

Potential of Hyperspectral Imaging for Rapid Prediction of Anthocyanin Content of Purple-Fleshed Sweet Potato Slices During Drying Process

Yunhong Liu¹ · Yue Sun¹ · Anguo Xie¹ · Huichun Yu¹ · Yong Yin¹ · Xin Li¹ · Xu Duan¹

Received: 22 March 2017 / Accepted: 21 May 2017 / Published online: 7 June 2017
© Springer Science+Business Media New York 2017

Abstract This study aimed to investigate the potential of hyperspectral imaging technique in tandem with chemometrics analysis for rapid and nondestructive determination of anthocyanin content within purple-fleshed sweet potato (PFSP) during drying process. Hyperspectral images of PFSP in the spectral range of 371–1023 nm were obtained during contact ultrasound-assisted hot air drying (CUHAD) process, and the reference anthocyanin contents of PFSP were measured by a traditional method. Partial least square regression (PLSR) and least-square support vector machine (LS-SVM) were applied to establish the calibration models based on raw extracted spectrum and spectrum preprocessed by four different methods. In order to simplify the calibration model, three algorithms including PLSR, LS-SVM, and multiple linear regression (MLR) were used to build models based on ten optimal wavelengths selected by regression coefficients (RC) method. The results showed that the RC-MLR yielded best results with the coefficient of determination for calibration (R_C^2) of 0.868 and coefficient of determination for prediction (R_P^2) of 0.866. Finally, distribution maps were developed based on an image processing algorithm to visualize anthocyanin content of PFSP at different drying periods which cannot be achieved by conventional methods. The overall results demonstrated that hyperspectral imaging technique is a useful tool for rapid and nondestructive determination of the anthocyanin content during drying process.

Keywords Hyperspectral imaging · Purple-fleshed sweet potato · Anthocyanin content · Drying process · Nondestructive analysis

Introduction

In recent years, increasing concern has been paid to the study of purple-fleshed sweet potato (PFSP) which is an excellent source of many nutrients such as dietary fiber, chlorogenic, ferulic acid, flavonoid, carotenoid, and anthocyanin (Kotíková et al. 2016; Cipriano et al. 2015). PFSP is widely cultivated in China, Japan, Korea, and New Zealand and could be produced as juice, powder, brewed drink, and dye (Steed and Truong 2008; Yoshimoto et al. 2005). However, fresh PFSP is easily deteriorated due to its high moisture content; therefore, certain technique such as drying is often used to remove the large amount of water for restraining microbiological and physico-chemical changes to extend the storage time and protect the quality (Dincer et al. 2002; Akpınar et al. 2003; Liu et al. 2014). Nevertheless, there are several factors during drying process including high temperature, long drying time, and quick air velocity which would be relevant to the degradation of quality, leading to undesired food flavor, effective compound degradation, enzymatic browning, etc. Hence, quality supervision of agri-food products including PFSP during drying process is very essential.

Anthocyanins, as a group of very important bioactive constituents in PFSP, possess pharmacological properties such as scavenging reactive oxygen species, decreasing capillary permeability and fragility, attenuating the proliferation of hepatic stellate cells, resisting mutagenesis or carcinogenic activity, and improving visual acuity (Li et al. 2013). What is more, anthocyanins obtained from PFSP have shown stronger radical-scavenging activity than anthocyanins from red

✉ Yunhong Liu
lyunhong@hotmail.com

¹ College of Food and Bio-engineering, Henan University of Science and Technology, Luoyang 471003, China

cabbage, elderberry, grape skin, and purple corn (Kubow et al. 2016). High concentration anthocyanins in PFSP have led to an interest in their applications as food colorants and dietary antioxidants due to their health benefits (Li et al. 2013). Therefore, the evaluation and monitor of anthocyanins are very important for obtaining dried PFSP products with high quality.

Several methods available for detecting anthocyanin content are normally classical assays such as ultraviolet-visible spectrophotometry and high-performance liquid chromatography (HPLC) (Ruttarattanamongkol et al. 2016; Liang et al. 2008). These methods are precise, but time-consuming, inefficient, destructive, and expensive (Xiong et al. 2015a). They are not suitable for fast or on-line determination of anthocyanin content. Consequently, rapid, accurate, and non-destructive technique should be applied to monitor the PFSP quality during drying. Visible-near-infrared (Vis-NIR) spectroscopy as one of the promising optical detection technique has been proven to be able to fast and nondestructively monitor food quality (Wu et al. 2012). Studies have been reported on the ability of spectroscopy technology for evaluating the anthocyanin content of fruit and vegetables such as berries (Ribera-Fonseca et al. 2016), sweet cherry (Pappas et al. 2011), black Goji berry (Li et al. 2016a), and flowering tea (Huang et al. 2014a). However, spectroscopy can only obtain the spectral information without acquiring the spatial information (Wu et al. 2012).

Hyperspectral imaging as one of the successfully used spectral imaging techniques has the ability to acquire both spatial information and spectral information simultaneously (Zhang et al. 2017). Recently, hyperspectral imaging technique has been applied successfully to determine vegetable quality, and there were some researches about the rapid detection of anthocyanin content in agricultural food. Yang et al. (2015a) used hyperspectral imaging technique coupled with model fusion to rapidly determine the anthocyanin content in lychee pericarp during storage and obtained the highest coefficients of determination (R^2) of 0.891 and 0.872 for the training sets and the testing sets, respectively. Chen et al. (2015) investigated the capability of hyperspectral imaging technique for predicting anthocyanin content of wine grapes during ripening, yielding good prediction accuracy with coefficient of prediction ($R_p^{2,P}$) of 0.9414, and root mean square error of prediction (RMSEP) of 0.0046. Moreover, it was reported that the anthocyanins in red grapes could be determined by hyperspectral imaging and chemometrics with R^2 of 0.72 and RMSE of 0.78 mg (Martínez-Sandoval et al. 2015). Anthocyanin content and other attributes such as brix, pH, and extractable polyphenols (total phenolic and flavanols) in whole port wine grape berries could also be simultaneously inspected by hyperspectral imaging

(Fernandes et al. 2015; Gomes et al. 2017; Nogales-Bueno et al. 2015). The above several studies show that the application of hyperspectral imaging is possible for predicting anthocyanin content of vegetables.

