ORIGINAL PAPER

Ripeness Classification of Astringent Persimmon Using Hyperspectral Imaging Technique

Xuan Wei · Fei Liu · Zhengjun Qiu · Yongni Shao · Yong He

Abstract Nondestructive detection of fruit ripeness is crucial for improving fruits' shelf life and industry production. This work illustrates the use of hyperspectral images at the wavelengths between 400 and 1,000 nm to classify the ripeness of persimmon fruit. Spectra and images of 192 samples were investigated, which were selected from four ripeness stages (unripe, mid-ripe, ripe, and over-ripe). Three classification models—linear discriminant analysis (LDA), soft independence modeling of class analogy, and least squares support vector machines were compared. The best model was LDA, of which the correct classification rate was 95.3 % with the input consisted of the spectra and texture feature of images at three feature wavelengths (518, 711, and 980 nm). Feature wavelengths selection and texture feature extraction were based on successive projection algorithm and gray level co-occurrence matrix, respectively. In addition, using the same input of ripeness detection to make an investigation on firmness prediction by partial least square analysis showed a potential for further study, with correlate coefficient of prediction set r_{pre} of 0.913 and root mean square error of prediction of 4.349. The results in this work indicated that there is potential in the use of hyperspectral imaging technique on non-destructive ripeness classification of persimmon. The experimental results could provide the theory support for studying online quality control of persimmon.

Keywords Hyperspectral imaging · Persimmon ripeness · Texture feature · Linear discriminant analysis (LDA) · Gray level co-occurrence matrix (GLCM)

Abbreviations

CCR	Correct classification rate
GLCM	Gray level co-occurrence matrix
LDA	Linear discriminant analysis
LS-SVM	Least squares support vector machines
PC	Principal component
PCA	Principal component analysis
PLS	Partial least squares
r_{pre}	Correlation coefficient of prediction
RMSEP	Root mean square error of prediction
ROI	Regions of interest
SIMCA	Soft independence modeling of class analogy
SPA	Successive projection algorithm
SSC	Soluble solids content
SVM	Support vector machines

Introduction

Persimmon (*Diospyros*) is native to East Asia and has been cultured in Italy, Korea, Brazil, Israel, Spain, the USA, New Zealand, and Australia (Telis et al., 2000). The fully matured fruits are sweet and palatable. It has been proved they contain many medicinally bioactive compounds, such as carotenoids, tannins, flavonoids, terpenoids, steroids, naphthoquinones, sugars, amino acids, minerals, and lipids (Mallavadhani et al., 1998).

Usually, persimmon fruit are classified as non-astringent and astringent types. Mature fruit of non-astringent cultivars can be eaten at the firm-stage when the pulp is crisp. On the contrary, fruit of astringent cultivars need some postharvest treatments, such as exposure to carbon dioxide in high concentrations, appropriate ethylene treatment or dried after peeling to remove astringency. Without those treatments, the fruit cannot be consumed until the flesh is very soft. In the past, astringent cultivars were commonly eaten as soft or dried fruit.

X. Wei · F. Liu · Z. Qiu · Y. Shao · Y. He (✉)
College of Biosystems Engineering and Food Science Zhejiang University, 866 Yuhangtang Road, Hangzhou 310058, China
e-mail: yhe@zju.edu.cn

Even now, soft or dried fruit are consumed and commercially produced in countries such as China, Italy, Japan, and Korea (Yamada et al., 2012).

In China, the main persimmon cultivar is the astringent type, which is harvested at the firm-stage and consumed after becoming soft. The method adopted to distinguish the ripening stage of persimmon is still relying on the people's experience and intuitionistic judgment. This approach is inefficient and subjective, and, thus, is unable to cope with large-scale production and trading, as well as modern fruits processing. Consequently, advanced technology for nondestructive measurements of postharvest quality and ripeness classification of persimmon is necessary.

In the 1980s, researchers began paying attention to the nondestructive technology on the ripeness evaluation of persimmon. Yamazaki and Suzuki (1980) developed color charts that were used commercially to determine optimum harvest time for different cultivars. Forbus et al. (1991) showed that delayed light emission had a high potential for application as a rapid, nondestructive technique for measuring persimmon maturity. Clark and MacFall (2003) applied qualitative and quantitative proton magnetic resonance imaging technique on persimmon during development and postharvest ripening. Taniwaki et al. (2009) used laser Doppler vibrometer to determine the changes in the physical properties of persimmon fruit during ripening. Zhang et al. (2011a, b) investigated the firmness and soluble solids content (SSC) of persimmon by visible and near-infrared spectroscopy.

Hyperspectral imaging is an emerging technique that integrates conventional imaging and spectroscopy to acquire both spatial and spectral information from a sample. It takes the advantages of the conventional RGB, near-infrared spectroscopy, and multispectral imaging. Each pixel in a hyperspectral image contains an entire spectrum also called spectral signature or spectral fingerprint, which mainly represent the light absorption properties of the object under investigation. The total spectral signature may also be influenced by the scattering properties of the tissue (Kamruzzaman et al., 2012). The three-dimensional data can be analyzed to characterize the physical and chemical features of the product being examined and to characterize the product in a way more reliably than the imaging or spectroscopy technique (Elmasry et al., 2012). The resulting spectrum acts like a fingerprint which can be used to characterize the composition of that particular pixel. It allows for the visualization of biochemical constituents of a sample, separated into particular areas of the image, since regions of a sample with similar spectral properties have similar chemical composition (Gowen et al., 2007). In recent years, there has been a growing interest in this technology from researchers for nondestructive analysis in fruits (Huang et al., 2012; Mendoza et al., 2011; Qin & Lu, 2008). Detection of the ripening process

also has got some attentions; the assessment of fruits ripening such as peach and banana have been investigated based on hyperspectral imaging technique (Cen et al., 2011; Lleo et al., 2011; Rajkumar et al., 2012), which showed a potential use of hyperspectral imaging for the fruits maturity/quality assessment. However, major bottlenecks such as high costs and difficulties in high-speed data acquisition and processing have limited the use of this technology in a real-time assessment (Kamruzzaman et al., 2013). The advantage of obtaining abundant information makes it a very useful tool for selecting some important wavelengths and what is more important is that it can be used to know what attributes, how much, and where they are located in the sample under study (Lorente et al., 2012). Then the on-line grading system could be done by using multispectral imaging technique (Rajkumar et al., 2012).

