Advance Journal of Food Science and Technology 7(8): 652-657, 2015
DOI:10.19026/ajfst.7.1623
ISSN: 2042-4868; e-ISSN: 2042-4876

Submitted: December 20, 2014 Accepted: January 27, 2015 Published: March 15, 2015

**Research Article**

# Egg Freshness Detection Based on Hyperspectral Image Technology

[1, 2]Qiaohua Wang, [1]Kai Zhou, [1]Caiyun Wang and [2, 3]Meihu Ma
[1]College of Engineering,
[2]The National Egg Processing Technology, R and D Center,
[3]College of Food Science and Technology, Huazhong Agricultural University, Wuhan 430070, China

**Abstract:** The storage period of eggs will affect the freshness, therefore the quality and effectiveness of inspection methods to determine freshness is vital. The transmission spectra of egg samples were obtained by hyperspectral image system. Spectrum data and images of egg samples were extracted by ENVI software. After filtering and de-noising, the sensitive waveband (550~900 nm) of spectrum curve were processed by length function. The spectral values of characteristic wavelengths of 625, 650, 675, 750 and 810 nm, respectively, were selected as the characteristic parameters of spectrum. At the same time, R, G, B components of egg samples' spectral images were chosen as the characteristic parameters of spectrum. After analyzing the 8 characteristic parameters by using Principal Component Analysis (PCA) and Stepwise Multiple Linear Regression (SMLR), four parameters were selected as spectral parameters of egg samples which could represent 97.66% spectral information. Using the four parameters ($\lambda_{650}$, $\lambda_{675}$, $X_R$ and $X_G$) as the independent variable, Haugh unit prediction model of egg samples was established by Multiple Linear Regression (MLR). The correlation coefficient R of calibration set was 0.925 and the correlation coefficient R of prediction set was 0.908. This suggested that the egg freshness detection with hyperspectra was feasible.

**Keywords:** Egg, freshness, hyperspectra, MLR, PCA, SMLR

## INTRODUCTION

In recent years, Chinese egg production has ranked first in the world (Yali, 2008). For the egg industry, quality classification is very important. Along with the increase of storage time, the quality of eggs declines and decayed eggs also can affect the surrounding eggs (Wu and Han, 1989). Foreign countries have already established a set of egg classification standards and developed sets of highly automation testing equipment. Domestic quality detection and classification of eggs need to be improved which can adapt to the requirements of large scale automation (Danming, 2012).

Because of short detection time, high efficiency and good economic benefit, nondestructive detection of egg quality meets the requirements of modern production requirements. Many scholars have successfully researched egg quality inspection by using near infrared technology (Hou *et al*., 2009; Yanying, 2012; Yande *et al*., 2010) and have established the model to detect egg freshness. Based on the spectrum values of visible light, identification model of Haugh unit value and albumen PH value have been established, respectively by using Partial Least Squares (PLS). The result shows the effects of the two are very similar (Bart *et al*., 2006). With visible and near infrared spectroscopy technique, the detection model of egg freshness and protein content is established successfully by using PLS (Nicolas *et al*., 2009). There are many researches on egg quality with visible light and near infrared at home and abroad. However, there is little study on the detection of egg quality based on hyperspectrum. The study on hatching egg using hyperspectrum is of good effect (Wei *et al*., 2012). The main objective of the present study is to investigate the relationship between the spectral value of stored egg and its freshness and to provide a new method for the egg freshness nondestructive inspection.

## MATERIALS AND METHODS

**Test egg samples:** Two hundred Roman pink shell eggs with one day old were obtained from the Jiu Feng Chicken Farm in Wuhan. The eggs were stored in the temperature of 24°C with the relative humidity of 70%.