As for the potential of hyperspectral imaging for rapid determination of the quality of agricultural food during drying process, the moisture content changes of prawn, beef, mango slice, and grass carp slice were investigated (Wu et al. 2012, 2013; Pu and Sun 2015; Qu et al. 2017). Huang et al. (2014b) also demonstrated that the hyperspectral imaging had the capability for real-time determination of moisture content and color for soybean at different dehydrated levels. Although there were several researches about the application of hyperspectral imaging technique for nondestructive determination of moisture content and color during drying process, to the best of our knowledge, rare researches about the determination of other food quality attributes during drying process by hyperspectral imaging technique have been found. And although it is very essential to detect the anthocyanin content change of PFSP during drying process, the study about the nondestructive and rapid estimation of anthocyanin content during drying has not been reported yet. Therefore, this study concerned about the implementation of hyperspectral imaging for anthocyanin content inspection of PFSP during drying process.

In our previous research, the drying characteristics of contact ultrasound assisted hot air drying (CUHAD) for PFSP have been studied and the CUHAD method showed significant advantages to drying process comparing with traditional hot air drying (Liu et al. 2017). Therefore, the CUHAD technique is employed for the dehydration of PFSP in this study. The aim of this study was to investigate the potential application of hyperspectral imaging technique in visible and near infrared regions of 371–1023 nm for the anthocyanin content inspection of PFSP during CUHAD process. The specific aims of this article were to (1) acquire hyperspectral images of PFSP samples during CUHAD in the spectral region of 371–1023 nm; (2) extract average spectral data from region of interests in samples images; (3) build calibration models based on full wavelengths information and compare the performance of diverse spectral preprocessing methods; (4) select optimal wavelengths using regression coefficient algorithm and establish three regression algorithms based on the optimal wavelength information with the anthocyanin content of PFSP determined by traditional conventional method during different drying periods; and (5) develop image processing algorithms for the visualization of anthocyanin content of PFSP in all pixels within an image to form distribution maps of anthocyanin content of PFSP during drying process.

Materials and Methods

Sample Preparation and Drying Procedure

Fresh PFSP materials were obtained from local Danis supermarket in Henan Province and were stored at 2 to 4 °C in a refrigerator. A total of 15 fresh PFSP tubers were used in this experiment. Prior to drying experiment, the PFSP materials were washed, peeled, and cut into slices with 40 mm in diameter and 5 mm in thickness using a stainless-steel food slicer. The PFSP slices were dried by the CUHAD system which was elaborated in our previous study (Liu et al. 2017) and a similar drying process was carried out. In this study, the power and frequency of ultrasound were fixed at 60 W and 28 kHz, respectively. And the drying temperature and air velocity were set as 40 °C and 1 m/s, respectively. During drying process, nine groups at different drying periods (0, 0.5, 1, 2, 3, 4, 5, 6, and 7 h, respectively) were tested. In each group, 13~15 PFSP slices were removed from the drying system, consequently obtaining a total of 132 PFSP slices. The numbers of PFSP slices for each group are listed in Table 1. Then, the obtained samples were scanned by a hyperspectral imaging system and the anthocyanin contents were measured as described below.

Determination of Anthocyanin Content

In this study, anthocyanin contents in PFSP samples dried at different drying times were measured based on the method proposed by Yang et al. (2015a). The mass of each PFSP slice was recorded for every sample. Each sample was firstly crushed and then immersed in 1% HCl-methanol (50 mL) for 20 min with supersonic condition. Afterward, the solution

was filtered while the filter residue was further added with 1% HCl-methanol (50 mL). The procedure was repeated three times. All the filtered solutions were collected together and centrifuged at 10,000 rpm under 4 °C for 20 min in a cold centrifuge (TGL-16, Xiangyi Centrifuge Instrument Co., Hunan, China). Finally, the absorbance of the supernatants was measured at 530 nm and 600 nm, respectively, by ultraviolet-visible spectrophotometer (T6, Beijing Purkinje General Instrument Co., Beijing, China). The anthocyanin content of the sample was expressed by cyaniding-3-glucoside (CGE) and calculated by using the following equation:

$$\text{Anthocyanin} \left(\text{mg/g} \right) = (A_{530\text{nm}} - A_{600\text{nm}}) \times \frac{V \times n \times M_W}{\varepsilon \times m} \quad (1)$$

where $A_{530\text{nm}}$ and $A_{600\text{nm}}$ are the absorbance at 530 and 600 nm, respectively; V is the total volume of extract; n is the dilution ratio (it is one in this study); M_W is the relative molecular weight of CGE (449.4); ε is the molar extinction coefficient ($29,600 \text{ M}^{-1} \text{ cm}^{-1}$); and m is the mass of the sample.

Hyperspectral Imaging Acquisition and Calibration

PFSP samples dried at different drying times were taken out of the CUHAD system for hyperspectral image scanning. The slices were placed on the translation platform and then conveyed to the detector's field of view for scanning line by line. The speed of the translation platform was set as 1.1 mm/s and the exposure time for each hyperspectral image was 90 ms. The details of the hyperspectral imaging system can be found elsewhere (Sun et al. 2017). Therefore, a total of 132 images

Table 1 Statistics for the anthocyanin content of PFSP during drying process

Different drying period (h)	Calibration set						Prediction set					
	Amount of sample	Min ^a	Max ^b	Mean ^c	SD ^d	SE ^e	Amount of sample	Min	Max	Mean	SD	SE
0	8	2.677	3.457	2.900	0.245	0.087	5	2.776	3.412	2.932	0.241	0.108
0.5	10	1.641	3.862	2.433	0.661	0.209	5	2.028	2.987	2.515	0.385	0.172
1	11	1.255	2.520	2.084	0.406	0.122	4	1.270	2.176	1.736	0.337	0.169
2	9	1.130	2.353	1.483	0.428	0.142	6	1.223	2.094	1.631	0.362	0.148
3	10	0.843	1.353	1.152	0.121	0.038	5	0.914	1.301	1.081	0.152	0.068
4	10	0.740	1.131	0.882	0.100	0.032	4	0.862	1.018	0.949	0.106	0.053
5	10	0.528	1.113	0.824	0.181	0.057	5	0.741	0.950	0.813	0.072	0.032
6	9	0.516	0.799	0.623	0.087	0.029	6	0.588	0.685	0.648	0.032	0.013
7	11	0.253	0.677	0.580	0.141	0.043	4	0.535	0.677	0.640	0.061	0.031

^a Represents minimum value of anthocyanin content

^b Represents maximum value of anthocyanin content

^c Represents mean value of anthocyanin content

^d Represents standard deviation

^e Represents standard error

were obtained and then stored in a raw format before being processed. In order to minimize the influence of the dark current from camera and eliminate the differences in physical configuration of the imaging system, the original images were calibrated with white and black reference images according to a normal procedure (Xie et al. 2015).