In this study, we mainly adopted successive projection algorithm (SPA) and level co-occurrence matrix (gray level co-occurrence matrix (GLCM)) method to acquire the different input of classifiers based on the spectra and images. SPA is a popular tool for variable selection in multivariate calibration and classification, the goal of which is finding a small, representative set of spectral variables with an emphasis on the minimization of collinearity (Pontes et al., 2005). In machine visual systems, texture feature analysis plays an important role for classification and presentation. And among them, the most popular method is the GLCM method. It is based on the second-order texture features, which takes into account the specific position of a pixel relative to another. It has been widely used for texture analysis in many applications (Gomez et al., 2012; Tournier et al., 2012; Yang et al., 2012).

We combined the feature spectra and images to ripening stage classification of persimmon fruit based on hyperspectral imaging technique in the visible and short-wave near-infrared spectral region between 380 and 1,030 nm. Using hyperspectral images, the objectives of this work are as follows: (1) find whether some specific locations could be used to persimmon ripeness classification; (2) optimize the calibration models by wavelength feature and texture feature using GLCM parameters; and (3) analysis the relationship between firmness and ripeness, provide a method to persimmon classification for further processing such as making persimmon cakes.

Materials and Methods

Persimmon Samples

In this study, persimmons of local astringent cultivars (*Diospyros oleifera* Cheng 'Fangshi') were used that are widely planted in eastern China for soft consumption. During the harvest season, fruit of uniform size were harvested from

orchards in Hangzhou (30°16′ N and 120°3′ E) when the entire fruit were mostly yellow with little green. After the arrival in the laboratory, samples were kept at 25 °C for 24 h. Their physical parameters were measured before hyperspectral image acquisition. Mass was determined by an electronic balance (YP2001N, Shanghai Precision & Scientific Instrument Co., Ltd., Shanghai, China). The perimeter and height of the fruit were obtained with ruler tape and vernier caliper (Shenhan measurements tools Co., Ltd., Shanghai, China), respectively. After the hyperspectra acquisition, the SSC was measured by an Abbebenchtop refractometer (WAY-2S, Shanghai Precision and Scientific Instrument Co., Ltd., Shanghai, China). All the measurements of parameters were repeated three times. Mass/height/perimeter/firmness/SSC was used as dependent and correspondingly ripeness as factor, respectively in one-way ANOVA by IBM SPSS (Version 20.0, Inc., Chicago, IL, USA).

Acquisition of Images

Experimental Procedure

Samples were classified into four groups: unripe, mid-ripe, ripe, and over-ripe according to the days (d) after harvest in this work. The fruit were harvest with a little green at the stylar end side. They were preserved in the laboratory for 5 to 7 d when all the green in skin was receded, and these were considered as the unripe group. Similarly, fruits kept for 8 to 11 d, 12 to 14 d, and 15 to 16 d from the harvest time were considered as the mid-ripe, ripe, and over-ripe group, respectively. During ripening, the climacteric persimmon fruits produce large quantities of ethylene (Kader, 1999), and the flesh becomes softer gradually. In order to remove outliers, firmness of every sample was measured three times at their equator by a fruit hardness tester (FHM-5, Takemura Electric Work. Ltd., Japan), after acquisition of the hyperspectra. Using Grubbs's test, calculated mean value and standard derivation were applied to identify and eliminate samples with extreme firmness for each group.

Hyperspectral Images Acquisition

For hyperspectral imaging of persimmon, a laboratory-type spectrum measurement device was designed (Fig. 1). The system mainly consisted of a high-performance back-illuminated CCD camera (C8484-05, Hamamatsu City, Japan), an imaging spectrograph (ImSpectorV10, Spectral Imaging Ltd., Oulu, Finland) covering the spectral range of 380–1,030 nm, and an assembled light unit containing two 150-w quartz tungsten halogen lamps (Oriel Instruments, USA). A mobile platform (ST-1212-300, Tanlian Co., Taiwan, China), with its speed controlled by a computer equipped with spectral image system V10E software (ISUZU OPTICS COPR,

Taiwan, China), was used for sample movement. While the samples were moved on the X-axis, the detector linearly scanned along the Y-axis. The acquired images were corrected with white and dark reference using Eq. 1 as:

$$R = \frac{I_{\text{sample}} - I_{\text{dark}}}{I_{\text{white}} - I_{\text{dark}}} \tag{1}$$

Calibration of the images was done by the hyperspectral imaging analyzer software. Considering the difference between the stylar end side and the stem-calyx end side of persimmon, we acquired the hyperspectral images of two sides one after another.

Data Analysis

### Feature Wavelengths Selection

Spectral data were collected from the images preprocessed by the ENVI (Version 4.7, ITT Visual Information Solutions, Boulder, USA) software. Four square-shaped regions of interest (ROI) in the same plane at 90° interval with about 6,000 pixels in total were acquired (Fig. 2).

Feature wavelengths were selected by the SPA method, which starts with one wavelength, then incorporates a new one at each iteration, until a specified number (N) of wavelengths is reached (Araújo et al., 2001). SPA was coded in Matlab 7.10 R2010b (The MathWorks, Natick, USA).

