INTELLIGENT AUTOMATION & SOFT COMPUTING, 2017
https://doi.org/10.1080/10798587.2017.1293927

Check for updates

Hyperspectral Reflectance Imaging for Detecting Typical Defects of Durum Kernel Surface

Feng-Nong Chen[a,b#], Pu-Lan Chen[c#], Kai Fan[a] and Fang Cheng[d]

[a]College of Life Information Science and Instrument Engineering, Hangzhou Dianzi University, Hangzhou, People's Republic of China; [b]School of Technology and Health, KTH Royal Institute of Technology, Stockholm, Sweden; [c]Department of Economic Management, Cancer Affiliated Hospital of Xinjiang Medical University, Urumqi, China; [d]College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou, People's Republic of China

ABSTRACT

In recent years, foodstuff quality has triggered tremendous interest and attention in our society as a series of food safety problems. The hyperspectral imaging techniques have been widely applied for foodstuff quality. In this study, we were undertaken to explore the possibility of unsound kernel detecting (Triticum durum Desf), which were defined as black germ kernels, moldy kernels and broken kernels, by selecting the best band in hyperspectral imaging system. The system possessed a wavelength in the range of 400 to 1,000 nm with neighboring bands 2.73 nm apart, acquiring images of bulk wheat samples from different wheat varieties. A series of technologies of hyperspectral imaging processing and spectral analysis were used to separate unsound kernels from sound kernels, including the Principal Component Analysis (PCA), the band ratio, the band difference and the best band. According to the selected bands, the best accuracy was 95.6, 96.7 and 98.5% for 710 black germ kernels, 627 break kernels and 1,169 healthy kernels , respectively. The result shows that the method based on the band selection was feasible.

Abbreviations: CCD: Charge-coupled Device; PC: Personal Computer; PCA: Principal Component Analysis; PLSDA: Partial Least Lquares Discriminant Analysis; ANN: Artificial Neural Networks; SVM: Support Vector Machine

KEYWORDS

Foodstuff quality; Hyperspectral imaging; Unsound kernel detection; Effective band selection

I. Introduction

Foodstuff safety is now an extremely important issue that concerns everyone in the world. The unsound kernel would affect the baking quality of bread, thus lowers the premium paid for wheat. It is time consuming to detect the kernel one by one. So, the quality assessment of bulk materials becomes more and more pressing. The appropriate information about the quality of products will be contributed to control the bulk handling process or to subdivide products into different quality classes. The sales contract often specifies the allowable percentage of unsound kernels by weight. The amount of unsound kernels is an important quality parameter in grain products. And if the amount is big, that will weaken the quality of picked products and lead economic losses for the proprietor. There is also a health concern due to the possible concomitant production of mycotoxin. Unsound kernels, along with other damages, will lower the official grade (Delwiche & Kim, 2000).

Cereal researchers have been examining the kernel morphology with digital image analysis more than 30 years since 1985 (Zayas reported on the use of image analysis to distinguish between morphological characteristics of the soft red winter wheat kernel for the first time). In many respects, this vision-based method imitated the actions of the trained inspector to examine size, shape, texture, and color of the kernel (Delwiche, Yang, & Graybosch, 2013). Most early studies have exploited the size and shape characteristics with 2D projections of single kernels, whereupon multivariate analysis has been applied to get the geometrical parameters. Majumdar et al. (Majumdar & Jayas, 2000a, 2000b, 2000c, 2000d) have developed Classification models by combining two or three feature sets (morphological, colour, textural) to classify individual kernels, the research chose different features to test on independent data-sets. Vitreous kernels are mostly related to quality, machine vision system can determine the percentage of vitreous, starchy, piebald and shrunken kernels, using a trans-illuminated image of one layer of non-singulated kernels (in bulk) acquired by a digital camera (Venora, Grillo, & Saccone, 2009). Wavelet analysis is a popular tool for characterization and classification of image texture (Choudhary, Mahesh, Paliwal, & Jayas, 2009; Choudhary, Paliwal, & Jayas, 2008; Qing, Tao, et al., 2016), and it is a signal processing technique for multi-resolution image texture analysis performed by decomposing the images into multiple wavelet components using a filter bank as suggested.

If using PCA to reduce the huge data of features (Wiwart, Suchowilska, Lajszner, & Graban, 2012), the applied method may facilitate the identification of hybrids between common wheat and other Triticum species, and the selection process in creative breeding. Non-destructive quality evaluation of agricultural products has become a major area of interest for the agricultural processing industry. Researchers have been working to find techniques for evaluating internal quality attributes of agricultural and food products non-destructively.

CONTACT Fang Cheng fcheng@zju.edu..cn; Feng-Nong Chen fnchen@hdu.edu.cn
#Feng-Nong Chen and Pulan Chen are contributed equally to this study and share first authorship.