After image acquisition and calibration, segmentation was performed to isolate the sample from the background, and the region of interest (ROI) should be recognized according to the spectrum difference between sample and background. In this study, the operation procedure was built as follows: hyperspectral image should be first open in ENVI 5.1 software (Research Systems Inc., Boulder, CO, USA) at the special spectral where sample and background exist a distinct difference in reflectance, then the preliminary ROI should be selected using the 'ROI tool', finally the 'Grow' function key should be pressed down. By this means, the abnormal pixels of the sample could be divided according to the deviation of spectra (Yang et al. 2017). And the spectral data within each ROI was extracted and calculated as the mean spectra of the whole sample.

Spectral Preprocess

Before the multivariate modeling, all the PFSP samples at different drying periods (0, 0.5, 1, 2, 3, 4, 5, 6, 7 h) were divided into two groups including the calibration set and the prediction set. Two thirds ($n = 88$) of the samples were allocated for calibration set, and the remaining one third of the samples were composed of the prediction set. As the large amount of spectral data obtained from hyperspectral images exist many different types of random noise, such as spectral variations caused by physical characteristics of samples and discrepancies in instrument response, spectral data often should be preprocessed to improve the accuracy and robustness of calibration models (Xiong et al. 2015b). Among various preprocessing methods, four different extensively used techniques including moving average, Savitzky-Golay filter (S_G filter), multiplicative scatter correction (MSC), and standard normal variate (SNV) were chosen for sample spectral normalization.

Multivariate Data Analysis

A number of multivariate regression methods including linear regression methods (such as principal component regression, MLR, PLSR, etc) and non-linear regression methods (such as genetic algorithms, support vector machines, etc) have been performed for modeling (Qin et al. 2013). In this study, PLSR, MLR, and LS-SVM were applied.

For modeling, the spectral data was used as independent (X) variable whose columns represent the reflectance value within each waveband and the rows represent the sample

number, and the anthocyanin content of PFSP was used as dependent (Y) variable, respectively. PLSR, as an effective and reliable algorithm, is particularly suitable when there is multi-collinearity among X values (Wu et al. 2012). It projects the original X variables onto a small number of latent variables (LVs) to simplify the relationship between the spectral data and reference quality matrix (Dong and Guo 2015). Leave-one-out cross-validation (LOOCV) was performed in the calibration set to decrease the model over-fitting and to confirm the proper number of LVs. And the robustness of the model was evaluated by coefficients of determination (R^2) and root mean square error (RMSE).

MLR is another multivariate chemometrics approach which is widely used to analyze the quantitative relationship between the spectral data and the observed data by fitting a linear equation (Cheng et al. 2015). However, it fails to run if the number of independent variables in hyperspectral data is more than the number of the samples, and it has the disadvantage of easy affection by collinearity between variables (Kamruzzaman et al. 2016). As the number of the samples is much less than the number of spectral wavelengths, the optimal wavelengths should firstly be selected to reduce the dimensionality, and then, the MLR could be performed to establish calibration model.

Support vector machine (SVM) is a supervised learning method that could solve the problem of non-linear regression and classification. LS-SVM is a modified algorithm of SVM. The principle of LS-SVM is that it maps input variables into a high-dimensional space and involves equality constraints instead of inequality constraints, working with a least square cost function (Yang et al. 2017). The transfer function of this algorithm includes three commonly used types: linear kernel, polynomial kernel, and radial basis function (RBF) (Yang et al. 2015a). Comparing with linear kernel and polynomial kernel functions, RBF function is capable of dealing with nonlinear relationships between the spectra and target attributes and reduces the computational complexity of training procedure (Dong and Guo 2015). Therefore, RBF was selected as the kernel function of LS-SVM. A grid search and leave-one-out cross-validation were applied to choose the optimal values of the important parameters for the model, including the regularization parameter (γ) which determines the trade-off between the training error minimization and smoothness of the estimated function, and the kernel function parameter (σ^2). All the computation and modeling process were implemented using LS-SVM toolbox (LSSVM v1.5, Suykens, Leuven, Belgium) and Matlab R2013a software (The Math Works Inc., USA).

Optimal Wavelength Selection

In this study, the spectral data acquired from hyperspectral images of PFSP samples contain 1288 wavelength bands

within the spectral range of 371–1023 nm. The extracted data possesses the characteristics of high dimensionality with redundancy and multicollinearity, which could slow down the speed of computation (Wu et al. 2012). Therefore, the elimination of wavelengths containing irrelevant information and the selection of optimal wavelengths carrying the most important information may reduce the burden of data computation and even improve the accuracy and robustness of models.

In this study, the regression coefficients (RC), also called β -coefficients, were applied to select the most informative optimal wavelengths corresponding to PLSR model with full spectral variables. The wavelengths with the large weighted regression coefficients (regardless of the sign) were considered as the optimal wavelengths for anthocyanin content prediction (Cheng et al. 2015). Then, the new optimized PLSR, MLR, and LS-SVM models were established and compared on the basis of the selected optimal wavelengths. All the analysis procedures were operated with the aid of Matlab 2013a software.

Model Assessment

There are several statistical criteria to assess the performance of established multivariate analysis models, and it is necessary to apply accurate, effective, and reliable methods. Coefficients of determination and root mean square errors of calibration (R_C^2 , RMSEC), cross-validation (R_{CV}^2 , RMSECV), and prediction (R_p^2 , RMSEP) are commonly used evaluation indicators (Cheng et al. 2013; Cheng and Sun 2014). What is more, the residual predictive deviation (RPD) also shows the prediction ability. In general, the values of R^2 as close as one and the RMSE as close as zero are considered as good predictive capacity.

Visualization of Anthocyanin Content

In order to better understand and monitor the changes of anthocyanin content of PFSP during drying process, the distribution map of anthocyanin content of PFSP was developed using obtained prediction model. Based on the quantitative relationship between spectral information in each pixel of a hyperspectral image with physicochemical attributes of samples, the concentration of anthocyanin content in this study can be calculated at each pixel in the sample, and therefore, the spatial evolution of anthocyanin content during drying process could be visualized in the distribution map. The calculated distribution maps were displayed with different colors, in which high concentration of anthocyanin content was represented by red while low concentration of anthocyanin content was represented by blue. Thus, one can easily acquire the interesting information about the distribution of anthocyanin content of PFSP during drying process, which cannot be

accomplished by applying conventional UV spectrophotometry or HPLC methods. All steps for visualization purpose were carried out using Matlab R2013a.