Texture Feature Acquisition

Texture is one of the essential characteristics used in identifying objects or regions of interest of an image. GLCM, a popular texture analysis method, estimates image properties related to the second-order statistics. The image in its digital form is usually stored L_y in the computer as a two-dimensional array. If $L_x = (1, 2, \cdots, N_x)$ and $L_y = (1, 2, \cdots, N_y)$ are the horizontal and vertical spatial domains, then $L_x \times L_y$ is the set of resolution cells and the image I is a function which assigns some gray-tone value $G = (1, 2, \cdots, N)$ to each cell; $I:L_x \times L_y \rightarrow G$. Texture information contained in image I is defined as specified by the statistical distribution of the spatial relationships of gray level properties. This information can be summarized in GLCM that are matrices of relative frequencies $P(i; j; d; \theta)$ with which two neighboring pixels separated by distance d at orientation θ occur in the image, one with gray level i and the other with gray level j. In this research, the textural features were calculated from GLCM when the direction equals to 0° and the distance equals to 1, respectively. In order to estimate the similarity among different GLCM, Haralick et al. (1973) proposed a set of statistical features extracted from them. To reduce the computational

Fig. 1 Schematic diagram of hyperspectral imaging system
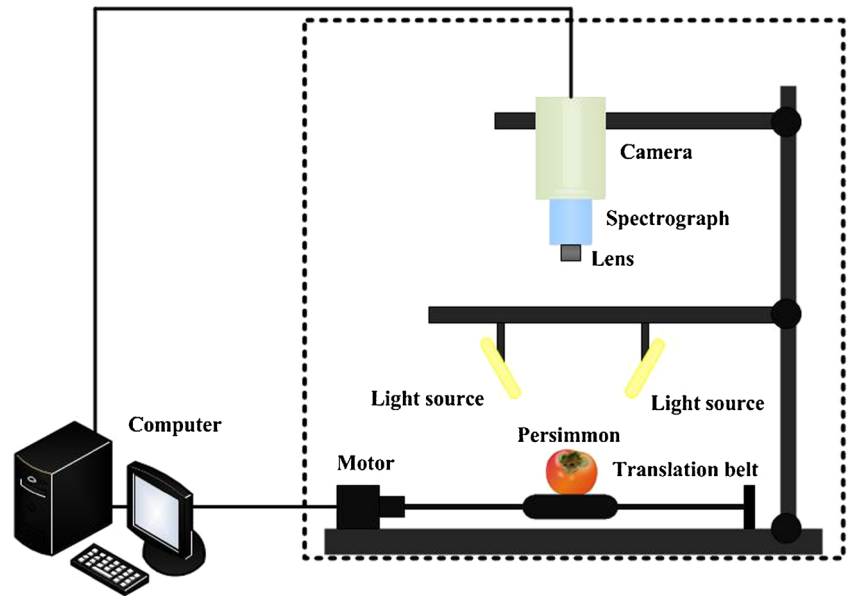
complexity, only four features were used for analysis. They were widely (ElMasry et al., 2007; Li et al., 2011; Tournier et al., 2012) used to select the optimal feature and indicated in Table 1.

Correlation feature is a measure of gray-tone linear dependencies in an image. Contrast is a difference moment of the P and it measures the amount of local variations in an image. Energy, also called angular second moment, is a measure of textural uniformity of an image. Energy reaches its highest value when gray level distribution has either a constant or a periodic form. Homogeneity, also called inverse difference moment, usually measures the closeness of the distribution of elements in GLCM to its diagonal (Gadelmawla, 2004).

Classification Models

The classification models were developed by the linear discriminant analysis (LDA), soft independence modeling of

class analogy (SIMCA), and least squares support vector machines (LS-SVM) methods.

LDA (also known as Fisher discriminant analysis) is a common supervised recognition method used in statistics, pattern recognition, and machine learning to find a linear combination of features that characterizes or separates two or more classes of objects or events (Wu et al., 1996).

In the SIMCA method, each class is regarded as an independent group and the characteristics of each class are extracted by principal component analysis (PCA). The distance of orthogonal projection between the principal component (PC) model of each class and the unknown data is calculated to determine class of the unknown (Maesschalck et al., 1999).

(a) calyx end side   **(b) stylar end side**
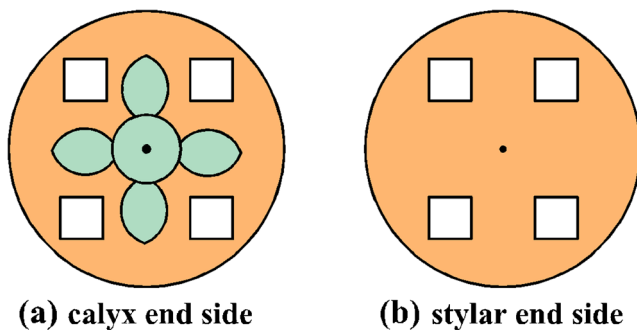
Fig. 2 Four square-shaped ROIs got from **a** stem-calyx end side, **b** stylar end side, respectively

Table 1 Expressions of textural features from GLCM

Texture features	Equation
Correlation	$\left[\sum_i \sum_j (ij)P(i,j) - \mu_x \mu_y\right] / \sigma_x \sigma_y$
Contrast	$\sum_i \sum_j (i-j)^2 P(i,j)$
Energy	$\sum_i \sum_j P(i,j)^2$
Homogeneity	$\sum \sum P(i,j)/(1 + (i-j)^2)$

Where $P(i,j)$ is $(i,j)th$ entry of the co-occurrence probability matrix, $\mu_x = \sum_i \sum_j i \cdot P(i,j)$, $\mu_y = \sum_i \sum_j j \cdot P(i,j)$, $\sigma_x = \sqrt{\sum_i \sum_j (i, \mu_x)^2 P(i,j)}$, $\sigma_y = \sqrt{\sum_i \sum_j (j, \mu_y)^2 P(i,j)}$

LS-SVM is the least squares versions of support vector machines (SVM) (Suykens & Vandewalle, 1999), which has been proved to be an effective and robust method for both classification and regression based on the structural risk minimization principle. With equality constraints in the formulation, LS-SVM obtains the solution by solving a set of linear equations, instead of quadratic programming for traditional SVM algorithms. More details about this algorithm can be found in other literatures (Borin et al., 2006; Chauchard et al., 2004; Liu et al., 2009).

Regression Model

The firmness prediction were conducted by the partial least squares (PLS) regression method, which generalizes and combines features from PCA and multiple regressions and has been widely applied in many areas (Haaland & Thomas, 1988).