© 2017 TSI® Press

The availability of advance technology has expanded avenues for nondestructive food quality determination, such as chemical composition, mineral content and fatty acid of seeds (Gharibzahedi, Mousavi, Jouki, & Ghanderijani, 2012). X-ray and computed tomography imaging techniques are two of them, which are gaining popularity nowadays in various fields of agriculture and food quality evaluation (Kotwaliwale et al., 2014). However, the error rate was high (Gorretta et al., 2006; Kwon & Nasrabadi, 2007; Wallays, Missotten, De Baerdemaeker, & Saeys, 2009).

Current hyperspectral imaging technology cannot be directly implemented in the online system for agricultural products sorting, because of the extensive time, which is caused by image acquisition and subsequent analysis of the huge data. However, hyperspectral imaging will be the most competitive candidates in our researches for it can not only determine important spectral bands, but also be implemented in a multispectral imaging system. These spectral bands can be obtained with different analysis methods, such as Principal Component Analysis (Maganioti et al., 2010). Ridgway et al. used the NIR spectra of healthy and insect-damaged wheat kernels to select the most important wavelengths to acquire NIR images, then according to the differences of the compared pictures to detect insect damage (Ridgway, Chambers, & Cowe, 1999). Singh et al. assessed the potential of shortwave NIR hyperspectral imaging for insect damage detection in wheat and made a comparison about the performance with colour imaging (Singh, Jayas, Paliwal, & White, 2010). Del detected toxigenic fungi among different hybrids of maize kernels, determined the damage degree using a tabletop reflectance hyperspectral imaging system and discriminated maize kernels of fungal contamination from healthy kernels with the method that combine hyperspectral imaging with multivariate statistical analysis of the data (Del Fiore et al., 2010). Nakariyakul et al. discriminated internally hyperspectral transmission spectra of damaged almond nuts from normal ones by only two sets of ratio features. However, most of them just analyze the raw data without further data analysis (Nakariyakul & Casasent, 2011).

The objective of the research was to detect the unsound kernels of wheat through studying the potential of hyperspectral imaging in the spectral region ranging from 400 to 1,000 nm. Specific objectives were as follows:

- To identify the spectral region and/or wavelengths, which are the most useful for the detection of unsound kernels.
- To develop the corresponding computer algorithms to identify and segregate unsound kernels from the normal ones.
- To verify the efficiency of the most useful band.

II. Materials and Methods

A. Grain Samples

The grain samples used in this study were *Xinchun* 8# and *Zhongmai* 998# harvested in 2009. These samples were obtained from several seed distributors in Xinjiang Uygur Autonomous Region and Henan province. Each group was a composite samples derived from many farms. Wheat category samples with different locations were mixed to obtain composite category samples. These samples were stored in closed plastic bags in a refrigerator to keep the material fresh and retain the initial moisture content. The moisture content of the samples was not measured, but it varied considerably, because the samples were collected at different stages of the season and different times of the day. To prepare the samples for hyperspectral measurements, unsound kernels were separated from the normal kernels. However, there were insufficient quantities available from unclean commercial samples; with the permission of the experimentation, these samples were handmade of kernels from normal samples in the laboratory (Mohammad, Emam-Djomeh, & Khazaei, 2012). Unsound kernels were visually selected and stored at the same temperature, but in relative different humidity conditions prior to imaging. To prepare the samples for image analysis, those kernel samples were selected from the bulk samples on the basis of visual judgment according to National standard of GB/T 21124-2007 and GB 1351-2008 (Chinese National Standard). Moreover, 710 black germ kernels, 627 break kernels and 1,169 healthy kernels were used to verify the efficiency of the most useful band.

B. Hyperspectral Imaging System

Figure 1 is the schematic of the hyper-spectral imaging system. It consist of a charge coupled device (CCD) camera system, Spectra Video Camera from HAMAMATSU, Inc. (Hamamatsu C8484–05G, Japan), equipped with an imaging spectrograph ImSpector V10E from Spectral Imaging Ltd. (Oulu, Finland). The ImSpector had a fixed-size internal slit to define the field of view for the spatial line and a prism/grating/prism system for the separation of the spectra along with the spatial line. To improve the spatial resolution of the hyperspectral images, an external adjustable slit was placed between the sample and the camera optical set. The image acquisition and recording were performed with a Pentium-based PC using a general purpose imaging software, SpectralCube 2.75b from Spectral Imaging, Inc. (Spectral Imaging Ltd., Finland.). A C-mount set with a focus lens and an aperture diaphragm allows focusing and aperture adjustments, for which the circular aperture was opened to its maximum and the external slit was adjusted with micrometer actuators to optimize light flow and resolution. The light source consisted of two 21 V, 150 W halogen lamps powered with a regulated DC voltage power supply from SCHOTT

