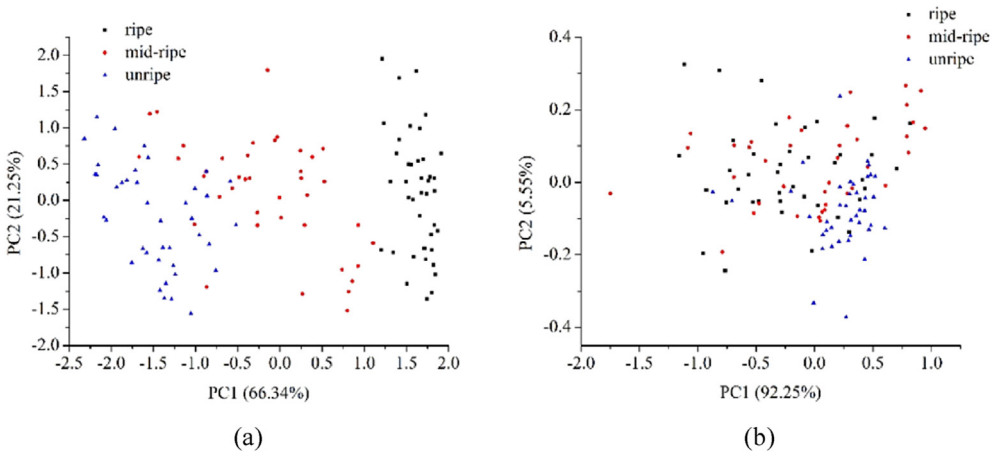


Fig. 4. 2D PCA scores scatter plot of 441.1–1013.97 nm (a) and 941.46–1578.13 nm (b).

Table 1
Classification results of SVM models on full spectra.

Spectral range	[a](c, g)		Calibration				[b]CV	Prediction			
			1	2	3	Accuracy	Accuracy	1	2	3	Accuracy
441.1–1013.97 nm	(147.0334, 0.0039)	1	40	0	0	100%		20	0	0	100%
		2	0	39	1	97.5%		1	17	2	85%
		3	0	0	40	100%		0	2	18	90%
		Total				99.17%	95%				91.67%
941.46–1578.13 nm	(147.0334, 0.0206	1	38	2	0	95%		20	0	0	100%
		2	0	35	5	87.5%		1	15	4	75%
		3	0	3	37	90%		1	6	13	65%
		Total				91.67%	78.33%				80%

[a] (c, g) were the parameters of SVM model, c was the penalty coefficient, and g as the kernel function parameter.
[b] CV was the abbreviation of cross validation.

that 10 optimal wavelengths were selected, and it could be found that some of the selected wavelengths (978.37 nm, 1069.03 nm and 1459.81 nm) were near 981.72 nm, 1075.75 nm and 1456.44 nm with quite small differences, thus 981.72 nm, 1075.75 nm, and 1456.44 nm were selected as optimal wavelengths instead of those near them, and 10 wavelengths were reduced to 5. The selected wavelengths of 441.1–1013.97 nm and 941.46–1578.13 nm were shown in Table 2. The selected optimal wavelengths at the range of 441.1–1013.97 nm were related to the color information.

4.5. SVM models on optimal wavelengths

As a consequence of wavelengths selection by PCA loadings, the

Table 2
Optimal wavelengths selected by PCA.

Spectral range	Number	Wavelengths (nm)
441.1–1013.97 nm	6	525.6, 533.01, 573.93, 606.38, 681.95, 692.1
941.46–1578.13 nm	5	981.72, 1075.75, 1146.35, 1334.95, 1456.44