Results and Discussions

Samples Selection

The changes in fruit firmness and SSC are closely associated with ripeness (Semmelmeyer & Ernst, 2006). It has been shown that the SSC changes during ripening were cultivar specific in persimmons (Glew et al., 2005). However, in fruits of non-astringent cultivars, firmness always declines during ripening, so samples with extreme firmness were eliminated for each group. Finally, 192 samples (48 for each stage) were adopted for further calculation. Mass and size of the samples did not change with maturation, but firmness decreased significantly as fruit gradually ripened. Mean SSC tended to increase with development but peaked at the over-ripe stage (Table 2).

The result of the ANOVA analysis (Table 3) of physical parameters and SSC in different ripening stage showed that there was no apparent discrepancy between the mass/height/perimeter of samples at different ripeness. Firmness presented significant differences among the groups and had a descending tendency along with the development of ripeness. The difference of SSC among unripe, mid-ripe, and ripe groups

Table 3 One-way ANOVA of quality parameters of persimmons

Parameters	Test of homogeneity		ANOVA			
	Levene statistic	Sig	Combine		Linear term	
			F	Sig.	F	Sig.
Mass	1.639	0.186	1.453	0.232	0.255	0.615
Height	0.654	0.582	1.125	0.169	0.232	0.633
Perimeter	0.335	0.800	0.772	0.513	0.202	0.654
Firmness	45.01	0.000	699.3	0.000	1,917.2	0.000
SSC	24.48	0.000	82.04	0.000	177.1	0.000

was not significant but both of them were significantly different with those of the over-ripe group.

Spectra Consideration

After removing the noise in the data, only wavelengths in the range from 400 to 1,000 nm were used for further analysis. In the average spectra of each group (Fig. 3a), it was observed that reflectance of fruit of both ripe and over-ripe fruits were lower than in that of unripe and mid-ripe ones. Some studies showed that with the change of ripeness, the contents and compositions of sugars, amino acids, carotenoids, and lipids were different. For example, the total carotenoid decreased until color break, then steadily gradually and drastically at the final ripening stage (Ebert & Gross, 1985). But there was no change in carotenoid absorption area (400 to 500 nm) of the spectra, and the presence of water could be observed clearly around 960 nm (Golic et al., 2003). So the change of spectra (600–900 nm) could not be the pigment but probably due to water and carbohydrate changes or scattering changes caused by tissue structure changes. Comparison of the spectra of all investigated fruit (Fig. 3b), however, showed that the different ripening stages could not be easily discriminated through the origin spectra. There were some overlaps especially between the neighboring groups.

Classification Models Based on Original Spectra

Firstly, the entire original spectra data were used for classification. Considering the discrepancy between stylar end side

Table 2 The physical parameters, firmness, and SSC of persimmons for each class

Classes	No.	Mass (N)	Height (mm)	Perimeter (mm)	Firmness (N)	SSC (Brix)
Unripe	48	14.22±0.34	45.3±4.57	223.9±1.78	25.0±4.37a	16.0±0.65b
Mid-ripe	48	14.84±0.28	45.0±6.02	227.4±1.62	12.2±2.15b	16.4±2.20b
Ripe	48	14.66±0.32	44.9±5.00	226.8±1.82	3.16±1.39c	16.8±3.38b
Over-ripe	48	14.04±0.27	44.6±4.60	225.1±1.39	1.32±0.36d	22.1±1.62a

Mean values in the same row with different letters are significantly different ($p \leq 0.05$)
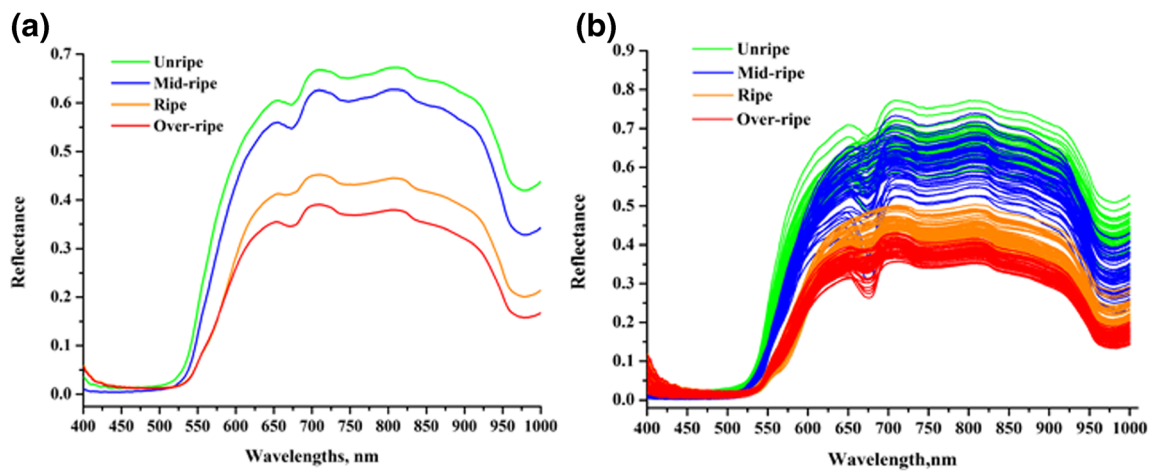
Fig. 3 Average spectra (**a**) and spectra (**b**) of all samples of four different ripen classes (the stylar end side)

and stem-calyx end side, we generated three different spectral subsets: (1) spectra of stylar end side, (2) spectra of stem-calyx end side, and (3) average spectra of the two sides. For each group, 2/3 of the fruits were randomly chosen (128 samples) as the calibration set and the rest 1/3 (64 samples) were adopted as the predicted set. As LDA would be difficult to calculate when the number of samples was far less than variables, only the SIMCA and LS-SVM classification models were developed to make a comparison. Their performances were presented in Table 4.

It can be seen that both of the two classifiers using spectra of the stem-calyx end side had better performance than those using spectra of stylar end side. The correct classification rate (CCR) of the later was higher with 9.4 and 6.2 % for the SIMCA and LS-SVM classifier, respectively. A possible explanation for this result was that in the visible and short-wave near-infrared spectral regions, the wavelengths might represent the water, carbohydrate, and texture differences of the fruits. The postharvest maturation of this cultivar often begins from the stylar end of persimmon, so the changes come first at this end. In this part, the characteristics of fruit at a given

period were close to the status at the next ripening stage; hence, similar chemical constituents would result in some interference. In comparison of the power of classifiers with the third type input and with the first type input, the accuracies were close between the two. As in production, reducing the calculations can certainly improve the detective speed, the spectra and images of stem-calyx end side were used for further analysis.