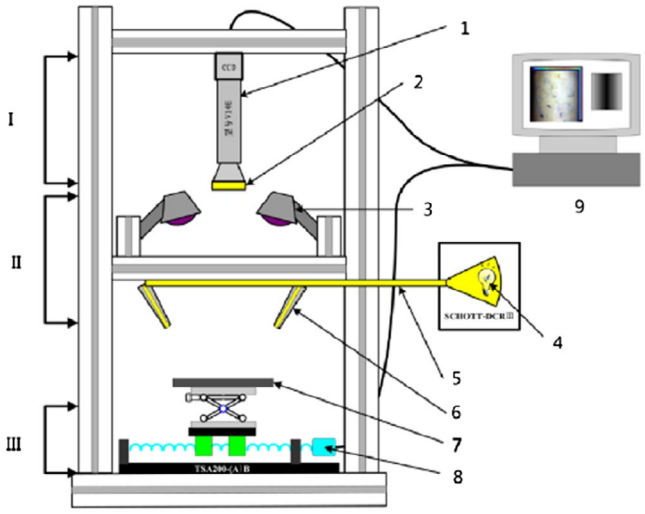

Figure 1. Schematic of the Hyper-spectral Imaging System (I) Optical Sensor Module; (II) Light Source Module; (III) Sampling Module (1) CCD; (2) Lens; (3) UV Light; (4) Visible and Near-infrared Light; (5) Extension Cable; (6) Line Source; (7) Mobile Platform; (8) Stepper Motor; (9) PC.

North America, Inc. (SCHOTT, USA). The light was remote controllable using built-in integrated RS232 interface. The samples were placed on a controllable move platform (Zolix, TSA200B, China) with an adjustable speed AC motor control speed master from Beijing Zolix Instrument Co., Ltd. (Zolix, SC300,China). The samples were scanned line by line with an adjustable scanning rate, illuminated by the two line sources as it passed through the camera's field of view.

C. Spectral Calibration

Dark current measurements were made with a lens cap covering the camera lens and turning off the lamps. The measurement data, based on the reflection principle, was collected from a slab of Spectral and placed in the same location as the wheat samples moving. The operation sequences consist of the acquisitions of dark current, the reflectance from a stationary reference target and reflectance from the sample. To correct inherent spatial non-uniformities due to the light intensity on the target area, the sample reflectance ($R_{i\lambda}$) at an individual pixel (i) and a given wavelength (λ) was calculated as follows:

$$R_{i\lambda} = RC_{\lambda}\frac{I_{i\lambda} - DC_{i\lambda}}{REF_{i\lambda} - DC_{i\lambda}} \quad (1)$$

where I represents a sample intensity, DC is the dark current, REF is the reference intensity, and RC is the correction factor for the reference slab. A value of RC = 1.0 at all wavelengths was used for the matter of simplification, due to Spectralon's flat response across the wavelength region and its high absolute reflectance (0.99).

D. Image Acquisition Method

For each wheat sample, normal kernels and unsound kernels were selected through the appearances. Kernels were placed on a black cloth (if possible, chose unreflecting black background) in alternating rows with unsound and normal kernels. The gap between neighboring rows was random.

The spectral line acquisition was performed at a room with the temperature of 20 °C. The exposure and acquisition time was set 15 ms for reflectance. The size of the hyperspectral image cube was 1,344 (length) ×600 (width) ×1024 (height). The proper speed of platform was 1.1 mm/s. After dark current correction at each pixel, monochromatic images were converted into percent reflectance using the commercial software package ENVI 5.0 from Research Systems, Inc. (Boulder, CO, USA).

E. Analysis Methods

MATLAB (The Math Works Inc., Natick, MA, USA) programs including Image Processing and Statistics Toolbox together with SPSS and HALCON were used for hyperspectral images processing and analyzing, such as, image segmentation, image classification and data analysis.

a. Feature Extraction

There are several methods for features extraction. With intuitive properties and simplicity of implementation, image thresholding enjoys a central position in applications of image segmentation. In this study, we use thresholding and morphological operation to extract the contour of kernels. According to thresholding, we extracted the object part, thin gulfs and small holes were gotten rid of using morphological operation.

b. Data Analysis

Data analysis is a fundamental task due to the great quantity of analytical information. Supervised pattern recognition aims to establish a classification model based on the experimental data, which was used to assign unknown samples to a previously defined sample class that based on its pattern of measured features. In this study, several kinds of pattern recognition methods have been applied to data analysis. They are Partial Least Lquares Discriminant Analysis (PLS-DA), Principal Component Analysis—Artificial Neural Networks (PCA-ANN) and Support Vector Machine Discriminant Nnalysis (SVM-DA).The detail you would see at the next section.

Partial Least Squares modeling is a multivariate projection method for modeling a relationship between dependent variables Y and independent variables X (Flood, Connolly, et al., 2016). The principle of PLS is used to find the components in the input matrix X that describe the relevant variations as much as possible and has a great correlation with the target value Y, then gives less weight to the variations that are irrelevant. Therefore, PLS models both X and Y simultaneously to find the latent variables (LVs) in X that will predict the latent variables in Y. PLS maximizes the co-variance between matrices X and Y.