**Test equipment:** Hyperspectral image system (Hypersis-VNIR-CL, American) was used to acquire the spectra and image information of egg samples on transmission mode, which is mainly composed of a computer, a tunable light source, a hyperspectral image

**Corresponding Author:** Qiaohua Wang, College of Engineering, Huazhong Agricultural University, Wuhan 430070, China

unit, a mobile device, a sample holder station and the image acquisition software (Wei *et al*., 2012). The hyperspectral image unit consists a CCD camera, a zoom lens and an image spectrometer whose spectral resolution was 8 nm. The effective waveband of the system ranged from 400 to 1000 nm. The wavelength accuracy of this system was 0.5 nm. Its effective pixels was 1392 by 1040. In order to avoid the interference of the external light source, the others were placed in a black box except the computer.

The egg quality analyzer (05-UM-001 Version E, Orka Technology Ltd, Israel) which was composed of a weight measuring system, a sliding disc, a Liquid Crystal Display (LCD) monitor, a control panel and a built-in printer, was used to determine the freshness of egg samples.

**Parameters setting and correction of hyperspectral image system:** In order to obtain accurate data, some parameters of hyperspectral image system needed to be set before image acquisition. In this study, the parameters were set as: 0.5 sec exposure time, 1.8 mm/sec move speed, 180 mm move distance and the scan distance ranged from 120 to 300 mm. The images were recorded with spatial dimension of 1392 by 500 pixels (Wei *et al*., 2012).

In order to exclude the impact of the camera dark current and external noise (Polder *et al*., 2003), the camera must be calibrated with a white and dark references before acquisition (Iqbal *et al*., 2013):

$$R = (Ro\text{-}Rd) / (Rw\text{-}Rd)$$

where,

- $R$ = The spectral image file of egg samples after correction
- $Ro$ = The original spectral image file of egg samples
- $Rd$ = The dark image file under the condition of total darkness
- $Rw$ = The white image file under the condition of whiteness (Wei *et al*., 2012)

**Test method and steps:**

- Select 20 eggs randomly and measure them on the storage time of 0, 3, 6, 9, 12, 15, 18, 21, 24 and 27 days, respectively
- Place the eggs one at a time vertically on the hole of sample holder station on the transmitted light box and ensure there is no light leakage from the transmitted light box. Move the sample holder station to a suitable position. Acquire the transmitted hyperspectral image which contains 520 wavebands through the camera
- After collecting spectrum data, measure the weight and Haugh unit values of egg samples
- Record all data

## RESULT ANALYSIS

The spectral data and images of egg samples were obtained respectively by ENVI4.7 analysis software. This study used MATLAB to extract their characteristic parameters and stepwise regression to identify significant parameters influencing the quality of egg samples greatly through Statistical Analysis System (SAS). Finally, calibration model was established between these parameters and the quality of eggs.

**Selection of sensitive wavebands and characteristic wavelengths:** The spectrum of egg samples included all information about the quality of eggs as well as contained noise and interference. Sometimes, the noise and interference information strength were greater than the eggs' and concealed the details of egg quality information. Therefore, the information deemed unnecessary must be weakened or eliminated and be separated from the information related to eggs' quality. In order to highlight the spectral information of egg quality, pretreatment was carried out on the original spectrum.

The high spectrum contained a large amount of data which was very difficult to analyze; Effective information could represent the majority of data. Therefore, this study had to extract effective information from the high spectrum. Here, the effective information included the characteristic wavelengths and parameters of spectrum.

With ENVI 4.7, this study selected 5 square areas (100*100 pixels) in each egg and extracted their spectral data. One of them was in the center of egg yolk, the others were distributed over albumen. MATLAB was used to calculate the average of their spectra and draw corresponding waveform curves. In the range of 400-500 and 900-1000 nm, the values of spectrum curve were low, which meant there was little effective information to reflect the quality of egg. In the range of 500~900 nm, the spectrum curve appeared for the first time to peak and valley and the influence of noise was small. This indicated the waveform curve could better reflect the characteristics of egg samples. Median filtering was used to filter out the noise of spectrum curve in the sensitive band. In this study, it was better to set 0.01 gradient by using the length function. And then select extreme points which was to be the characteristic spectral parameters on the spectral curve.