Optimal Wavelengths Selection

Generally, it is important to select the wavelengths that contribute to the quality attribute of the product. The hyperspectral images contain abundant information, which is time consuming for data processing and unsuitable for on-line detection. In order to reduce the dimensional and redundant information, SPA was applied to select the feature wavelengths. This processing was conducted in Matlab.

Table 4 The classification results of the ripeness by SIMCA and LS-SVM with the original spectra

Input	Classifier	Calibration			Prediction		
		No.	Missed	CCR (%)	No.	Missed	CCR (%)
1	SIMCA	128	33	74.2	64	18	71.9
	LS-SVM	128	8	93.8	64	6	90.6
2	SIMCA	128	40	68.8	64	24	62.5
	LS-SVM	128	12	90.6	64	10	84.4
3	SIMCA	128	34	73.4	64	17	73.4
	LS-SVM	128	9	93.0	64	7	89.1

CCR correct classification rate, *1* spectra of stem-calyx end side, *2* spectra of stylar end side, *3* average spectra of two sides

Fig. 4 Scores scatter plot of PC1 and PC2 of PCA conducted on feature wavelengths

Table 5 Classification details of the prediction set by different models using the feature wavelengths selected by SPA

Models	Groups	Classification				Total	CCR%
LDA	1	15	1	–	–	16	93.8
	2	2	13	1	–	16	81.2
	3	–	–	12	4	16	85.7
	4	–	–	1	15	16	93.8
Total		17	14	14	19	64	85.9
SIMCA	1	9	5	–	–	16	56.2
	2	2	13	–	–	16	81.2
	3	–	–	10	2	16	62.5
	4	–	–	4	12	16	75.0
Total		11	18	5	14	64	68.8
LS-SVM	1	11	5	–	–	16	68.8
	2	3	12	1	–	16	75.0
	3	–	–	14	2	16	87.5
	4	–	–	1	15	16	93.8
Total		14	17	16	17	64	81.2

1 unripe stage, *2* mid-ripe stage, *3* ripe stage, *4* over-ripe stage

After calculating, three wavelengths (518, 711, and 980 nm) were chosen.

PCA conducted on the spectra of selected optimal wavelengths of the calibration set displayed that the first two components explained 99.9 % (PC1—97.4 % and PC2—2.5 %) of the differences of persimmon fruit at different ripening stages (Fig. 4). Most of the observed values were distinguished into different groups. It indicated that the

selected wavelengths could be used for discriminating the ripeness of persimmon.

Classification models were developed using only the selected wavelengths (Table 5). Compared to the models based on original spectra, the performance of both feature wavelengths-based SIMCA and LS-SVM models was worse. CCR of SIMCA and LS-SVM decreased from 71.9 to 68.8 % and from 90.6 to 81.2 %, respectively. The LDA classifier with CCR of 89.1 % acquired the best result. The misclassification mostly occurred between the unripe and mid-ripe groups, and between the ripe and over-ripe groups. As the circles in Fig. 5 showed, the samples could be clearly classified into two sets (unripe and mid-ripe vs. ripe and over-ripe). This may imply that the chemical composition of fruit of the unripe and mid-ripe group and the ripe and over-ripe group, respectively, were rather similar. Results in Table 5 also revealed that in both the LDA and LS-SVM classifier, the highest CCR of 93.8 % appeared in the over-ripe group. In the analysis of SSC, the over-ripe group also showed significant difference with other three groups, so maybe both the outer and internal character appeared significant changes compared with the other groups. To sum up, except the SIMCA model, the other models achieved good predictive accuracy and the feature wavelengths could be considered valid.

Texture Features at the Images of Feature Wavelengths

One of the objectives for this work was to determine the calibration by feature wavelengths at texture feature using GLCM. In order to improve the accuracy of the classification

Fig. 5 Texture parameters of persimmon at different ripening stage using gray level co-occurrence matrix GLCM
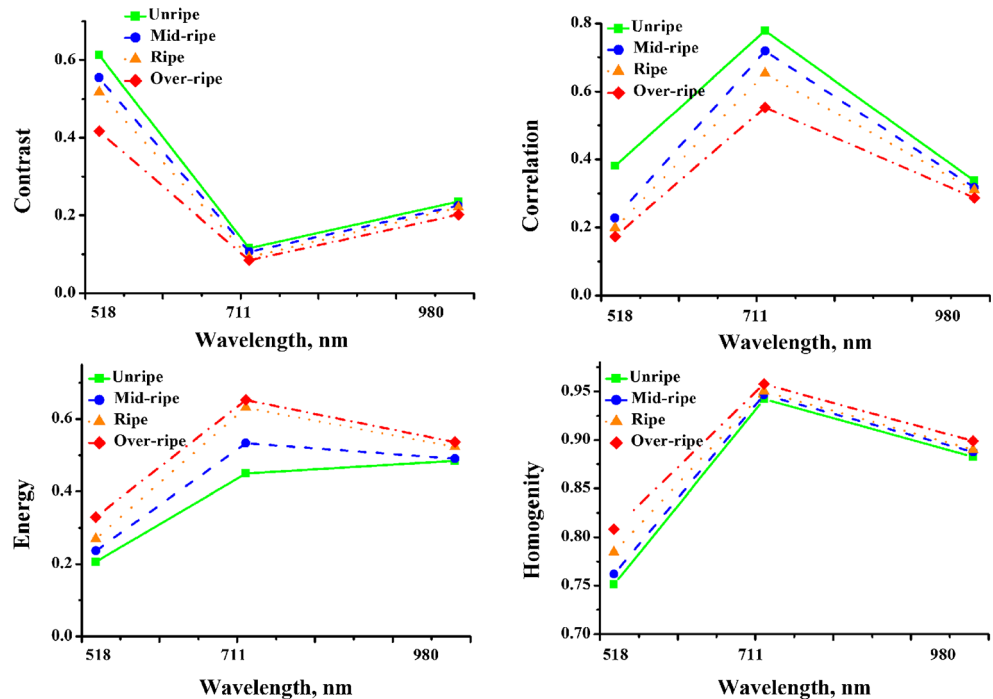
Table 6 The texture feature images at 518, 711, 980nm of persimmons of four different ripen classes