SVM is a supervised learning technique, based on the statistical learning theory (Cortes & Vapnik, 1995). It is applicable to cope with both classification and regression problems. In the case of classification, SVM is a method for obtaining the "optimal" boundary of two classes in a vector space independently on the probabilistic distributions of training vectors in the data-set. If the two classes are linearly separated, the aim of SVM is to find the "optimal" hyperplane boundary, which exactly separates them, classifying not only the training set but also unknown samples. The "optimal" boundary is defined as a hyperplane has the maximum distance from both sets. Although the distribution of the sets is not known, this boundary is expected to be the optimal classification of the sets, since this boundary is the most isolated one from both sets. The training vectors closest to the boundary are called support vectors. The margin is the minimal distance between the separating hyperplane and the closest data points. So, SVM learning machine seeks for an optimal separating hyperplane where the margin is maximal.

In artificial neural network (Hopfield, 1982), the most versatile and widely used type of networks is the back-propagation ANNs(BPNN). The term back-propagation refers to the way the error computed at the output side is propagated backward from the output layer to the hidden layer and finally to the input layer. In back-propagation ANNs, the data is fed forward into the network without feedback, and the neurons can be fully or partially interconnected. The learning strategy of back-propagation is based on an algorithm that corrects the weights within each layer in proportion to the error obtained from the previous layer.

III. Results

Figure 2 shows the mean spectra reflectance value of normal and unsound kernel tissue in the range from 400 to1,000 nm.

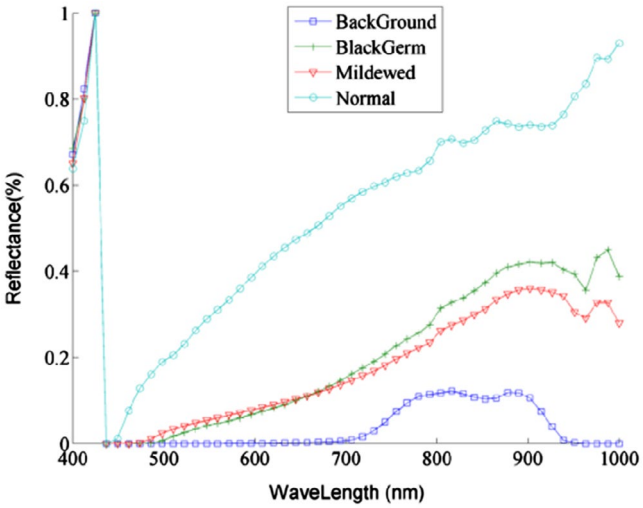

Figure 2. Representative ROIs Averaged Reflectance Spectra of Normal and Unsound Kernel.

The mean spectra were calculated based on pixels of the region of interest (ROIs) from all wheat samples. Here, we selected 4 pixels as the region of interest (ROI) with a squire containing 2*2 and calculated its average value, because 4 pixels are enough to get the useful information, otherwise it will take up more time to calculate. The mean reflectance of unsound kernels' tissue was consistently lower than that of normal kernels' tissue except the range from 400 to 500 nm, where considerable spectral overlapping was observed from four types of tissue. Moreover, generally not all ranges of high spectral were necessary for the purpose of discrimination. In this study, no image appeared below the band of 452 nm. The biggest difference of the reflectance value between normal and unsound kernels was ranging from 930 to 1000 nm, and further analyses to find greater difference does not necessarily yield for higher detecting accuracies. What's more, this range provided a low signal-to-noise ratio. So the spectral range of our research was selected from 452 to 930 nm.

A. Principal Component Analysis on Hyperspectral Cube Data

In order to minimize the hyperspectral images' data to be processed, it was desirable to find an optimal bandwidth without sacrificing detection results at the same time. PCA on hyperspectral images was performed to reduce spectral dimensionality and enhance image features. An optimal threshold value was chosen to segregate unsound kernels from the normal ones based on the principal component images (Ariana, Lu, & Guyer, 2006).

To get PCA images, all of selected image bands are as input data, and then the output image bands are ranked by eigenvalue, which is the result based on PCA algorithm. The first principal component (Figure 3a) covered 96.2% of the image variation across the wavebands, but no clear features appeared. Only specific damage and mildewed were visible in the region, black germs were difficult to distinguish. Compared with other principal components (Figure 3), the second PC(Figure 3b) enhanced the contrast between the normal and abnormal portions of the wheat. And it was more apparent between black germs and mildewed. It was also easy to extract abnormal features from PC2. The second principal component alone was sufficient to distinguish the unsound kernels from the normal ones so that it was chosen. The corresponding spectral contributions for the features were reported in Figure 4.

Although it was easy to extract features from the unsound kernel, the computation of the second principal component consists of all spectral images for a spectral region. For real time applications, it was more desirable to use fewer (two or three) wavelengths in order to accelerate the speed of image acquisition and analysis.