In Fig. 1, curve inflection point appeared in the wavelength of 625, 650, 675, 750 and 810 nm, respectively. These wavelengths were chosen as the characteristic wavelengths and the spectral values were marked as $\lambda_{625}$, $\lambda_{650}$, $\lambda_{675,}$ $\lambda_{750}$ and $\lambda_{810}$ at these wave lengths as five characteristic parameters of the egg samples. Then principal component analysis was used to determine the characteristic parameters of a relatively large impact on the quality of egg samples.

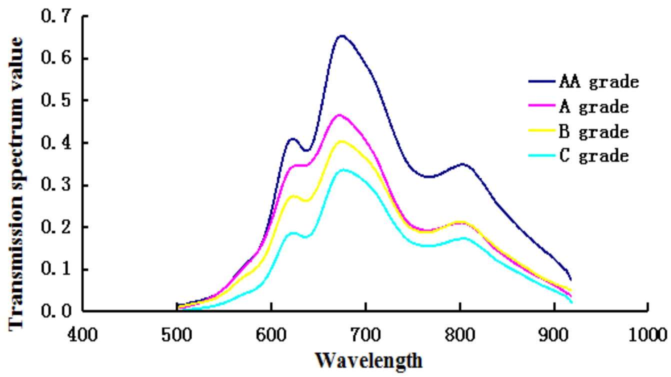
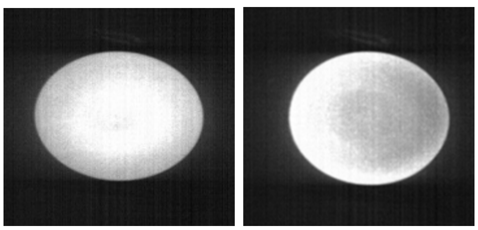



Fig. 1: Spectral curves of egg samples with different fresh degree



(a) Early period (b) Late period

Fig. 2: Original spectral images of the same egg sample in the early and late period

Table 1: Eigenvectors of principal component analysis

| | Eigenvalue | Difference | Proportion | Cumulative |
|---|---|---|---|---|
| 1 | 4.691 | 2.638 | 0.5863 | 0.5870 |
| 2 | 2.053 | 1.425 | 0.2566 | 0.8448 |
| 3 | 0.627 | 0.185 | 0.0784 | 0.9233 |
| 4 | 0.442 | 0.323 | 0.0553 | 0.9766 |
| 5 | 0.119 | 0.069 | 0.0149 | 0.9916 |
| 6 | 0.051 | 0.041 | 0.0063 | 0.9979 |
| 7 | 0.020 | 0.002 | 0.0012 | 0.9991 |
| 8 | 0.007 | | 0.0009 | 1.0000 |

**Spectral image preprocessing:** With the increase of storage time, the change of physical, chemical, biological of egg samples caused the change of internal moisture, PH etc. Therefore, the transmission rates would also change. These changes were reflected in their spectral images. With ENVI spectral analysis software, parameters of R, G, B were selected to synthesize color images of egg samples (Shubin, 2010). As shown in Fig. 2, the original spectral image of egg samples was preserved in the main image window of ENVI4.7.

Figure 2a was an original spectral image of egg sample in the early stage of storage, Fig. 2b was an original spectral image of the same egg sample at the end of storage period. From the original spectral images of egg samples, it was clear that the color of egg sample images deepened along with the increase of storage time. The noise in the images and the background of it would interfere in the extraction of feature parameters of spectral images. Therefore, the background elimination and noise cancellation were achieved by using the Otsu method to find a suitable threshold value.