Results and Discussion

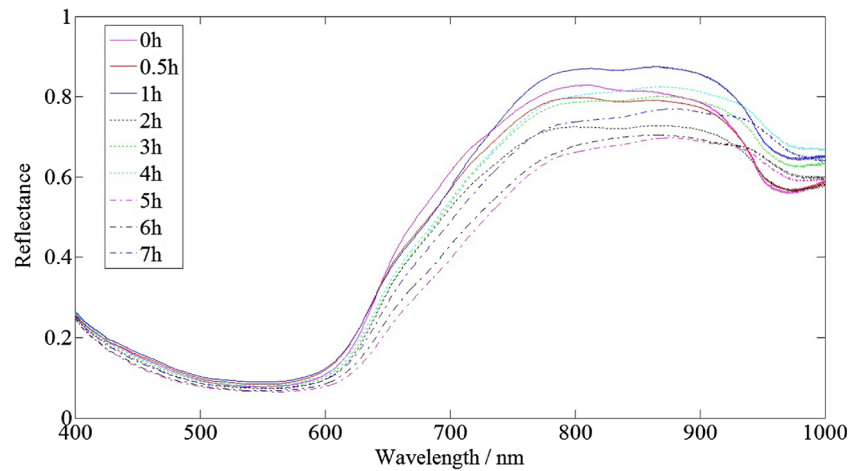
Statistics of Reference Anthocyanin Content

The statistics of the reference anthocyanin content determined by conventional ultraviolet-visible spectrophotometry method are summarized in Table 1. The wide variation that ranged from 0.253 to 3.862 (mg/g) was important to generate stable and robust prediction model. Table 1 also shows that the mean values of anthocyanin content decreased along with the increase of drying time, and especially decreased rapidly during the first several hours. These changes were consistent with the fact that the anthocyanin degraded during drying process. As a subgroup of polyphenolic flavonoids, anthocyanins are expected to be sensitive to temperature, light, oxygen, and other exogenous parameters. CUHAD, as a kind of heat-providing drying method, may cause anthocyanin oxidation and hydration reactions inside PFSP, leading to anthocyanin degradation along thermal treatments and oxygen exposure during drying process (Weber et al. 2017; Xu et al. 2014). What is more, with the increase of drying time, the microenvironment of PFSP samples changed and the freezable water content decreased which may slow down the mobility of macromolecules and then weaken the diffusion-controlled deteriorative process, and the degradation speed of anthocyanin was therefore decreased (Xin et al. 2013; Dinani et al. 2015).

Spectral Features of PFSP Samples

Due to low signal-to-noise ratios at wavelength before 400 nm and after 1000 nm, only the spectral wavelengths in the range of 400 to 1000 nm were used for further analysis. Figure 1 shows the mean relative reflectance spectral curves in the ROI of the samples at different drying periods. It could be seen that the general trends of the whole spectra for examined samples during drying were similar. The prominent valley at around 780 and 970 nm might be attributed to the O–H stretch third and second overtone (water) band (He et al. 2013). What is more, the low reflectance spectra in purple spectral region might indicate the reason why PFSP usually looks dark purple and blue (Wu et al. 2012). PFSP is an excellent source of chemical components including moisture, sugar, chlorogenic, ferulic acid, flavonoid, and anthocyanin. And the obtained spectral reflectance values are the response of these biochemistry constituents. Since some of the compositions in PFSP would change during drying process and the constituents were very complicated, reflectance curves of the samples appeared

Fig. 1 The mean spectra values of PFSP samples in wavelength range of 400–1000 nm



to vary and overlap in Vis-NIR spectral region, leading to complicated spectra (Yang et al. 2015b).

Modeling Based on Full Wavelengths

The selection of an optimal calibration model is very important for spectral analysis and makes great contribution to the subsequent prediction of anthocyanin content. The performance of diverse calibration models should be compared to select the best one. PLSR and LS-SVM were developed to establish the relationship between the spectral data extracted from the ROI of the samples and the corresponding anthocyanin content analyzed by conventional chemical method. Table 2 presents the performance of PLSR and LS-SVM with and without spectral preprocess for predicting anthocyanin content under full range spectra. As shown in Table 2, it can be seen that PLSR based on the original data obtained the accuracy with R_C^2 of 0.876, R_p^2 of 0.835, RMSEC of 0.296 mg/g, and RMSEP of 0.635 mg/g. When LS-SVM was applied to establish the calibration model, higher values of R_C^2 of 0.921, R_p^2 of 0.876, RMSEC of 0.236 mg/g, and RMSEP of 0.291 mg/g were acquired comparing to the results of PLSR, indicating that LS-SVM exhibited better performance for predicting

anthocyanin content. What is more, the *RPD* value of LS-SVM with 2.866 was higher than the *RPD* value of PLSR. In order to further explore effective ways to improve the ability of prediction model, several spectral pretreatments were applied to correct spectral data, eliminate undesirable random noise, and enhance the spectral resolution. From the view of Table 2, it was obvious that all the pretreatments were useful for the improvement of PLSR predictive model comparing to the previous result without pretreatment under the same model, and PLSR model of MSC filtering yielded the best results in accuracy with R_p^2 of 0.857 and RMSEP of 0.632 mg/g. As for LS-SVM models with diverse pretreatments, almost the whole methods led to decreased accuracy except for MSC spectral pre-process method with better ability of R_C^2 of 0.961, R_p^2 of 0.882, RMSEC of 0.166 mg/g, and RMSEP of 0.275 mg/g. The effectiveness of the other pretreatment methods decreased possibly because of the loss of useful details and information caused by spectral correction. The plots of measured and predicted anthocyanin content of PFSP samples using both PLSR and LS-SVM with MSC preprocess are shown in Fig. 2. By contrasting the accuracy and robustness of the two algorithms, the results indicated that LS-SVM was more suitable for the estimation of anthocyanin content

Table 2 Performance of PLSR and LS-SVM with and without preprocess under full range spectra

Pre-treatment methods	PLSR					LS-SVM				
	R_C^2	RMSEC (mg/g)	R_p^2	RMSEP (mg/g)	RPD	R_C^2	RMSEC (mg/g)	R_p^2	RMSEP (mg/g)	RPD
Original data	0.876	0.296	0.835	0.635	1.313	0.921	0.236	0.876	0.291	2.866
Moving average	0.875	0.298	0.839	0.631	1.322	0.887	0.283	0.834	0.335	2.490
S_G filter	0.884	0.288	0.841	0.654	1.275	0.910	0.254	0.867	0.301	2.771
MSC	0.882	0.290	0.857	0.632	1.320	0.961	0.166	0.882	0.275	3.033
SNV	0.889	0.281	0.836	0.652	1.279	0.918	0.241	0.874	0.293	2.846