	518nm	711nm	980nm
Unripe			
Mid-ripe			
Ripe			
Over-ripe			

models, texture information was extracted from the images at feature wavelengths. Images were obtained from a part of ROIs which were used to extract spectra before through software ENVI at the stem-calyx end side; there were 576 images each with 50×50 pixels. Part of the images at the feature wavelengths is shown in Table 6. They exhibited some differences: images at 980 nm appeared brighter for the unripe and mid-ripe groups than the ripe and over-ripe groups clearly, which is the same for images at 711 nm but more obvious. This difference was not apparent at 518 nm.

The texture features correlation, contrast, energy, and homogeneity were calculated in Matlab. Figure 5 shows the means of the four texture parameters of persimmon at different ripening stages. It can be seen that contrast and correlation of the unripe group was the highest compared with the other three groups at different wavelengths. This means that unripe fruit contain higher local variations and more smooth texture (ElMasry et al., 2007). The similar conclusion also could be obtained by analyzing the mean value of homogeneity, which was the lowest for the unripe group with the highest contrast. The energy of fruits of the unripe group was lower than in those of the other groups, which meant that the unripe persimmons were less homogeneous (Baraldi & Parmiggiani, 1995). The texture parameters of 980 nm showed smaller variations. In general, as shown in Fig. 5, there were no intersections of all the parameters; therefore, it is possible for ripeness classification based on these statistics.

With these obtained statistical texture features, three different subsets were adopted as different inputs to the classifiers. The subsets were spectra of feature wavelengths, texture features at feature wavelengths, and spectra of feature wavelengths combined the texture features, respectively. Predictive results were illustrated in Fig. 6. From all of the models, the LDA model with the input of spectra combined the texture features had the highest CCR of 95.3 %, which increased by 9.4 % compared with the LDA model with the input of spectra of feature wavelengths. When only using the subset of texture features as the input, the LS-SVM classifier showed best result

with the CCR of 84.4 %. SIMCA always had the lowest CCR and when based on subsets of texture feature, it was entirely ineffective. Figure 7 showed the scatter plots of the prediction set by using the first two discriminant factors obtained from the LDA modeling with the spectra combined texture features. It revealed that there also were some difficulties to recognize some unripe samples from mid-ripe and to recognize some ripe from over-ripe by the LDA classifier. The distinction between the neighboring stages still was not very clear. Generally, almost all of the models in this work have made misclassifications in some samples, indicating the difficulties in distinguishing samples with similar physiological status.

Prediction of Firmness

Because of the firmness had obvious linear correlations with the maturity periods in ANOVA analysis, the input of classifiers also could act as the input of firmness prediction if the data processing strategy used to ripeness classification was
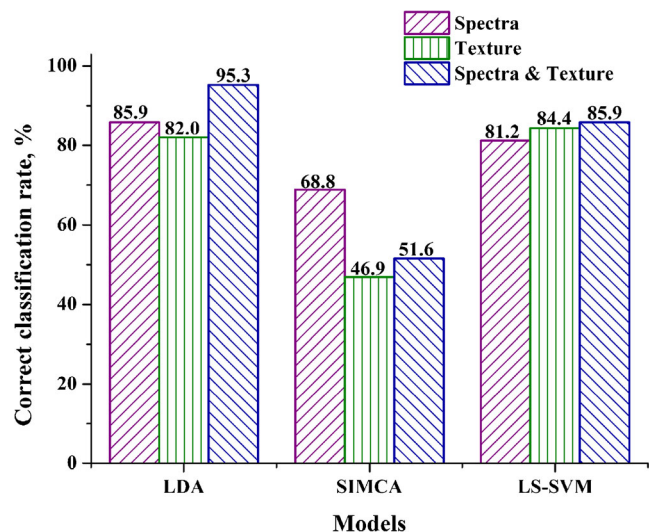

Fig. 6 Predictive results of different models using three types of subsets
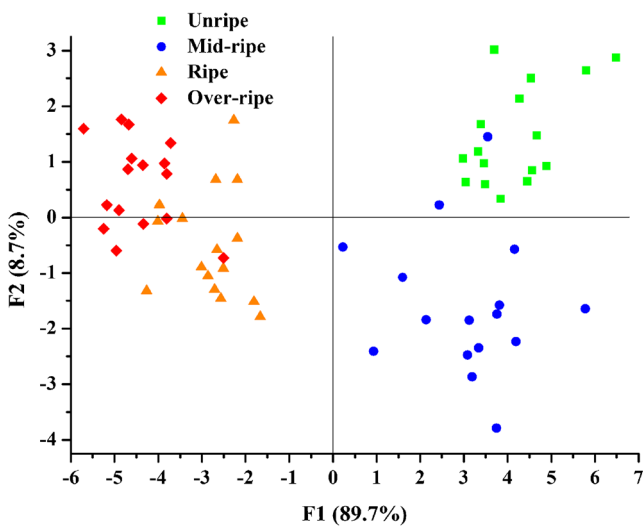
Fig. 7 Scatter plot of the prediction set by using the first two discriminant factors obtained from linear discriminant analysis modeling with the spectra and texture features