As blue, red and green could synthesize all the colors of nature, computer used RGB model to display the color of each pixel in picture based on changing the ratio of R, G, B. Each color was synthesized with different content of R, G, B (Zheng *et al*., 2010). Each color difference of pixel could reflect the difference between pixels in image. Therefore, the component values of $X_B$, $X_R$, $X_G$ were extracted from spectral images of egg samples by MATLAB and they were used as three characteristic parameters to evaluate egg quality. Then principal component analysis was used to determine the characteristic parameters of a relatively large impact on the quality of egg samples.

**Principal component analysis:** With the MLR method, the characteristic parameters of egg samples' spectral curve and spectral image were looked as two sets of independent variables to establish prediction model of Haugh unit values, respectively. Predictions of the two models were poor, which indicated the information of using the two sets of parameters for modeling separately was not enough. Therefore, both the spectral parameters and image parameters were used to establish model.

Through analyzing the correlation of multiple variables, principal component analysis was a method to reduce the number of variables without losing the original information and choose a few variables which could represent the overall index of the original variables (Baojin, 2007). In the study, the number of spectral data parameters and spectral image parameters was too much, which led to a large amount of actual operation task. Therefore, it was necessary to reduce the number of parameters on the basis of high accuracy.

Table 1 showed the eigenvalues of sample covariance matrix, in which the figure of 1, 2, 3, 4, 5, 6, 7 and 8, respectively represented eight principal components. In the table, the eigenvalue column represented eigenvalues of sample covariance matrix.

Table 2: Summary of stepwise selection

| Step | Variable entered | Partial R-square | Model R-square | C (p) | F-value | Pr>F |
|---|---|---|---|---|---|---|
| 1 | $\lambda_{675}$ | 0.565 | 0.565 | 62.850 | 189.57 | <0.0001 |
| 2 | $\lambda_{650}$ | 0.070 | 0.635 | 31.538 | 27.83 | <0.0001 |
| 3 | $X_G$ | 0.036 | 0.671 | 16.401 | 15.78 | <0.0001 |
| 4 | $X_R$ | 0.031 | 0.702 | 3.691 | 14.85 | <0.0001 |

Table 3: Linearity hypothesis test of model

| | df | S.S. | M.S. | F-value | Significance F-value |
|---|---|---|---|---|---|
| Regression analysis | 4 | 11051.6642 | 2762.9161 | 215.5238 | $5.98\times10^{-60}$ |
| Residual error | 145 | 1858.8337 | 12.8195 | | |
| Total | 149 | 12910.4979 | | | |

S.S.: Sum of square; M.S.: Mean square

Table 4: Regression coefficients remarkable test of model

| | Coefficients | S.E. | t-value | p-value | Lower 95% |
|---|---|---|---|---|---|
| Intercept | -34.8216 | 5.9203 | -5.8817 | $2.69\times10^{-08}$ | -46.5228 |
| $\lambda_{650}$ | -37.5865 | 6.6225 | -5.6756 | $7.27\times10^{-08}$ | -50.6756 |
| $\lambda_{675}$ | 101.7419 | 5.4713 | 18.5956 | $3.19\times10^{-40}$ | 90.9281 |
| $X_R$ | -0.0882 | 0.0150 | -5.8958 | $2.51\times10^{-08}$ | -0.1177 |
| $X_G$ | 0.3309 | 0.0374 | 8.8517 | $2.75\times10^{-15}$ | 0.2570 |

S.E.: Standard error

Proportion column showed the variance contribution rates of the principal component. Cumulative column listed the accumulative contribution rate of principal components. From the table, it was found that variance cumulative contribution rate of the first four principal components was 0.9766, which meant that the four principal components could represent 97.66% information of the original variable. The first four principal components retained a vast majority of hyperspectral information of egg samples. From the first to the fifth principal component, variance contribution rates became very small which could be ignored. Therefore, as long as the characteristic parameters of the first four principal components were found, the model could be established.

**Establish and validate the model:** The results of principal component analysis showed that only 4 main variables could represent the vast majority of the general information of the original variables. Therefore, parameters of spectral curve and spectral image which retained a vast majority of information of the original variables were chosen with the SMLR. And then the model between the parameters and Haugh unit value of egg samples was established. The following would select 4 main characteristic variables and establish the SMLR discriminant model of the Haugh unit value.