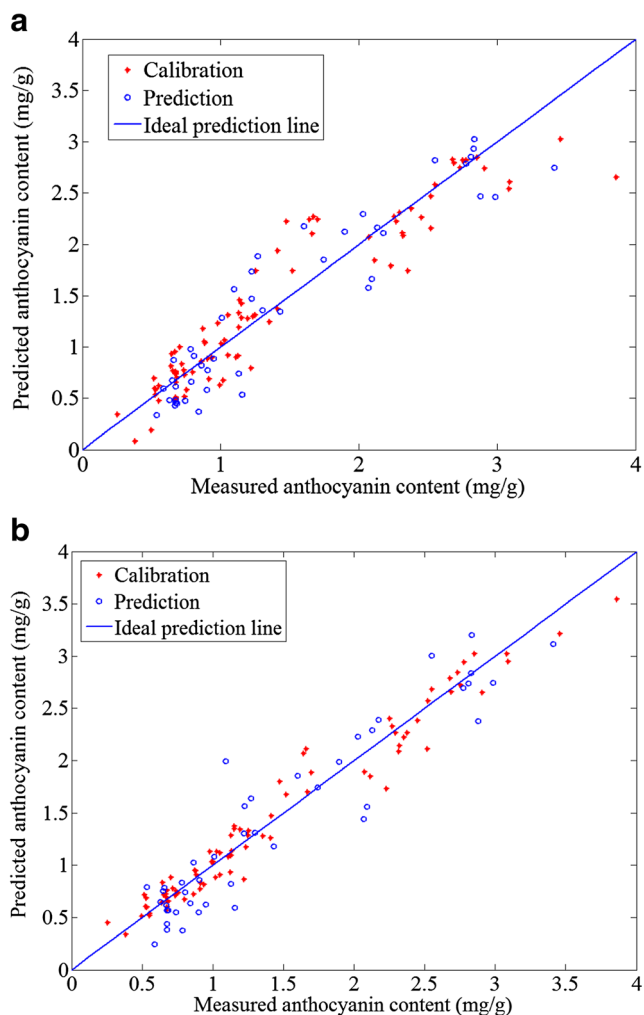
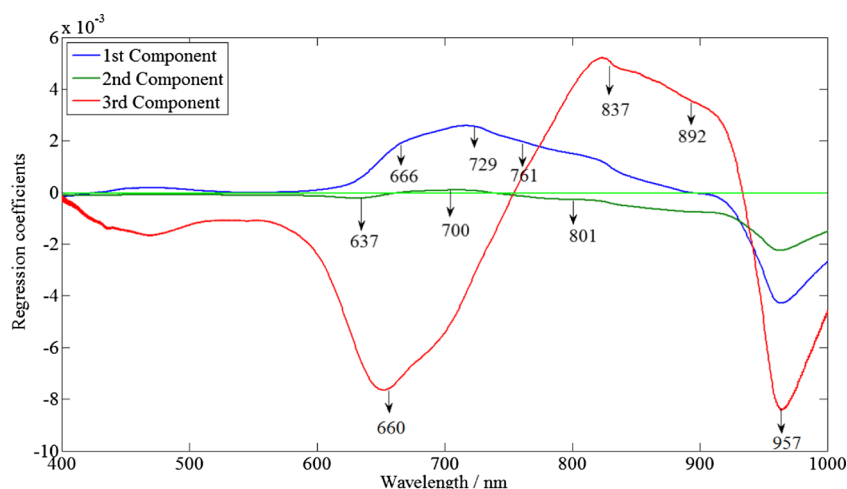


Fig. 2 Measured and predicted anthocyanin content for both PLSR model (a) and LS-SVM model (b) with MSC filtering under full range spectra

than PLSR. And according to the results of MSC preprocess for both PLSR model and LS-SVM model, it is necessary to use spectral pretreatment in terms of the current research.

Fig. 3 Selection of optimal wavelengths based on regression coefficients of PLSR calibration model



Modeling Based on Optimal Wavelengths

Hyperspectral data with hundreds and thousands of contiguous wavelengths for each pixel of image is a great issue for data processing. Such a large amount of data would cause the necessity of expensive instruments and complexity of on-line application (Cheng et al. 2016). Therefore, the selection of optimal wavelengths is very important for decreasing the burden of computation and simplifying the prediction model. In this study, RC method was applied to choose the key wavelengths containing the most valuable information related to the desired attributes from the whole spectral range (Fig. 3). As a result, ten optimal wavelengths including 637, 660, 666, 700, 729, 761, 801, 837, 892, and 957 nm were selected, and the corresponding simplified calibration and prediction models were established. In order to find a better simplified calibration model using the identified optimal wavelengths, two linear regression algorithms namely PLSR and MLR and a non-linear regression algorithm namely LS-SVM were performed to analyze the relationship between the reference measured anthocyanin content of samples and the obtained reduced spectral data.

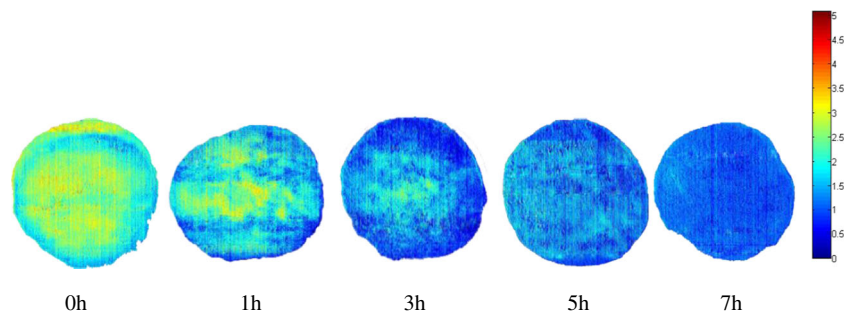
Table 3 presents the accuracy and robustness of models for the prediction of anthocyanin content during drying process based on PLSR, LS-SVM, and MLR. As shown in Table 3, the simplified RC-PLSR model performed slightly declined prediction accuracy with R_C^2 of 0.883, RMSEC of 0.288 mg/g, R_P^2 of 0.830, and RMSEP of 0.350 mg/g comparing to the prediction ability of PLSR model developed by using the whole spectral data with MSC filtering. Although the elimination of variables was approximately 99.2%, the RC-PLSR model behaved generally comparable to the PLSR model based on full wavelengths. Thus, it could be demonstrated that RC algorithm is useful and effective for the selection of key wavelengths in predicting anthocyanin content during drying. Pu and Sun (2015) also developed the successful estimation models for moisture content of mango slices during drying with RC and PLSR analysis. Furthermore, the prediction of

