RAPID AND NONDESTRUCTIVE DETECTION OF ADULTERANTS OF ORGANIC SPELT FLOUR USING HYPERSONTICAL IMAGING

Wen-Hao Su, Da-Wen Sun

1-Food Refrigeration and Computerised Food Technology (FRCFT)-School of Biosystems and Food Engineering-Agriculture & Food Science Centre-University College Dublin (UCD)-National University of Ireland-Belfield-Dublin 4-Ireland-Phone: +(353-1)7167342-Fax: +(353-1)7167493-E-mail: (dawen.sun@ucd.ie)

ABSTRACT – A multispectral real-time imaging system was proposed to determine the fidelity of organic spelt flour (OSF) from three categories of adulterants including rye flour (RF), organic wheat flour (OWF) and spelt flour (SF). A set of mutual wavelengths (1145, 1192, 1222, 1349, 1359, 1396, 1541, and 1567 nm) was chosen from the spectral range of 900 to 1700 nm by first-derivative and mean centering iteration algorithm (FMCIA) for all investigated flour samples. Then these selected feature wavelengths were utilized in partial least squares regression (PLSR) and multiple linear regression (MLR) models to devise multispectral imaging system. Better performances for quantitative measure of adulterants were emerged in simplified PLSR models. To visualize the adulterants in OSF samples, the distribution maps were drawn by computing the spectral response of each pixel on corresponding spectral images at specific frequencies using a quantitative identification function.

KEYWORDS: spectral imaging; adulteration; authentication; chemometric analysis; visualization

1. INTRODUCTION

The hexaploid spelt (Triticum spelta L.) was the predominant cereal food cultivated in Europe from the 5th century and has been substituted by wheat (Triticum aestivum L.) since the 20th century. The spelt yield is much lower, because the husks take up loss of about 30% and the milling procedure requires an extra step for husk separation. With an excellent source of dietary fiber, vitamins and minerals besides carbohydrate and protein, spelt is suggested to have higher nutritional potential and better taste than common wheat, and can be widely used for brewing, baking and production of pasta. In recent years, spelt whose price often achieves twice that of comparable wheat products is undergoing a renaissance in Europe and North America. The advantage of spelt is that it is more resistant to adverse climate and poor soil conditions like wet, cold soils and at high altitudes. Moreover, spelt carries high-level resistance to several fungal pathogens in terms of Pythium aristoporum Vanterpool, Fusarium spp and yellow rust. Therefore, spelt can survive in low-input farming systems without using much fertilizer and pesticide, which makes it more suitable for organic agriculture. It is known that the aim of organic agriculture is to minimise the use of external inputs and avoid using any chemical additive. In organic agriculture, the lower yield and cumbersome qualification process to some extent result in much higher prices of organic spelt products. Even so, the global market for organic food is still huge and increasing significantly, especially in Europe where the largest organic food market exists.

Hyperspectral imaging has been applied not only to the determination of their locations based spatial distribution but also the authentication and quantification of spectral properties based chemical component. In order to realize the on-line application, high-efficiency and cost-effective multispectral imaging systems can be eventually established using qualified algorithms and feature wavelengths based
on hyperspectral imaging. Thus, it is essential to identify the most characteristic wavelengths from hundreds of contiguous spectral bands to develop effective models. The criteria for the selection of wavelength is based on the information content of spectral bands to enhance the prediction accuracy of the calibration model maximally. In this study, we present an approach termed first-derivative and mean centering iteration algorithm (FMCIA), which has the latent capacity to select an optimum wavelength combination in full spectral range. In order to demonstrate the effectiveness of the proposed algorithm, partial least squares regression (PLSR) and multiple linear regression (MLR) models are constructed to evaluate the inveracity of RF, OWF and SF in OSF, respectively. The accuracy and predictability of these simplified models are verified and compared. To visualize the result, spatial distribution maps of adulteration are generated by image processing.