Table 2 showed the specific process of how to choose the main characteristic parameters through the SMLR method. And the four variables were used in the regression model in the order of $\lambda_{675}$, $\lambda_{650}$, $X_G$ and $X_R$. All variables in the regression model was significant under the level of $\alpha = 0.15$ and unselected variables was not significant under the level of $\alpha = 0.15$. It also demonstrated the increase value of $R^2$, the cumulative value of $R^2$, the corresponding F value and the p value when each new parameter entered the model. In Table 2, when $X_R$ entered the model, C (P) was 3.691 which was close to the variable number of 4. This showed that the established model was relatively better. Meanwhile, p values were less than 0.15, which also showed that variables of $\lambda_{650}$, $\lambda_{675}$, $X_R$ and $X_G$ could retain a higher contribution rate for modeling. However, variables $\lambda_{625}$, $\lambda_{750}$, $\lambda_{810}$ and $X_B$ with lower contribution rate for modeling should be removed.

Therefore, the four chosen variables ($\lambda_{650}$, $\lambda_{675}$, $X_R$ and $X_G$) were used to establish the prediction model of Haugh unit value by MLR:

$$Hu = -34.8216 - 37.5865\lambda_{650} + 101.7419\lambda_{675} - 0.0882X_R + 0.3309X_G$$

The correlation coefficient: $R = 0.925$.

where,

- $H_u$ = The Haugh unit value
- $\lambda_{650}$ = The characteristic value of spectral curve in the wavelength of 650
- $\lambda_{675}$ = The characteristic value of spectral curve in the wavelength of 675
- $X_R$ = The R component value of spectral image
- $X_G$ = The G component value of spectral image

The regression analysis of model showed that the correlation coefficient R was 0.925, which indicated the model was reliable. In Table 3, the F = 215.5238 was far greater than significance F = 2.4341 after the hypothetical test, which showed high reliability of the model.

The results of significant test for each regression coefficient were shown in Table 4. For $\lambda_{650}$, the t = 5.6756 was larger than the significance t = 1.9765. For $\lambda_{675}$, the t = 18.5956 was larger than the significance t = 1.9765. For $X_R$, the t = 5.8958 was larger than the significance t = 1.9765. For $X_R$, the t = 8.8517 was larger than the significance t = 1.9765.

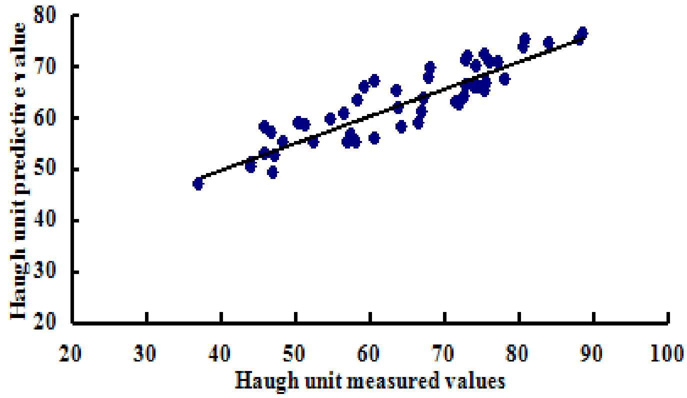



Fig. 3: Performance of PLSR model

For intercept, the t = 5.8817 was larger than the significance t = 1.9765. This showed that the regression coefficients at significance level of $\alpha = 0.15$ could pass the t test. All p values of regression coefficients were less than 0.01, which showed that the confidences of the regression coefficients could reach more than 99%. Thus, using the four chosen spectral parameters ($\lambda_{650}$, $\lambda_{675}$, $X_R$ and $X_G$) to establish the Haugh unit value prediction model of egg samples were feasible.