Table 3 Performance of PLSR, MLR, and LS-SVM models for predicting anthocyanin content during drying based on optimal wavelengths

Model	Calibration		Cross-validation		Prediction		
	R_C^2	RMSEC (mg/g)	R_{CV}^2	RMSECV (mg/g)	R_p^2	RMSEP (mg/g)	RPD
PLSR	0.883	0.288	0.874	0.296	0.830	0.350	2.382
MLR	0.868	0.300	0.859	0.315	0.866	0.302	2.763
LS-SVM	0.963	0.163	0.914	0.209	0.848	0.340	2.454

anthocyanin content throughout the ripening process of wine grape and the determination of total pigments in red meat were also reported by using RC method (Chen et al. 2015; Xiong et al. 2015b). From the view of Table 3, it also illustrated the accuracy of optimal RC-LS-SVM and RC-MLR model with R_C^2 of 0.963, 0.868 and R_p^2 of 0.848, 0.866, respectively. There were similar results between the LS-SVM algorithm and the PLSR algorithm that R_p^2 of RC-LS-SVM reduced about 3.83% compared with the LS-SVM model built based on the whole spectra with MSC filtering. The RC-MLR model behaved the most effective in estimating anthocyanin content when it was compared to the performance of RC-PLSR and RC-LS-SVM algorithms. In general, the prediction accuracy of a regression model is regarded as good when its R^2 value is higher than 0.80 (Yang et al. 2017), and the RPD of the RC-MLR was 2.763 which indicated that the model was adequate for analytical purposes. As the characteristic of easier interpretation and less colinearity than PLSR and LS-SVM, MLR is confirmed as the most suitable model for the development of a multispectral system for nondestructive measurement of anthocyanin content in PFSP slices during drying. The quantitative function was acquired and described as follows:

$$\begin{aligned}
 Y_{\text{anthocyanin}} = & 2.2304 + 13.4487X_{637\text{nm}} - 55.7173X_{660\text{nm}} \\
 & + 43.921X_{666\text{nm}} \\
 & + 12.5494X_{700\text{nm}} - 20.9731X_{729\text{nm}} \\
 & + 5.6521X_{761\text{nm}} \\
 & + 7.2235X_{801\text{nm}} - 1.4655X_{837\text{nm}} \\
 & + 5.4789X_{892\text{nm}} - 9.8640X_{957\text{nm}}
 \end{aligned} \quad (2)$$

Fig. 4 Distribution maps of anthocyanin contents of PFSP at different drying periods

where $X_{i \text{ nm}}$ is the reflectance spectral response at wavelength of $i \text{ nm}$ and $Y_{\text{anthocyanin}}$ is the predicted anthocyanin content.

Visualization of Anthocyanin Content

The visualization of distribution maps of anthocyanin content would be very useful for better understanding the change of anthocyanin within PFSP during drying process. In this study, the optimized RC-MLR model was applied to transfer the spectrum of each pixel in hyperspectral image into the predicted values, and therefore, the distribution map was formed. Figure 4 shows the distribution maps for five samples at different drying periods (0, 1, 3, 5, and 7 h, respectively). From Fig. 4, anthocyanin content within PFSP was significantly different at diverse drying time especially at the first several hours. The color in distribution map of sample dried at 0 h exhibited yellow and even some pixels appeared to be red, which indicated that there were large amounts of anthocyanin within PFSP sample. As the increase of drying time, the density of yellow color became less and tended to green, and the blue color appeared more, demonstrating the decline of anthocyanin content during drying process. Another phenomenon worthy of attention in Fig. 4 was that the anthocyanin content on the center of sample was relatively higher than that at the edge of sample. This was more obvious in the first several hours, when the blue regions started to expand to the center. This phenomenon indicated that anthocyanin content degraded more on the edge of sample than inside, which might be due to the fact that the heated condition of the edge of sample is better than that inside sample during drying process (Li et al. 2016b). When PFSP was cut into slices, the mechanically damaged tissue cells were more easily exposed to oxygen

and released endoenzymes consequently, led to the occurrence of biochemical reaction, and thus deteriorated the anthocyanin (Fратиanni et al. 2013; Liu et al. 2015). Moreover, more probability for microorganism was accessed to infect the microstructure of PFSP and therefore decompose anthocyanin (Lee et al. 2011).

It is very important to generate the distribution map of anthocyanin content of PFSP during drying based on the spatial information that hyperspectral imaging provided. The distribution map is very helpful for real-time problem detection and could improve the on-line monitor, e.g., adjusting the parameters of drying equipment such as drying temperature and ultrasonic power to shorten the drying time of the last half drying period and to avoid serious anthocyanin degradation according to the anthocyanin change regulation and degradation degree. The results indicated the advantages of hyperspectral imaging technique that could not be realized either by traditional chemical method or spectroscopy technique.

Conclusions

In this paper, a Vis/NIR hyperspectral imaging system covering the range of 371–1023 nm was investigated for the estimation of anthocyanin content in PFSP during drying process. The results showed that both PLSR and LS-SVM models based on full wavelengths with and without diverse spectral preprocessing produced different results. PLSR models with spectral pretreatments were all useful for the improvement of PLSR predictive model especially for MSC filtering which yielded best results in accuracy with R_p^2 of 0.857 and RMSEP of 0.632 mg/g. As for LS-SVM models, almost the whole pretreatment methods led to decreased accuracy except for MSC spectral pre-process method with best ability of R_p^2 of 0.882 and RMSEP of 0.275 mg/g. Then, RC was applied to select optimal wavelengths, and ten key wavelengths (637, 660, 666, 700, 729, 761, 801, 837, 892, and 957 nm) were chosen. Based on the optimal wavelengths, simplified RC-PLSR, RC-LS-SVM, and RC-MLR models were established, in which RC-MLR behaved the best performance with R_p^2 of 0.866 and RMSEP of 0.302 mg/g. Finally, the distribution maps were created based on the optimized RC-MLR model to visualize anthocyanin content change during drying process. On the whole, the encouraging results of this study indicated the potential application of hyperspectral imaging technique for fast, accurate, and non-destructive determination of anthocyanin content within PFSP during drying process.