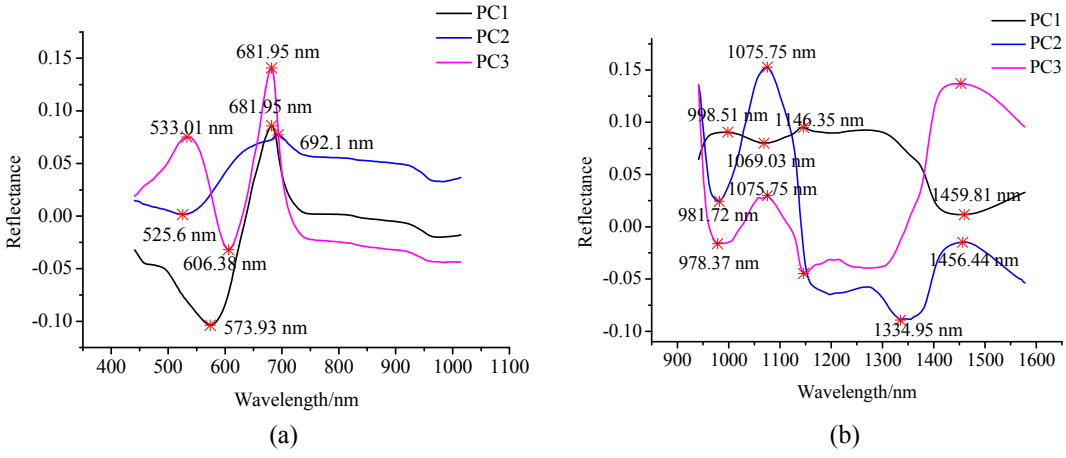


Fig. 5. PCA loading plots and the selected optimal wavelengths of 441.1–1013.97 nm (a) and 941.46–1578.13 nm (b).

number of wavelengths reduced to 6 and 5 in the two spectral ranges. The selected optimal wavelengths were used to build SVM models instead of the full spectra. The performances of developed SVM models using selected optimal wavelengths were shown in Table 3. It could be noted that SVM model on the optimal wavelengths at 441.1–1013.97 nm obtained similar performance as SVM model on corresponding full spectra, while the SVM model on optimal wavelengths at 941.46–1578.13 nm performed much worse than SVM model on corresponding full spectra. The classification results of ripe, mid-ripe and unripe samples in Table 3 met with the corresponding results in Table 1, ripe samples could be misclassified as mid-ripe samples, and unripe samples could be misclassified mid-ripe samples, while mid-ripe samples could be misclassified as ripe and unripe samples. However, the ripe and unripe samples were rarely misclassified as each other.

4.6. SVM models on texture features

Texture features of ripe, mid-ripe and unripe samples were extracted from the gray-scale images at optimal wavelengths. For images at 441.1–1013.97 nm, ROI with 50 × 50 pixels were defined in the sample region of each strawberry. Four texture features, including correlation, contrast, entropy and homogeneity, were extracted, and a new dataset with 24 texture features was formed for each sample. For images at 941.46–1578.13 nm, ROI with 20 × 20 pixels were defined in the sample region of each strawberry and 20 features of each sample were extracted. SVM models were built on the extracted texture features, and the results were shown in Table 4. SVM models using texture features at 441.1–1013.97 nm performed better than SVM models using texture features at 941.46–1578.13 nm, with classification accuracy of calibration and prediction of 90% and 85%, respectively. The SVM model using texture features at 941.46–1578.13 nm obtained poor classification accuracy of the prediction set of 71.67%, indicating that the application of texture features of 941.46–1578.13 nm for strawberry ripeness evaluation should be further studied. The overall results indicated that it was feasible to use texture feature for strawberry ripeness stage evaluation. Based on texture features, classification results of ripe, mid-ripe and unripe samples showed similar phenomenon as Tables 1 and 3.

4.7. SVM models on data fusion

The advantage of HSI was to provide spatial and spectral information simultaneously. As the previous analysis in this study, SVM models on spectral data used mainly spectra data, and SVM models on texture data used mainly spatial data, the results showed that both spectral data and texture features could contribute to evaluate strawberry ripeness. The idea of using both spectral information and texture features for analysis (called data fusion) in spectral imaging has been proposed in previous literature (Wei et al., 2014; Liu et al., 2014; Huang et al., 2013; Zhang et al., 2012). In this study, SVM models using both spectral data of optimal wavelengths and texture features were built. The results were shown in Table 5. SVM model using combined dataset of 441.1–1013.97 nm obtained better performance with classification accuracy of calibration and prediction of 99.17% and 95%, while SVM model using combined dataset of 941.46–1578.13 nm obtained acceptable results with classification accuracy of calibration and prediction of 95% and 83.33%. Meanwhile, the similar phenomenon was observed in Table 5 as in Tables 1, 3 and 4 that ripe samples could be misclassified as mid-ripe samples, unripe samples could be misclassified mid-ripe samples, and mid-ripe samples could be misclassified as unripe samples by using the combined datasets.