B. Difference and Ratio of Two Wavelengths

Followed by image segmentations, difference or ratio of two wavelengths was used in this study. According to the band contributor of PC2, two extreme wavelengths were 617 and 925 nm, certainly, there is a fluctuating range at this extreme wavelength, but the range should be evaluated.

For the difference of two wavelengths, we tried to calculate by Equation (2) to extrude unsound kernel features:

$$Rt = R_{925\ \mathrm{nm}} - R_{617\ \mathrm{nm}} - Th_{617\ \mathrm{nm}} \tag{2}$$

where $R_{925\ \mathrm{nm}}$ and $R_{617\ \mathrm{nm}}$ are relative reflectance value at 925 and 617 nm, respectively. $Th_{617\ \mathrm{nm}}$ was the result of gray-level image segmentation with the method of optimal global and adaptive thresholding.

The ratio between two wavelengths not only decreased the effect of light source, but also increased the difference between different bands, as calculated by Equation (3)

$$BV_{i,j,r} = BV_{i,j,k}/BV_{i,j,l} \tag{3}$$

where $BV_{i,j,k}$, $BV_{i,j,l}$ were the relative reflectance's of the same pixel at k and l band, respectively. The ratio algorithm for two extreme wavelengths was found by using correlation analysis among all possible wavelengths. According to the correlation analysis of all wavelengths between 452 nm to 930 nm, wavelengths 453 nm and 902 nm were found to own the minimum correlation coefficient. After performance by Eq. (3), the result showed the normal aspect of kernel, including sound kernels and normal aspects of unsound kernels. So it should be calculated by another algorithm to get the unsound features. Compared with the difference of two wavelengths, the ratio algorithm was unfit.

C. Unsound Kernel Detection Based on Extreme Band

The high-level band obtained from above indicators analysis was 617 nm, which was a more effective wavelength to extract unsound kernel features than other wavelengths. Here, we can extract the abnormal area from the kernel by thresholding. First of all, it was easy to find the thresholding (T1) to extract wheat region from the background and the thresholding (T2) to extract wheat normal regions, but, the area of normal wheat was different with the difference of thresholding, because the pixel value was gradient between the normal region and the background region. To solve the problem, erosion, the simplest solutions, was chosen to shrink the large area by proper structuring element; those algorithms can be realized by mathematical morphology. At last, the unsound kernel features can be extracted by performing an exclusive OR operation on the corresponding image.

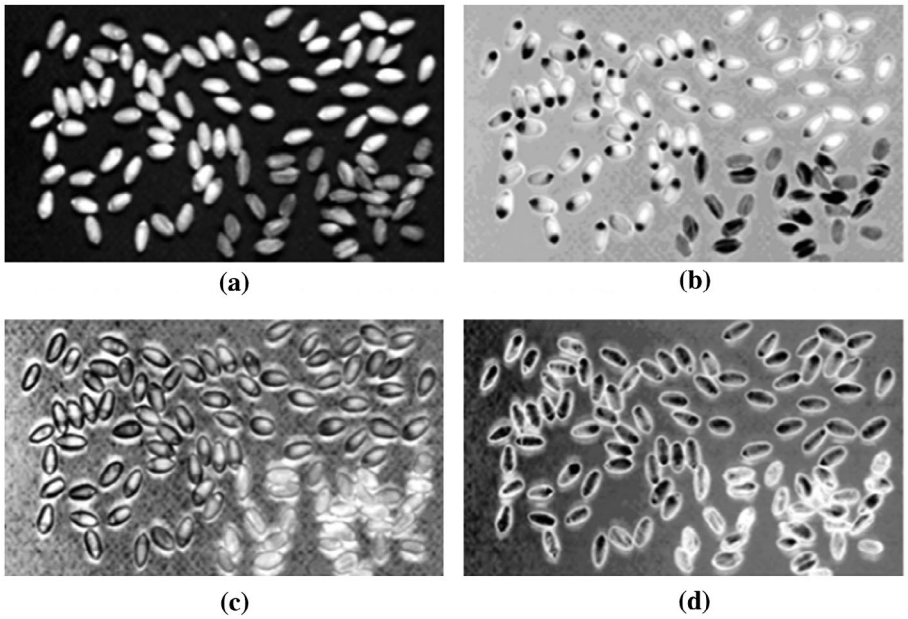

Figure 3. The PC Images of PCA on Different Spectral Regions of 400 to 1,000 nm (a) Principal Component 1(a), 2(b), 3(c) and 4(d).