2. MATERIALS AND METHODS

2.1. Sample Preparation and Spectral Image Acquisition

The organic spelt and organic common wheat were produced from organic system where the organic food standards such as non-using synthetic fertiliser and pesticide were strictly implemented. The whole OSF samples (Doves Farm Foods Ltd., Berkshire, UK) micronized by grinding were legally qualified organic (EU Organic - GB-ORG-05, EU/non-EU Agriculture). The OWF samples were certified by the Organic Trust in Ireland (IE-ORG-03, EU/non EU Agriculture). Other whole flour samples including RF and SF were produced in conventional system without following organic standards. All these samples were collected and transported to the laboratories. Firstly, a total of 60 pure samples of each flour variety were prepared with a weight of 35 g for each sample and all the 240 (60 × 4) samples used as the calibration set. Then OSF samples were adulterated by respectively mixing with RF, OWF and SF in the range of 3-75% (w/w), at approximately 3% increments. Specifically, the RF, OWF and SF samples were separately weighed together with OSF samples, thoroughly mixed and homogenized to get a total sample weight of 35 g each time. Finally, 150 samples (6 samples per adulterant level × 25 levels) were obtained for each adulteration type (RF, OWF and SF adulterated with OSF). Among them, 90 samples (4 samples per adulterant level × 25 levels) were randomly selected as the calibration set and the remaining 60 samples (2 samples per adulterant level × 25 levels) were selected as the prediction set. All these acquired samples were respectively placed in circular transparent plastic jars one by one and imaged using the laboratory-based pushbroom hyperspectral imaging system mentioned by Kamruzzaman, ElMasry, Sun, and Allen (2012).

2.2. Extraction of the Region of Interest

The fundamental purpose of extraction of spectral data from the test samples is to identify the region of interest (ROI) and obtain the average spectrum representing the ROI. The raw hyperspectral images should be first calibrated and transformed into relative-reflectance images based on the captured white (~99% reflectance) and dark reference images dark (0% reflectance) (Menesatti et al., 2009; Wu & Sun, 2013a). In the relative-reflectance image, as there are incomputable mussy fringes and noises from 897-957 nm and 1625-1753 nm, only the spectral image from 957 to 1625 nm (200 bands) are distinct and useful. In order to remove the background from the corrected hyperspectral image, a binary mask image was constructed by subtracting a low-reflectance band image from a high-reflectance band image within the same corrected hyperspectral image. The segregated flour part was then treated as the ROI to be used for extracting spectral data from the sample. A mean spectrum was then generated by averaging the spectral reflectance values of all pixels within the ROI at each wavelength of 957 to 1625 nm. The same procedure was repeated to obtain mean spectrum for hyperspectral image of every tested sample.
2.3. Multivariate Data Analysis

The quantitative analysis of OSF adulteration was explored by PLSR and MLR. These models were developed through a specific algorithm utilizing the spectra in the spectral data matrix (input variables, X-block) and correlated values in column vector (Y variables). To acquire the best prediction ability, X pre-processing techniques such as first derivative (1st Der) (Savitsky Golay smoothing, 7 points window, 1 order polynomial) and mean centring (MC), and Y pretreatment with MC were applied. In this study, the calibration models were evaluated based on a full cross validation of venetian blind protocol with 10 data splits and a test set validation. A calibration phase and a validation phase calculating for the determination coefficients and the root mean square error in calibration ($R^2_C$, RMSEC), in cross validation ($R^2_{CV}$, RMSECV) and in prediction ($R^2_P$, RMSEP) are employed in the models for presenting model performance. The superiority of spectral imaging exists in transferring multivariate analysis models to each pixel of the image, generating a visualization map with adulteration proportions. All the multivariate data and image analysis were executed by house written scripts and PLS toolbox 7.9.5 (Eigenvector Research, Manson, WA, USA) within Matlab 7.12 software (The Mathworks Inc., Natick, MA, USA).

3. RESULTS AND DISCUSSION
3.1. Spectral Characteristics of Flour Samples

The average spectra of each adulteration level from 3% to 75% including rye flour adulteration (RFA), organic wheat flour adulteration (OWFA) and spelt flour adulteration (SFA) are depicted in Figure 1(a, b and c), respectively. The analogous spectral trend throughout the wavelength range was noticed, along with the variances in the magnitude of spectral absorption for different adulteration proportions. The comparison of the mean spectra of RFA, OWFA, and SFA is demonstrated in Figure 1(d). The average spectra curve of SFA had the lowest reflectance, followed by OWF and RFA. Differences observed in spectral profiles among different varieties may be related to variations in structure properties of the tested samples. In addition, the variations in spectral features among such samples may be attributed to the differences in their chemical constituents in terms of carbohydrates, protein, microelement and other nitrogen-containing compound.

Figure 1 - The spectra of different flour adulteration.
3.2. Selection of Optimum Wavelengths

To improve the processing speed and realize on-line monitoring, it is of vital importance to qualify several optimal wavelengths instead of hundreds of variables. Wavelength selection can also improve the model prediction ability and furnish more robust models as well as models that can be transferred more readily. In this study, FMCIA was put forward for selection of characteristic variable in the spectral range of 957 to 1625 nm. The original mean and StdDev spectra of all the samples from different flour varieties are described in Figure 2 (a). The first derivative (1st Der) spectra and the spectra processed by 1st Der and mean centring (MC) were shown in Figure 2 (b and c). Based on this procedure, it will be helpful to establish a model with higher precision and to determine feature wavelengths for multispectral imaging system. As can be seen in Figure 2 (d), the loading plot of StdDev coefficient resulting from FMCIA was used for choosing characteristic wavelengths. The variables corresponding to the peak and trough of coefficient presented higher differences and would play an important role in established models. By means of this approach, eight sensitive wavelengths (1145, 1192, 1222, 1349, 1359, 1396, 1541, and 1567 nm) were finally survived.