reasonable. So, we conducted firmness prediction using the PLS regression analysis with the spectra of feature wavelengths and the texture features. The model was evaluated by root mean square error of validation (RMSECV), root mean square error of prediction (RMSEP), and correlation coefficient (r) as shown in Fig. 8. The results with r_{pre} of 0.913 and RMSEP of 4.349 indicated that the selected wavelengths and texture feature could be used for the nondestructively prediction of persimmons firmness in this research. Kato (1990) showed that high-quality persimmon fruit should have firmness of 1.5 to 2.5 kg (14.7 to 24.5 N; measured with a Universal hardness meter), which could own the jelly-like texture, and had advantage for physical handling associated with marketing. So it is useful to determine fruit firmness according to the needs of persimmons industry.
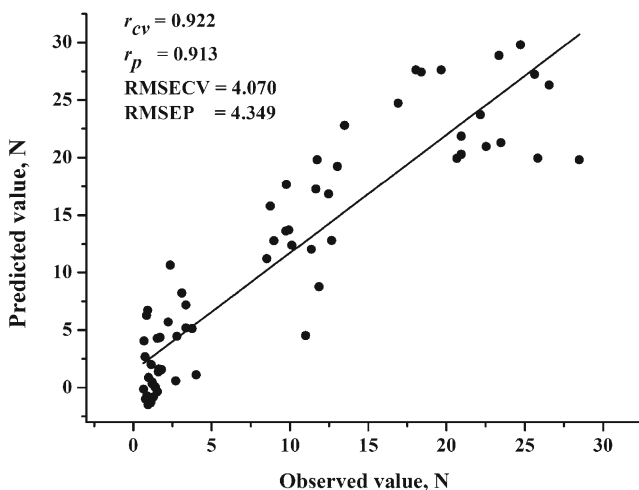
Fig. 8 Firmness predicted using the PLS models

Conclusions

The results of this study suggested that hyperspectral imaging together with chemometrics could be nondestructively used to classify the ripening stage of persimmons. Out of 476 wavelengths, only three wavelengths (518, 711, and 980 nm) were selected as the optimum wavelengths and found to be suitable for ripeness classification. Based on the spectra and texture at these three optimum wavelengths of fruits' stem-calyx end side, discrimination accuracy of 95.3 % was achieved using the LDA classifier. The same dataset for ripeness classification also got a good performance of firmness prediction by PLS in the regression analysis. More research is needed to include more samples as well as different regions and different postharvest treatments to ascertain the discrimination power of this method.

References

Araújo, M. C. U., Saldanha, T. C. B., Galvão, R. K. H., Yoneyama, T., Chame, H. C., & Visani, V. (2001). The successive projections algorithm for variable selection in spectroscopic multicomponent analysis. *Chemometrics and Intelligent Laboratory Systems, 57*(2), 65–73.

Baraldi, A., & Parmiggiani, F. (1995). An investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters. *Geoscience and Remote Sensing, IEEE Transactions on, 33*(2), 293–304.

Borin, A., Ferrão, M. F., Mello, C., Maretto, D. A., & Poppi, R. J. (2006). Least-squares support vector machines and near infrared spectroscopy for quantification of common adulterants in powdered milk. *Analytica Chimica Acta, 579*(1), 25–32.

Cen H., Lu R., Mendoza F.A. & Ariana D.P. (2011). Peach maturity/quality assessment using hyperspectral imaging-based spatially resolved technique. In: Proc. SPIE 8027, Sensing for agriculture and food quality and safety III, 80270L, 02 June 2011, Bellingham, USA.

Clark, C. J., & MacFall, J. S. (2003). Quantitative magnetic resonance imaging of 'Fuyu' persimmon fruit during development and ripening. *Magnetic Resonance Imaging, 21*(6), 679–685.

Chauchard, F., Cogdill, R., Roussel, S., Roger, J. M., & Bellon-Maurel, V. (2004). Application of LS-SVM to non-linear phenomena in NIR spectroscopy: development of a robust and portable sensor for acidity prediction in grapes. *Chemometrics and Intelligent Laboratory Systems, 71*(2), 141–150.

De Maesschalck, R., Candolfi, A., Massart, D. L., & Heuerding, S. (1999). Decision criteria for soft independent modelling of class analogy applied to near infrared data. *Chemometrics and Intelligent Laboratory Systems, 47*(1), 65–77.

Ebert, G., & Gross, J. (1985). Carotenoid changes in the peel of ripening persimmon (*Diospyros kaki*) cv Triumph. *Phytochemistry, 24*(1), 29–32.

Elmasry, G., Kamruzzaman, M., Sun, D.-W., & Allen, P. (2012). Principles and applications of hyperspectral imaging in quality evaluation of agro-food products: a review. *Critical Reviews in Food Science and Nutrition, 52*(11), 999–1023.

ElMasry, G., Wang, N., ElSayed, A., & Ngadi, M. (2007). Hyperspectral imaging for nondestructive determination of some quality attributes for strawberry. *Journal of Food Engineering, 81*(1), 98–107.

Forbus, W. R., Payne, J. A., & Senter, S. D. (1991). Nondestructive evaluation of Japanese persimmon maturity by delayed light emission. *Journal of Food Science, 56*(4), 985–988.

Gadelmawla, E. S. (2004). A vision system for surface roughness characterization using the gray level co-occurrence matrix. *NDT & E International, 37*(7), 577–588.

Glew, R., Ayaz, F., Millson, M., Huang, H. S., Chuang, L. T., Sanz, C., & Golding, J. (2005). Changes in sugars, acids and fatty acids in naturally parthenocarpic date plum persimmon (*Diospyros lotus* L.) fruit during maturation and ripening. *European Food Research and Technology, 221*(1–2), 113–118.

Golic, M., Walsh, K., & Lawson, P. (2003). Short-wavelength near-infrared spectra of sucrose, glucose, and fructose with respect to sugar concentration and temperature. *Applied Spectroscopy, 57*(2), 139–145.

Gomez, W., Pereira, W. C. A., & Infantosi, A. F. C. (2012). Analysis of co-occurrence texture statistics as a function of gray-level quantization for classifying breast ultrasound. *IEEE Transactions On Medical Imaging, 31*(10), 1889–1899.

Gowen, A. A., O'Donnell, C. P., Cullen, P. J., Downey, G., & Frias, J. M. (2007). Hyperspectral imaging—an emerging process analytical tool for food quality and safety control. *Trends in Food Science and Technology, 18*(12), 590–598.

Haaland, D. M., & Thomas, E. V. (1988). Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. *Analytical Chemistry, 60*(11), 1193–1202.