Acknowledgments The authors express their sincere appreciation to the National Natural Science Foundation of China (No. U1404334), the College Young Teachers Development Program of Henan province (No.

2015GGJS-048), and the Science and Technology Project of Henan Province (No. 172102310617) for supporting this study financially.

Compliance with Ethical Standards

Funding This study was funded by the National Natural Science Foundation of China (No. U1404334), the College Young Teachers Development Program of Henan province (No. 2015GGJS-048), and the Science and Technology Project of Henan Province (No. 172102310617).

Conflict of Interest Yunhong Liu has no conflict of interest. Yue Sun declares that she has no conflict of interest. Anguo Xie declares that he has no conflict of interest. Huichun Yu declares that she has no conflict of interest. Yong Yin declares that he has no conflict of interest. Xin Li declares that she has no conflict of interest. Xu Duan declares that he has no conflict of interest.

Ethical Approval This article does not contain any studies with human participants or animals performed by any of the authors.

Informed Consent Not applicable.

References

- Akpinar E, Midilli A, Bicer Y (2003) Single layer drying behavior of potato slices in a convective cyclone dryer and mathematical modeling. *Energ Convers Manag* 44(10):1689–1705
- Chen S, Zhang F, Ning J, Liu X, Zhang Z, Yang S (2015) Predicting the anthocyanin content of wine grapes by NIR hyperspectral imaging. *Food Chem* 172:788–793
- Cheng JH, Sun DW (2014) Hyperspectral imaging as an effective tool for quality and safety analysis and control of fish and other seafoods: current research and potential applications. *Trends Food Sci Technol* 37(2):78–91
- Cheng JH, Dai Q, Sun DW, Zeng XA, Liu D, Pu HB (2013) Applications of non-destructive spectroscopic techniques for fish quality and safety evaluation and inspection. *Trends Food Sci Technol* 34(1): 18–31
- Cheng JH, Sun DW, Pu HB, Wang QJ, Chen YN (2015) Suitability of hyperspectral imaging for rapid evaluation of thiobarbituric acid (TBA) value in grass carp (*Ctenopharyngodon idella*) fillet. *Food Chem* 171:258–265
- Cheng JH, Sun DW, Pu H (2016) Combining the genetic algorithm and successive projection algorithm for the selection of feature wavelengths to evaluate exudative characteristics in frozen-thawed fish muscle. *Food Chem* 197:855–863
- Cipriano PDA, Ekici L, Barnes RC, Gomes C, Talcott ST (2015) Pre-heating and polyphenol oxidase inhibition impact on extraction of purple sweet potato anthocyanins. *Food Chem* 180:227–234
- Dinani ST, Hamdami N, Shahedi M, Havet M, Queveau D (2015) Influence of the electrohydrodynamic process on the properties of dried button mushroom slices: a differential scanning calorimetry (DSC) study. *Food Bioprod Process* 95:83–95
- Dincer I, Hussain MM, Sahin AZ, Yilbas BS (2002) Development of a new moisture transfer (bi-re) correlation for food drying applications. *Int J Heat Mass Transfer* 45(8):1749–1755
- Dong J, Guo W (2015) Nondestructive determination of apple internal qualities using near-infrared hyperspectral reflectance imaging. *Food Anal Methods* 8:2635–2646
- Fernandes AM, Franco C, Mendes-Ferreira A, Mendes-Faia A, Costa PLD, Melo-Pinto P (2015) Brix, pH and anthocyanin content