Figure 2 - Selection of sensitive wavelengths based on FMCIA.

3.3. MLR and PLSR Models at Optimum Wavelengths

For purpose of verifying the validity of selected characteristic wavelengths and rapid quantitative detection of adulterants in OSF, MLR and PLSR models at feature wavelengths were raised and their predictive accuracies were compared. Both MLR and PLSR models using only feature wavelengths showed very good performance, and the property of models pretreated by MC was much better than models without pre-processing. Moreover, it revealed that both FMCIA-MC-MLR and FMCIA-MC-PLSR models possessed nearly the same prediction accuracy. However, compared with FMCIA-MC-MLR model, the predictive ability of the FMCIA-MC-PLSR model was a little bit better with \( R^2_p \) of 0.933 for detecting SFA in OSF. In addition, it was realized that the performance of the FMCIA-MC-PLSR model was almost comparable to that of the model developed with full spectra.
Therefore, the FMCIA-MC-PLSR model was considered the best model for measuring OSF adulterants. With the qualification of seven LVs, the OWFA in OSF was detected to obtain highest accuracy ($R^2_P = 0.974$) followed by RFA ($R^2_P = 0.965$) and SFA ($R^2_P = 0.933$). The FMCIA-MC-PLSR model (mean $R^2_P = 0.957$) could be utilized for designing a high-efficiency multispectral system for on-line inspection.

3.4. Visual Detection

As an advanced spectroscopy and imaging technique, hyperspectral imaging can not only be used for the identification and quantification of adulterants in the spectral dimension but also has the potential to realize visualization of adulteration in the spatial dimension, which means that based on their spectral characteristics this technique has superiority to recognize the varieties, gradients and spatial distributions of specific samples by spatial visualization of the image in each pixel. Figure 3 showed five typical distribution maps of adulteration percentages of SF in OSF by predicting variations in adulteration proportion based on FMCIA-MC-PLSR model. With the increase of adulteration level from 3% to 75%, the overall image colors changed from blue (low level) to red (high level), which indicated the corresponding adulteration in proportion to spectral differences of pixels. It was found that the adulteration degree varied in disparate samples and even within different locations of one same sample. These visual discrimination can be finally carried out by hyperspectral imaging, but it is hard to realize using naked eyes. Based on precise models recognized, the results were graphic to show the potential of hyperspectral imaging for non-invasive and rapid detection of adulteration in OSF.

Figure 3 - Prediction maps of adulteration percentages of SF in OSF based on FMCIA-MC-PLSR model.

4. CONCLUSIONS

The most vital challenge for this current study is to devise a new wavelength selection approach for real-time determination of three categories of adulterants in OSF using hyperspectral imaging. The results demonstrate that the FMCIA is an effective method of selecting characteristic wavelengths. The efficient multispectral imaging system can be established to detect OSF adulteration for a real-time practical application based on hyperspectral imaging in tandem with multivariate analyses. Instead of qualifying specific sets of characteristic wavelengths for different flours, mere eight wavelengths (1145, 1192, 1222, 1349, 1359, 1396, 1541, and 1567 nm) were chosen for all flour samples. These wavelengths were verified in different simplified models for designing an accurate and rapid multispectral system. Based on the feature wavelengths, the optimum model performances with the visualization of adulteration were found in FMCIA-MC-PLSR model. Moreover, even though the adulterants (RF, OWF and SF) investigated in this study are just three categories, it is possible to detect the presence of other unknown adulterated flour categories or exotic matters based on the methodology used. Additionally, the proportions of adulterated OSF samples from 3% to 75% are un-continuously
with the interval of 3%, which means that the interferent percentages in OSF might affect the prediction accuracy. In future study, the impacts of these potential factors will be explored to obtain more precise and reliable models.

ACKNOWLEDGMENTS

The authors would like to acknowledge the UCD-CSC Scholarship Scheme supported by University College Dublin (UCD) and China Scholarship Council (CSC).

REFERENCES

Zhang, Q., Saleh, A. S., & Shen, Q. (2013). Discrimination of edible vegetable oil adulteration with used frying oil by low field nuclear magnetic resonance. Food and Bioprocess Technology, 6(9), 2562-2570.