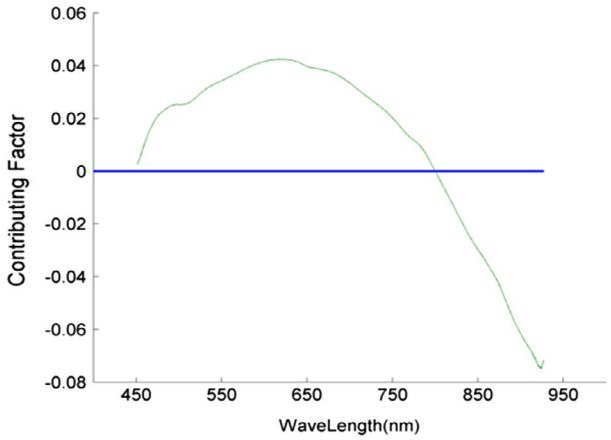

Figure 4. Spectral Range Contribution for the Feature of the Second Principal Component.

Table 1. Extracted Features of Individual Wheat Kernels.

Morphological Features			
No.	Code	Feature explanation	Formula
1	A	Area of kernel	$\frac{1}{2}\sum_{i=0}^{r-1}\left(x_i y_{i+1} - x_{i+1} y_i\right)$
2	W	Width of kernel	$\left(\frac{2\left[\mu_{2,0}+\mu_{0,2}+\sqrt{\left(\mu_{2,0}-\mu_{0,2}\right)^2+4\mu_{1,2}^2}\right]}{\mu_{0,0}}\right)^{\frac{1}{2}}$
3	L	Length of kernel	$\left(\frac{2\left[\mu_{2,0}+\mu_{0,2}-\sqrt{\left(\mu_{2,0}-\mu_{0,2}\right)^2+4\mu_{1,2}^2}\right]}{\mu_{0,0}}\right)^{\frac{1}{2}}$
4	Fc	Roundness factor	$\frac{P^2}{4\pi A}$
5	Rl	Diameter ratio	$\frac{L}{W}$
6	P	Perimeter	$\sum_{i=0}^{r-1}\sqrt{\left(x_{i+1}-x_i\right)^2+\left(y_{i+1}-y_i\right)^2}$
7	Ra	Rectangular ratio	$\frac{A}{L*W}$
Texture features			
8–15	Energy	4 level wavelet energy of R	$\frac{1}{A}\sum_{x=1}^{N}\sum_{y=1}^{M}\lvert I(x,y)\rvert^2$
16–23	Entropy	4 level wavelet entropy of R	$\frac{1}{A}\sum_{x=1}^{N}\sum_{y=1}^{M}\lvert I(x,y)\rvert^2 \log \lvert I(x,y)\rvert^2$

Note: Where *I(x,y)* is intensity expressed in grey level, μ_i is the *i*th central moment of the intensity distribution.

IV. Discussion

As we know, the data of hyperspectral was so huge that it was difficult to be analysed online. So, it was vital to find the best band. In this study, according to the contributing factor after PCA analysis for hyperspectral data, the best band was 617 nm, of course, there is a fluctuating range at this band, and the range should be assessed based on actual situation. The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has the largest possible variance, and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components. In this study, the second Principal Component (PC2), as Figure 3(b), the component can show flaw clearly, also, it's easy to extract features after image processing, therefore, PC2 is much better than others. For PC2, different hyperspectral bands have different contributions, which can be measured by the contributing factor (Figure 4); those useful bands were similar to red band in RGB. Therefore, we selected the red band as the best band to identify black germ kernels, break kernels and healthy kernels, although the red band is not the best one, from the view of economic, it can reduce unnecessary spending and complex steps. A software package was developed to extract various morphological features in red band, which were listed in the Table 1. The performance of wheat class discrimination models was developed by the methods of PLSDA, SVMDA and PCA-ANN; those algorithms are combination between feature extraction and pattern recognition.

The samples of each category were split into 2 subsets at random in order to avoid man-made interference. Sixty percent of the samples were used to train and the other parts (40%) were used for testing. The percentage of sample correctly classified were 91.3%, 92.5%, 94.1% (black germ kernel, break kernel and healthy kernel, respectively) for PLSDA, 95.6%, 96.7%, 98.5% for SVMDA and 90.3%, 91.6%, 93.5% for PCA-ANN.

Table 2. The Accuracy Comparison between Different Models.

Discriminate analysis	Black germ kernel (%)	Break kernel (%)	Healthy kernel (%)
PLSDA	91.3	92.5	94.1
SVMDA	95.6	96.7	98.5
PCA-ANN	90.3	91.6	93.5

Table 3. The Sensitivity and Specificity of Different Categories.

	Sensitivity (prediction)			Specificity (prediction)		
LVs	BGK	BK	HK	BGK	BK	HK
1	0.039	0.760	0.697	0.742	0.461	0.821
2	0.064	0.888	0.718	0.854	0.853	0.852
3	0.811	0.840	0.803	0.918	0.883	0.980
4	**0.910**	**0.848**	**0.803**	**0.925**	**0.912**	**0.989**
5	0.910	0.920	0.972	0.925	0.957	0.994
6	0.906	0.936	0.979	0.966	0.973	0.994
7	0.893	0.944	0.958	0.985	0.968	0.994
8	0.897	0.944	0.972	0.978	0.981	0.997

Note: Bold values represent the optimum LVS selection.