Using $\lambda_{675}$, $\lambda_{650}$, $X_G$ and $X_R$ to establish the prediction model of egg freshness, the correlation coefficient R of calibration set was 0.925. The effect of prediction set was shown in Fig. 3 and the correlation coefficient was 0.908.

## CONCLUSION

- The study finds that the hyperspectral sensitive waveband of stored eggs lies between 550~900 nm and characteristic wavelengths were 625, 650, 675, 750 and 810 nm, respectively.
- The result of principal component analysis shows that the four main characteristic parameters ($\lambda_{650}$, $\lambda_{675}$, $X_R$ and $X_G$) can represent 97.66% spectrum information of egg samples.
- Through MLR method, four parameters ($\lambda_{650}$, $\lambda_{675}$, $X_R$ and $X_G$) are selected to establish the Haugh unit value prediction model of egg samples. The correlation coefficient R is 0.925, which means that the prediction model of egg samples is very significant.
- The correlation coefficient R of calibration set is 0.925 and the correlation coefficient R of prediction set is 0.908. The result shows that using hyperspectral image technology to study the freshness of stored eggs is feasible.

## ACKNOWLEDGMENT

This study was supported by National Science Foundation of China (31371771), Corpus of Public Welfare Industry (agriculture) Scientific Research Projects (201303084) and National Science and Technology Support Programme (2015BAD19B05).

## REFERENCES

Baojin, W., 2007. Multivariate Analysis: Statistical Software and Data Analysis. Peking University Press, Beijing, China.

Bart, J.K., R.B. Flip, D.K. Bart, M. Kristof, T. Kokou, M.D. Eddy and G.D.B. Josse, 2006. Visible transmission spectroscopy for the assessment of egg freshness. J. Sci. Food Agr., 86(9): 1399-1406.

Danming, T., 2012. Study of egg quality testing technology based on near infrared transmission spectroscopy. China Jiliang University.

Hou, Z.C., N. Yang, J.Y. Li and G.Y. Xu, 2009. Egg quality prediction by using fourier transform near infrared reflectance spectroscopy (FT-NIR). Spectrosc. Spect. Anal., 29(8): 2063-2066.

Iqbal, A., D.W. Sun and P. Allen, 2013. Prediction of moisture, color and pH in cooked, pre-sliced turkey hams by NIR hyperspectral imaging system. J. Food Eng., 117(1): 42-51.

Nicolas, A.N., M. Ngadi and S. Prasher, 2009. Prediction of egg freshness and albumen quality using visible/near infrared spectroscopy. Food Bioprocess Tech., 10(1): 265-271.

Polder, G., W.A.M. Gerie and D.H. Van, 2003. Calibration and characterization of imaging spectrographs. Ellis. Hor. S., 11(3): 193-210.

Shubin, D., 2010. ENVI the Method of Remote Sensing Image Processing. Science Press.

Wei, Z., P. Leiqing and T. Kang, 2012. Detecting early embryo development of chicken hatching eggs by hyperspectral transmittance imaging. Trans. Chin. Soc. Agr. Eng., 28(21): 149-155.

Wu, S. and N. Han, 1989. Egg freshness optical non-destructive testing and grading. Trans. Chin. Soc. Agr. Eng., 5(4): 64-70.

Yali, J., 2008. Analysis of 2008 layer market. Henan J. Anim. Husbandry Vet. Med., 29(3): 48-48.

Yande, L., P. Yanying and S. Xudong, 2010. A study on variable selection of vis-NIR spectral information for online detection albumen pH of eggs. Acta Agr. Univ. Jiangxiensis, 32(5): 1075-1080.

Yanying, P., 2012. Non-destructive measurement of eggs quality using near infrared spectroscopy. East China Jiaotong University, pp: 194.

Zheng, Z., W. Yanping and X. Guixiang, 2010. Digital image processing and machine vision. The People's Posts and Telecommunications Press, China.