Haralick, R. M., Shanmugam, K., & Dinstein, I. H. (1973). Textural features for image classification. *Systems Man and Cybernetics IEEE Transactions on, SMC-3*(6), 610–621.

Huang, M., Zhu, Q., Wang, B., & Lu, R. (2012). Analysis of hyperspectral scattering images using locally linear embedding algorithm for apple mealiness classification. *Computers and Electronics in Agriculture, 89*, 175–181.

Kato, K. (1990). Astringency removal and ripening in persimmons treated with ethanol and ethylene. *American Society for Horticultural Science Alexandria, 25*(2), 205–207.

Kader, A. A. (1999). *Fruit maturity, ripening, and quality relationships*. Leuven: International Society for Horticultural Science.

Kamruzzaman, M., Barbin, D., ElMasry, G., Sun, D.-W., & Allen, P. (2012). Potential of hyperspectral imaging and pattern recognition for categorization and authentication of red meat. *Innovative Food Science & Emerging Technologies, 16*, 316–325.

Kamruzzaman, M., Sun, D.-W., ElMasry, G., & Allen, P. (2013). Fast detection and visualization of minced lamb meat adulteration using NIR hyperspectral imaging and multivariate image analysis. *Talanta, 103*, 130–136.

Li, X. L., Nie, P. C., Qiu, Z. J., & He, Y. (2011). Using wavelet transform and multi-class least square support vector machine in multi-spectral imaging classification of Chinese famous tea. *Expert Systems with Applications, 38*(9), 11149–11159.

Liu, F., Jiang, Y. H., & He, Y. (2009). Variable selection in visible/near infrared spectra for linear and nonlinear calibrations: a case study to determine soluble solids content of beer. *Analytica Chimica Acta, 635*(1), 45–52.

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. *Journal of Food Engineering, 104*(4), 612–620.

Lorente, D., Aleixos, N., Gómez-Sanchis, J., Cubero, S., García-Navarrete, O. L., & Blasco, J. (2012). Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. *Food and Bioprocess Technology, 5*(4), 1121–1142.

Mallavadhani, U. V., Panda, A. K., & Rao, Y. R. (1998). Pharmacology and chemotaxonomy of diospyros. *Phytochemistry, 49*(4), 901–951.

Mendoza, F., Lu, R., Ariana, D., Cen, H., & Bailey, B. (2011). Integrated spectral and image analysis of hyperspectral scattering data for prediction of apple fruit firmness and soluble solids content. *Postharvest Biology and Technology, 62*(2), 149–160.

Pontes, M. J. C., Galvão, R. K. H., Araújo, M. C. U., Moreira, P. N. T., Neto, O. D. P., José, G. E., & Saldanha, T. C. B. (2005). The successive projections algorithm for spectral variable selection in classification problems. *Chemometrics and Intelligent Laboratory Systems, 78*(1–2), 11–18.

Qin, J., & Lu, R. (2008). Measurement of the optical properties of fruits and vegetables using spatially resolved hyperspectral diffuse reflectance imaging technique. *Postharvest Biology and Technology, 49*(3), 355–365.

Rajkumar, P., Wang, N., Eimasry, G., Raghavan, G. S. V., & Gariepy, Y. (2012). Studies on banana fruit quality and maturity stages using hyperspectral imaging. *Journal of Food Engineering, 108*(1), 194–200.

Semmelmeyer, & Ernst. (2006). *OECD guidance on objective testing to determine the ripeness of fruit (Vol. 14)*. Skierniewice: Research Institute of Pomology and Floriculture.

Suykens, J. A. K., & Vandewalle, J. (1999). Least squares support vector machine classifiers. *Neural Processing Letters, 9*(3), 293–300.

Taniwaki, M., Hanada, T., & Sakurai, N. (2009). Postharvest quality evaluation of "Fuyu" and "Taishuu" persimmons using a nondestructive vibrational method and an acoustic vibration technique. *Postharvest Biology and Technology, 51*(1), 80–85.

Telis, V. R. N., Gabas, A. L., Menegalli, F. C., & Telis-Romero, J. (2000). Water sorption thermodynamic properties applied to persimmon skin and pulp. *Thermochimica Acta, 343*(1–2), 49–56.

Tournier, C., Grass, M., Zope, D., Salles, C., & Bertrand, D. (2012). Characterization of bread breakdown during mastication by image texture analysis. *Journal of Food Engineering, 113*(4), 615–622.

Wu, W., Mallet, Y., Walczak, B., Penninckx, W., Massart, D. L., Heuerding, S., & Erni, F. (1996). Comparison of regularized discriminant analysis linear discriminant analysis and quadratic discriminant analysis applied to NIR data. *Analytica Chimica Acta, 329*(3), 257–265.

Yamada M., Giordani E. & Yonemori K. (2012). Persimmon. In M. L. Badenes & D. H. Byrne (Eds.), *Fruit Breeding*, vol. 8 (pp. 663–693). New York: Springer.

Yamazaki, T., & Suzuki, K. (1980). Colour charts: useful guide to evaluate fruit maturity. 1. Colorimetric specifications of colour charts for Japanese pear, apple, peach, grape, persimmon, and citrus fruits. *Bulletin of the Fruit Tree Research Station A, 7*, 19–44.

Yang, X., Tridandapani, S., Beitler, J. J., Yu, D. S., Yoshida, E. J., Curran, W. J., & Liu, T. (2012). Ultrasound GLCM texture analysis of radiation-induced parotid-gland injury in head-and-neck cancer radiotherapy: an in vivo study of late toxicity. *Medical Physics, 39*(9), 5732–5739.

Zhang, P., Li, J. K., Meng, X. J., Zhang, P., Wang, B. G., & Xiao, Y. F. (2011a). Nondestructive determination of soluble solid content in Mopan persimmon by visible and near-infrared diffuse reflection spectroscopy. *Food Science, 6*, 191–194.

Zhang, P., Li, J. K., Meng, X. J., Zhang, P., Wang, B. G., & Xiao, Y. F. (2011b). Study on nondestructive measurement of firmness of Mopan persimmon using visible and near infrared diffuse reflection spectroscopy. *Science and Technology of Food Industry, 7*, 423–425,428.