The statistical parameters of the results obtained from the studied models for test set were summarized in Table 2. According to the result, SVMDA proved to be the most efficient method. And the band selection based on hyperspectral image system is feasible in this study.

Sensitivity and specificity are statistical measures to evaluate the performance of a binary classifier, in statistics they are known as performance parameters. Sensitivity (also called the true positive rate or the recall rate in some fields) measures the proportion of actual positive, which is correctly identified, such as, the percentage of sick people who are correctly identified as meeting the conditions. Specificity (sometimes called the true negative rate) measures the proportion of negatives, which is correctly identified, such as, the percentage of healthy people who are correctly identified as not meeting the conditions. These two measures are closely related to the concepts of type I and type II errors. We also got the Sensitivity and specificity, which can be get using different PLS components. The principle of PLS is used to find the components in the input matrix X that describes the relevant variations as much as possible and has a great correlation with the target value Y, then gives less weight to the variations that are irrelevant. We can see the results from Table 3, which shows the value of specificity and sensitivity with different PLS component. This method is also a way to choose the number of PLS.

Those methods are reference for durum kernel online detection, of course, there are several methods to detect kernel or something like those, different research has their own advantages. The goal of hyperspectral imaging in this paper is to obtain the spectrum for each pixel in the image of durum kernel, with the purpose of finding the flaws of kernel, identifying the defects of durum kernel.

Further research: We should try other combined methods to test the accuracy for unsound kernel identification. Also, we should make some further experiments to verify the efficiency: 1) wheat variety "A" vs. wheat variety "B", 2) commercial wheat class "A" vs. commercial wheat class "B", 3) wheat seed vs. other component of wheat plant, 4) wheat seed vs. other seed species, 5) seed vs. stone, 6) wheat damaged kernel (heat, frost, mold, etc.) vs. sound kernel, 7) wheat diseased kernel (scab, black tip, etc.) vs. healthy kernel, and so on.

Another method is to use wider spectral data to analysis this data, so as to get a variety of explanations for the kernel detection, such as ultraviolet rays, near-infrared and fluorescence.

V. Conclusion

In the spectral region ranging from 452 nm to 930 nm, the reflectance of hyperspectral image from unsound tissue was generally lower than that from normal tissue. The second principal component images received by principal component analysis of the wheat kernel samples were useful. Although the accuracy of PC2 was high enough for detection, this method was incredible as the data was too huge for real-time application; it was desirable to use fewer wavelengths for rapid image acquisition and processing. From waveband contributor of PC2, two extreme wavelengths were 617 nm and 925 nm. As for the method of features extraction at wavelength 617 nm, it will be readily seen that the method of image segmentation was easy to realize, and the best detection accuracy can reach 95.6%, 96.7% and 98.5% for 710 black germ kernels (BGK), 627 break kernels (BK) and 1169 healthy kernels (HK), respectively. The general detection performance analysis demonstrated that the wavelength of 617 nm can be implemented effectively in the machine vision system.

Acknowledgment

We thank the Project 61305037 supported by National Natural Science Foundation of China (NSFC).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Swedish Institute [grant number 05612/2015], national Natural Science Foundation of China [grant number 61305037].

Notes on contributors

Feng-Nong Chen received a Ph.D. in 2012 from Zhejiang University, and worked 1 year at KTH Royal Institute of Technology in Sweden as post doc. His research interests include computer vision, image processing and pattern recognition, especially in the field of agricultural product and Unmanned aerial vehicle of plant protection.

Pu-Lan Chen is a senior accountant in Cancer Affiliated hospital of Xinjiang Medical University. Her research interests include data analysis and economic management in her department.

Kai Fan received his Ph.D. in 2014 from Zhejiang University. Presently, he is a lecturer at HangzhouDianzi University. His research interests include nano electronic functional materials, electrochemical preparation of nano structural materials and electrochemical sensors and biosensors.

Fang Chen is a professor in Zhejiang University. Her research interests are image-based quality sensing technology for agro-products.

References

Ariana, D.P., Lu, R.F., & Guyer, D.E. (2006). Near-infrared hyperspectral. eflectance imaging for detection of bruises on pickling cucumbers. *Computers and Electronics in Agriculture, 53*, 60–70.

Choudhary, R., Paliwal, J., & Jayas, D.S. (2008). Classification of cereal grains using wavelet, morphological, colour, and textural features of non-touching kernel images. *Biosystems Engineering, 99*, 330–337.

Choudhary, R., Mahesh, S., Paliwal, J., & Jayas, D.S. (2009). Identification of wheat classes using wavelet features from near infrared hyperspectral images of bulk samples. *Biosystems Engineering, 102*, 115–127.

Cortes, C. & Vapnik, V. (1995). Support-Vector networks. *Machine Learning, 20*, 273–297.