- determination in whole port wine grape berries by hyperspectral imaging and neural networks. *Comput Electron Agric* 115:88–96
- Fratianni A, Albanese D, Mignogna R, Cinquanta L, Panfili G, Matteo MD (2013) Degradation of carotenoids in apricot (*Prunus armeniaca* L.) during drying process. *Plant Foods Hum Nutr* 68:241–246
- Gomes V, Fernandes A, Martins-Lopes P, Pereira L, Faia AM, Melo-Pinto P (2017) Characterization of neural network generalization in the determination of pH and anthocyanin content of wine grape in new vintages and varieties. *Food Chem* 218:40–46
- He HJ, Wu D, Sun DW (2013) Non-destructive and rapid analysis of moisture distribution in farmed Atlantic salmon (*Salmo salar*) filets using visible and near-infrared hyperspectral imaging. *Innovative Food Sci Emerg Technol* 18:237–245
- Huang X, Zou X, Zhao J, Shi J, Zhang X, Mel H (2014a) Measurement of total anthocyanins content in flowering tea using near infrared spectroscopy combined with ant colony optimization models. *Food Chem* 164:536–543
- Huang M, Wang Q, Zhang M, Zhu Q (2014b) Prediction of color and moisture content for vegetable soybean during drying using hyperspectral imaging technology. *J Food Eng* 128:24–30
- Kamruzzaman M, Makino Y, Oshita S (2016) Parsimonious model development for real-time monitoring of moisture in red meat using hyperspectral imaging. *Food Chem* 196:1084–1091
- Kotíková Z, Šule M, Lachman J, Pivec V, Orsák M, Hamouz K (2016) Carotenoid profile and retention in yellow-, purple- and red-fleshed potatoes after thermal processing. *Food Chem* 197:992–1001
- Kubow S, Iskandar MM, Sabally K, Azadi B, Ekbatan SS, Kumarathasan P, Das DD, Prakash S, Burgos G, Felde TZ (2016) Biotransformation of anthocyanins from two purple-fleshed sweet potato accessions in a dynamic gastrointestinal system. *Food Chem* 192:171–177
- Lee HH, Hong SI, Kim D (2011) Microbiological and visual quality of fresh-cut cabbage as affected by packaging treatments. *Food Sci Biotechnol* 20(1):229–235
- Li J, Li XD, Zhang Y, Zheng ZD, Qu ZY, Liu M, Zhu SH, Liu S, Wang M, Qu L (2013) Identification and thermal stability of purple-fleshed sweet potato anthocyanins in aqueous solutions with various pH values and fruit juices. *Food Chem* 136:1429–1434
- Li Y, Zou X, Shen T, Shi J, Zhao J, Mel H (2016a) Determination of geographical origin and anthocyanin content of black goji berry (*Lycium ruthenicum* Murr.) using near-infrared spectroscopy and chemometrics. *Food Anal Methods*. doi:10.1007/s12161-016-0666-4
- Li Y, Xu Y, Yuan Y, Zhao Z, Yuan Y (2016b) Multi-scale and multi-layer structural heat and mass transfer modeling and experiment on drying of rice in bin. *Trans Chin Soc Agric Eng* 32(2):258–265
- Liang Z, Wu B, Fan P, Yang C, Duan W, Zheng X, Liu C, Li S (2008) Anthocyanin composition and content in grape berry skin in *Vitis* germplasm. *Food Chem* 111:837–844
- Liu Y, Wu J, Miao S, Chong C, Sun Y (2014) Effect of a modified atmosphere on drying and quality characteristics of carrots. *Food Bioprocess Technol* 7(9):2549–2559
- Liu P, Mujumdar AS, Zhang M, Jiang H (2015) Comparison of three blanching treatments on the color and anthocyanin level of the microwave-assisted spouted bed drying of purple flesh sweet potato. *Dry Technol* 33:66–71
- Liu Y, Sun Y, Yu H, Yin Y, Li X, Duan X (2017) Hot air drying of purple-fleshed sweet potato with contact ultrasound assistance. *Dry Technol* 35(5):564–576
- Martínez-Sandoval JR, Nogales-Bueno J, Rodríguez-Pulido FJ, Hernández-Hierro JM, Segovia-Quintero MA, Martínez-Rosasa ME, Heredia FJ (2015) Screening of anthocyanins in single red grapes using a non-destructive method based on the near infrared hyperspectral technology and chemometrics. *J Sci Food Agric* 96:1643–1647
- Nogales-Bueno J, Baca-Bocanegra B, Rodríguez-Pulido FJ, Heredia FJ, Hernández-Hierro JM (2015) Use of near infrared hyperspectral tools for the screening of extractable polyphenols in red grape skins. *Food Chem* 172:559–564
- Pappas CS, Takidelli C, Tsantili E, Tarantili PA, Polissiou MG (2011) Quantitative determination of anthocyanins in three sweet cherry varieties using diffuse reflectance infrared Fourier transform spectroscopy. *J Food Compos Anal* 24:17–21
- Pu YY, Sun DW (2015) Vis-NIR hyperspectral imaging in visualizing moisture distribution of mango slices during microwave-vacuum drying. *Food Chem* 188:271–278
- Qin J, Chao K, Kim MS, Lu R, Burks TF (2013) Hyperspectral and multispectral imaging for evaluating food safety and quality. *J Food Eng* 118:157–171
- Qu JH, Sun DW, Cheng JH, Pu H (2017) Mapping moisture contents in grass carp (*Ctenopharyngodon idella*) slices under different freeze drying periods by Vis-NIR hyperspectral imaging. *LWT-Food Sci Technol* 75:529–536
- Ribera-Fonseca A, Noferini M, Jorquera-Fontena E, Rombolà AD (2016) Assessment of technological maturity parameters and anthocyanins in berries of cv. Sangiovese (*Vitis vinifera* L.) by a portable vis/NIR device. *Sci Hortic-Amsterdam* 209:229–235
- Ruttarattanamongkol K, Chittrakorn S, Weerawatanakorn M, Dangpium N (2016) Effect of drying conditions on properties, pigments and antioxidant activity retentions of pretreated orange and purple-fleshed sweet potato flours. *J Food Sci Technol MYS* 53(4):1811–1822
- Steed LE, Truong VD (2008) Anthocyanin content, antioxidant activity, and selected physical properties of flowable purple-fleshed sweet potato purees. *J Food Sci* 73(5):S215–S221
- Sun Y, Liu Y, Yu H, Xie A, Li X, Yin Y, Duan X (2017) Non-destructive prediction of moisture content and freezable water content of purple-fleshed sweet potato slices during drying process using hyperspectral imaging technique. *Food Anal Methods*. doi:10.1007/s12161-016-0722-0
- Weber F, Boch K, Schieber A (2017) Influence of copigmentation on the stability of spray dried anthocyanins from blackberry. *LWT-Food Sci Technol* 75:72–77
- Wu D, Shi H, Wang S, He Y, Bao Y, Liu K (2012) Rapid prediction of moisture content of dehydration prawns using online hyperspectral imaging system. *Anal Chim Acta* 726(9):57–66
- Wu D, Wang S, Wang N, Nie P, He Y, Sun DW, Yao J (2013) Application of time series hyperspectral imaging (TS-HSI) for determining water distribution within beef and spectral kinetic analysis during dehydration. *Food Bioprocess Technol* 6:2943–2958
- Xie A, Sun DW, Xu Z, Zhu Z (2015) Rapid detection of frozen pork quality without thawing by Vis-NIR hyperspectral imaging technique. *Talanta* 139:208–215
- Xin Y, Zhang M, Adhikari B (2013) Effect of trehalose and ultrasound-assisted osmotic dehydration on the state of water and glass transition temperature of broccoli (*Brassica oleracea* L. var. *botrytis* L.) *J Food Eng* 119:640–647
- Xiong Z, Sun DW, Xie A, Han Z, Wang L (2015a) Potential of hyperspectral imaging for rapid prediction of hydroxyproline content in chicken meat. *Food Chem* 175:417–422
- Xiong Z, Sun DW, Xie A, Pu H, Han Z, Luo M (2015b) Quantitative determination of total pigments in red meats using hyperspectral imaging and multivariate analysis. *Food Chem* 178:339–345
- Xu J, Su X, Lim S, Griffin J, Carey E, Katz B, Tomich J, Smith JS, Wang W (2014) Characterisation and stability of anthocyanins in purple-fleshed sweet potato P40. *Food Chem* 186:90–96
- Yang YC, Sun DW, Pu H, Wang N, Zhu Z (2015a) Rapid detection of anthocyanin content in lychee pericarp during storage using hyperspectral imaging coupled with model fusion. *Postharvest Biol Technol* 103:55–65

- Yang YC, Sun DW, Wang NN (2015b) Rapid detection of browning levels of lychee pericarp as affected by moisture contents using hyperspectral imaging. *Comput Electron Agric* 113:203–212
- Yang Q, Sun DW, Cheng W (2017) Development of simplified models for nondestructive hyperspectral imaging monitoring of TVB-N contents in cured meat during drying process. *J Food Eng* 192:53–60
- Yoshimoto M, Yamakawa O, Tanoue H (2005) Potential chemopreventive properties and varietal difference of dietary fiber from sweet potato (*Ipomoea Batatas* L.) root. *Japan Agric Res Q* 39(1):37–43
- Zhang C, Wang Q, Liu F, He Y, Xiao Y (2017) Rapid and non-destructive measurement of spinach pigments content during storage using hyperspectral imaging with chemometrics. *Measurement* 97:149–155