4.8. Comparison of SVM models on different datasets

In this study, four datasets extracted from hyperspectral images for each spectral range were studied, including full spectra (results in Table 1), optimal wavelengths (results in Table 3), texture features at optimal wavelengths (results in Table 4), and data fusion of combining optimal wavelengths and texture features (results in Table 5). The overall results of the SVM models for 441.1–1013.97 nm and 941.46–1578.13 nm, and the overall results indicated that both spectral data and texture features could be used for strawberry ripeness evaluation. It could be observed that SVM models using datasets for 441.1–1013.97 nm performed better than SVM models using datasets for 941.46–1578.13 nm, the reason might be that the samples used in this study showed more variations of external quality. SVM models using full spectra, optimal wavelengths and the combined dataset of 441.1–1013.97 nm all obtained the satisfactory classification results with classification accuracy of the calibration set and the prediction set over 90%, while SVM model using texture features obtained relative worse results with classification accuracy over 80%. However, SVM models using datasets for 941.46–1578.13 nm performed much worse, and the SVM models using optimal wavelengths and texture features obtained the worst results with classification accuracy of 66.67% and 71.67% of the prediction set, respectively. The results of SVM model using the combined dataset for 941.46–1578.13 nm were much better, with classification accuracy of calibration and prediction of 95% and 83.33%.

In all, SVM models using the combined datasets obtained similar or better results than all other corresponding SVM models, indicating that combination of spectral data and spatial information could be used for strawberry efficiently, revealing the advantage of hyperspectral imaging.

5. Conclusion

In this study, HSI was applied to evaluate the ripeness degree of strawberry. SVM models of spectral data at two different spectral

Table 3
Classification results of SVM models on optimal wavelengths.

Spectral range	(c, g)		Calibration				CV	Prediction			
			1	2	3	Accuracy	Accuracy	1	2	3	Accuracy
441.1–1013.97 nm	(256, 0.0206)	1	40	0	0	100%		19	1	0	95%
		2	0	37	3	92.5%		0	20	0	100%
		3	0	3	37	92.5%		0	3	17	85%
		Total				95%	95%				93.33%
941.46–1578.13 nm	(1.7411, 1)	1	33	4	3	82.5%		14	4	2	70%
		2	6	24	10	60%		2	11	7	55%
		3	0	6	34	85%		1	4	15	75%
		Total				75.83%	71.67%				66.67%

Table 4
Classification results of SVM models on texture features.

Spectral range	(c, g)		Calibration				CV	Prediction			
			1	2	3	Accuracy	Accuracy	1	2	3	Accuracy
441.1–1013.97 nm	(1.7411, 0.0625)	1	39	1	0	97.5%		20	0	0	100%
		2	2	32	6	80%		1	14	5	70%
		3	0	3	37	92.5%		0	3	17	85%
		Total				90%	91.67%				85%
941.46–1578.13 nm	(3.0314, 0.1895)	1	34	5	1	85%		16	3	1	80%
		2	3	26	11	65%		4	13	3	65%
		3	1	5	34	85%		0	6	14	70%
		Total				78.33%	65.83%				71.67%

Table 5
Classification results of SVM models on datasets combing optimal wavelengths and texture features.