Del Fiore, A., Reverberi, M., Ricelli, A., Pinzari, F., Serranti, S., Fabbri, A.A., Bonifazi, G., & Fanelli, C. (2010). Early detection of toxigenic fungi on maize by hyperspectral imaging analysis. *International Journal of Food Microbiology, 144*, 64–71.

Delwiche, S.R. & Kim, M.S. (2000). Hyperspectral imaging for detection of scab in wheat. *Proceedings of SPIE - The International Society for Optical Engineering, 4203*, 13–20.

Delwiche, S.R., Yang, I.C., & Graybosch, R.A. (2013). Multiple view image analysis of freefalling U.S. wheat grains for damage assessment. *Computers and Electronics in Agriculture, 98*, 62–73.

Flood, M.E., Connolly, M.P., Comiskey, M. C., & Hupp, A. M. (2016). Evaluation of single and multi-feedstock biodiesel - diesel blends using GCMS and chemometric methods. *Fuel, 186*, 58–67.

Gharibzahedi, S.M.T., Mousavi, S.M., Jouki, M., & Ghanderijani, M. (2012). Analysis of physicochemical and thermomechanical characteristics of Iranian black seed (Nigella oxypetala Boiss). *International Journal of Food Engineering, 8*, 1–14.

Gorretta, N., Roger, J.M., Aubert, M., Bellon-Maurel, V., Campan, F., & Roumet, P. (2006). Determining vitreousness of durum wheat kernels using near infrared hyperspectral imaging. *Journal of Near Infrared Spectroscopy, 14*, 231–239.

Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences, 79*, 2554–2558.

Kotwaliwale, N., Singh, K., Kalne, A., Jha, S.N., Seth, N., & Kar, A. (2014). X-ray imaging methods for internal quality evaluation of agricultural produce. *Journal of Food Science and Technology, 51*(1), 1–15.

Kwon, H. & Nasrabadi, N.M. (2007). Kernel spectral matched filter for hyperspectral imagery. *International Journal of Computer Vision, 71*, 127–141.

Maganioti, A.E., Hountala, C.D., Papageorgiou, C.C., Kyprianou, M.A., Rabavilas, A.D., & Capsalis, C.N. (2010). Principal component analysis of the P600 waveform: RF and gender effects. *Neuroscience Letters, 478*, 19–23.

Majumdar, S. & Jayas, D.S. (2000a). Classification of cereal grains using machine vision: I. Morphology models. *Transactions of the ASAE, 43*, 1669–1675.

Majumdar, S. & Jayas, D.S. (2000b). Classification of cereal grains using machine vision: Ii. Color models. *Transactions of the ASAE, 43*, 1677–1680.

Majumdar, S. & Jayas, D.S. (2000c). Classification of cereal grains using machine vision: III. *Transactions of the ASAE, 43*, 1681–1687.

Majumdar, S. & Jayas, D.S. (2000d). Classification of cereal grains using machine vision: IV. Combined morphology, color, and texture models. *Transactions of the ASAE, 43*, 1689–1694.

Mohammad, J., Emam-Djomeh, Z., & Khazaei, N. (2012). Physical properties of whole rye seed (Secale cereal). *International Journal of Food Engineering, 8*(4), 1–14.

Nakariyakul, S. & Casasent, D.P. (2011). Classification of internally damaged almond nuts using hyperspectral imagery. *Journal of Food Engineering, 103*, 62–67.

Qing, S., Tao, X., Tatsuo, Y., Nan, S., Hang, Z. (2016). Classified denoising method for laser point cloud data of stored grain bulk surface based on discrete wavelet threshold. *International Journal of Agricultural and Biological Engineering, 9*, 123–131.

Ridgway, C., Chambers, J., & Cowe, I.A. (1999). Detection of grain weevils inside single wheat kernels by a very near infrared two-wavelength model. *Journal of Near Infrared Spectroscopy, 7*, 213–221.

Singh, C.B., Jayas, D.S., Paliwal, J., & White, N.D.G. (2010). Identification of insect-damaged wheat kernels using short-wave near-infrared hyperspectral and digital colour imaging. *Computers and Electronics in Agriculture, 73*, 118–125.

Venora, G., Grillo, O., & Saccone, R. (2009). Quality assessment of durum wheat storage centres in Sicily: Evaluation of vitreous, starchy and shrunken kernels using an image analysis system. *Journal of Cereal Science, 49*, 429–440.

Wallays, C., Missotten, B., De Baerdemaeker, J., & Saeys, W. (2009). Hyperspectral waveband selection for on-line measurement of grain cleanness. *Biosystems Engineering, 104*, 1–7.

Wiwart, M., Suchowilska, E., Lajszner, W., & Graban, Ł. (2012). Identification of hybrids of spelt and wheat and their parental forms using shape and color descriptors. *Computers and Electronics in Agriculture, 83*, 68–76.