Spectral range	(c, g)		Calibration				CV	Prediction			
			1	2	3	Accuracy	Accuracy	1	2	3	Accuracy
441.1–1013.97 nm	(48.5029, 0.1088)	1	40	0	0	100%		20	0	0	100%
		2	0	39	1	97.5%		0	19	1	95%
		3	0	0	40	100%		0	2	18	90%
		Total				99.17%	95%				95%
941.46–1578.13 nm	(27.8576, 0.1088)	1	38	1	1	95%		19	1	0	95%
		2	0	38	2	95%		3	15	2	75%
		3	0	2	38	95%		0	4	16	80%
		Total				95%	70.83%				83.33%

ranges (441.1–1013.97 nm and 941.46–1578.13 nm) showed acceptable results. Optimal wavelengths were selected from the two spectral ranges, SVM models using optimal wavelengths from 441.1 to 1013.97 nm showed satisfactory results while the results of SVM models using optimal wavelengths from 941.46 to 1578.13 nm were much worse. Texture features of correlation, contrast, entropy and homogeneity were extracted from the gray-scale images at optimal wavelengths. SVM models built on the texture features obtained acceptable results. SVM models using the combined datasets of texture features and optimal wavelengths showed best results. In all, ripe samples could be misclassified as mid-ripe samples; unripe samples could be misclassified mid-ripe samples, while mid-ripe samples could be misclassified as ripe and unripe samples. However, the ripe and unripe samples were rarely misclassified as each other. SVM models on the datasets from 441.1 to 1013.97 nm showed much better performances than SVM models on the datasets from 941.46 to 1578.13 nm. The overall results indicated that hyperspectral imaging could be used to evaluate strawberry ripeness, and extracted spectral data and texture features were both effective for this procedure. In future study, more samples with wider variations of external and internal quality of strawberry fruits at different ripeness stage were needed for better identification of strawberry ripeness.

Acknowledgements

This work was supported by the National Key Foundation for Exploring Scientific Instrument of China (2014YQ47037703), Zhejiang Provincial Public Welfare Technology Research Projects (2014C32103), Natural Science Foundation of China (61273062).

References

Azodanlou, R., Darbellay, C., Luisier, J.L., Villettaz, J.C., Amado, R., 2004. Changes in flavour and texture during the ripening of strawberries. Eur. Food Res. Technol. 218 (2), 167–172.
Devos, O., Ruckebusch, C., Durand, A., Duponchel, L., Huvenne, J.-P., 2009. Support vector machines (SVM) in near infrared (NIR) spectroscopy: focus on parameters optimization and model interpretation. Chemom. Intell. Lab. Syst. 96 (1), 27–33.
Fadilah, N., Mohamad-Saleh, J., Halim, Z.A., Ibrahim, H., Ali, S.S.S., 2012. Intelligent color vision system for ripeness classification of oil Palm fresh fruit Bunch. Sensors 12 (10), 14179–14195.
Fuentes de Mendoza, M., De Miguel Gordillo, C., Marin Expoxito, J., Sanchez Casas, J., Martinez Cano, M., Martin Vertedor, D., Franco Baltasar, M.N., 2013. Chemical composition of virgin olive oils according to the ripening in olives. Food Chem. 141 (3), 2575–2581.
Giovenzana, V., Beghi, R., Malegori, C., Civelli, R., Guidetti, R., 2014. Wavelength selection with a view to a simplified handheld optical system to estimate grape ripeness. Am. J. Enol. Vitic. 65 (1), 117–123.
Haralick, R.M., Shanmuga, K., Dinstein, I., 1973. Textural features for image classification. Ieee Trans. Syst. Man Cybern. SMC3 (6), 610–621.
Huang, L., Zhao, J., Chen, Q., Zhang, Y., 2013. Rapid detection of total viable count (TVC) in pork meat by hyperspectral imaging. Food Res. Int. 54 (1), 821–828.
Ji, W., Koutsidis, G., Luo, R., Hutchings, J., Akhtar, M., Megias, F., Butterworth, M., 2013. A digital imaging method for measuring banana ripeness. Color Res. Appl. 38 (5), 364–374.
Kamruzzaman, M., ElMasry, G., Sun, D.-W., Allen, P., 2011. Application of NIR hyperspectral imaging for discrimination of lamb muscles. J. Food Eng. 104 (3), 332–340.
Langeron, Y., Doussot, M., Hewson, D.J., Duchene, J., 2007. Classifying NIR spectra of textile products with kernel methods. Eng. Appl. Artif. Intell. 20 (3), 415–427.
Larrain, M., Guesalaga, A.R., Agosin, E., 2008. A multipurpose portable instrument for determining ripeness in wine grapes using NIR spectroscopy. IEEE Trans. Instrum. Meas. 57 (2), 294–302.
Li, S.X., Zhang, Y.J., Zeng, Q.Y., Li, L.F., Guo, Z.Y., Liu, Z.M., Xiong, H.L., Liu, S.H., 2014. Potential of cancer screening with serum surface-enhanced Raman spectroscopy and a support vector machine. Laser Phys. Lett. 11 (6).
Liu, D., Pu, H., Sun, D.-W., Wang, L., Zeng, X.-A., 2014. Combination of spectra and texture data of hyperspectral imaging for prediction of pH in salted meat. Food Chem. 160, 330–337.
Lleo, L., Roger, J.M., Herrero-Langreo, A., Diezma-Iglesias, B., Barreiro, P., 2011. Comparison of multispectral indexes extracted from hyperspectral images for the assessment of fruit ripening. J. Food Eng. 104 (4), 612–620.
López Camelo, A.F., Gómez, P.A., 2004. Comparison of color indexes for tomato ripening. Hortic. Bras. 22 (3), 534–537.
Mendoza, F., Aguilera, J.M., 2004. Application of image analysis for classification of ripening bananas. J. Food Sci. 69 (9), 471–477.
Qin, J., Chao, K., Kim, M.S., 2012. Nondestructive evaluation of internal maturity of tomatoes using spatially offset Raman spectroscopy. Postharvest Biol. Technol. 71, 21–31.
Rodriguez-Pulido, F.J., Barbin, D.F., Sun, D.-W., Gordillo, B., Lourdes Gonzalez-Miret, M., Heredia, F.J., 2013. Grape seed characterization by NIR hyperspectral imaging. Postharvest Biol. Technol. 76, 74–82.
Serranti, S., Gargiulo, A., Bonifazi, G., 2012. Hyperspectral imaging for process and quality control in recycling plants of polyolefin flakes. J. Near Infrared Spectrosc. 20 (5), 573–581.
Shinya, P., Contador, L., Predieri, S., Rubio, P., Infante, R., 2013. Peach ripening: segregation at harvest and postharvest flesh softening. Postharvest Biol. Technol. 86, 472–478.
Tallada, J.G., Nagata, M., Kobayashi, T., 2006. Non-destructive estimation of firmness of strawberries (Fragaria x ananassa Duch.) using NIR hyperspectral imaging.

Environ. Control Biol. 44 (4), 245–255.
Tan, Y.A., Low, K.W., Lee, C.K., Low, K.S., 2010. Imaging technique for quantification of oil palm fruit ripeness and oil content. Eur. J. Lipid Sci. Technol. 112 (8), 838–843.
Tarkosova, J., Copikova, J., 2000. Determination of carbohydrate content in bananas during ripening and storage by near infrared spectroscopy. J. Near Infrared Spectrosc. 8 (1), 21–26.
Usenik, V., Stampar, F., Kastelec, D., 2014. Indicators of plum maturity: when do plums become tasty? Sci. Hortic. 167, 127–134.
Velez-Rivera, N., Blasco, J., Chanona-Perez, J., Calderon-Dominguez, G., de Jesus Perea-Flores, M., Arzate-Vazquez, I., Cubero, S., Farrera-Rebollo, R., 2014a. Computer vision system applied to classification of “Manila” mangoes during ripening process. Food Bioprocess Technol. 7 (4), 1183–1194.
Velez Rivera, N., Gomez-Sanchis, J., Chanona-Perez, J., Jose Carrasco, J., Millan-Giralolo, M., Lorente, D., Cubero, S., Blasco, J., 2014b. Early detection of mechanical damage in mango using NIR hyperspectral images and machine learning. Biosyst. Eng. 122, 91–98.
Wei, X., Liu, F., Qiu, Z., Shao, Y., He, Y., 2014. Ripeness classification of astringent persimmon using hyperspectral imaging technique. Food Bioprocess Technol. 7 (5), 1371–1380.
Zhang, X., Liu, F., He, Y., Li, X., 2012. Application of hyperspectral imaging and chemometric calibrations for variety discrimination of Maize seeds. Sensors 12 (12), 17234–17246.