Nondestructive Internal Defects Evaluation for Pear Using NIR/VIS Transmittance Spectroscopy

D. S. Ryu, S. H. Noh, H. Hwang

Abstract: Internal defects such as browning of the flesh and blackening and rot of the ovary of pear can be easily developed because of the inadequate environmental conditions during the storage and distribution of fruit. The quality assurance system for the agricultural product is to be settled in Korea. All defected agricultural products should be excluded prior to the distribution to enhance the commercial values. However, early stage on-line defect detection of agricultural product is very difficult and even more difficult in a case of the internal defects. The goal of this research is to develop a system that can detect and classify internal defects of agricultural produce on-line using VIS/NIR transmittance spectroscopy. And Shingo pear, which is one of the famous species of Korean pear, was used for the experiment. Soft independence modeling of class analogy (SIMCA) algorithm was employed to analyze the transmittance spectroscopic data qualitatively. On-line classification system was constructed and classification model was developed and validated. As a result, the correct classification rate (CCR) using the developed classification model was 96.1%.

Keywords: On-line, Pear, Internal Defects, VIS/NIR, Transmittance, Spectroscopy, SIMCA

Introduction

Owing to the current refrigerated storage technology, the storage period of most agricultural products has been extended and hence consumers could get the fresh ones at any time. Producers also could avoid the rapid price down caused by an excessive supply at harvest by means of controlling the amount of discharge to the market. However, if the storage condition becomes improper, the quality of stored fruits deteriorates and hence the amount of fruits with defects including rotten fruits rapidly increase. It is reported that the loss of fruit is from 20 to 30% during the storage and distribution of fruits. In addition, since the quality assurance system for agricultural products is to be settled in Korea, development of nondestructive quality evaluation technologies for fruit is urgently required.

There are various kinds of internal defects in agricultural products. Most of them are caused by the physiological disease or damage during storage. In a case of Korean pears, principal internal defects are developed during storage and are the discoloration or rot of the ovary and the flesh (Fig. 1). Other defect of pear is the 'spongy disorder' shown in Fig. 2. The flesh tissue of pear is softened and its structure becomes weak. Though the physiological process of the sponge disorder has not been made clear yet, this defect is considered as a refrigeration damage.

Several methods are available to detect the internal defects of fruit. Some of them are ultrasonic, Nuclear Magnetic Resonance (NMR), Magnetic Resonance Imaging (MRI), etc. NMR and MRI methods are quite good enough for nondestructive detection of internal defects but too expensive and are relatively slow at processing.

Recently, VIS/NIR transmittance spectroscopy has been widely used for the quantitative analysis of the internal constituents such as sugar contents, acidity of fruits and so on (Hwang, 1999; Hwang, 2000; Ryu, 2001). This method has been successfully applied to the automatic fruit sorter and used in the packing facility in Korea. In this paper, VIS/NIR transmittance spectroscopy method was used to develop the nondestructive internal quality evaluation of fruit primarily. Secondly the classification algorithms and models based on qualitative analysis methods were

The authors are **Dong Soo Ryu,** Ph.D., Director, Research Institute of Non-destructive Quality Evaluation, Life & Technology Co. Ltd., Suwon, KOREA, **Sang Ha Noh,** Professor, School of Biological Resources and Materials Engineering, College of Agriculture and Life Sciences, Seoul National University, Seoul, Korea, **Heon Hwang,** Professor, Department of Bio-Mechatronic Engineering, Faculty of Life Science and Technology, Sungkyunkwan University **Corresponding author:** Dong Soo Ryu, Ph.D., Director, Life & Technology Co. Ltd., Suwon, KOREA; e-mail: ryuds@ lifentech. com

developed and its classification performance was validated.

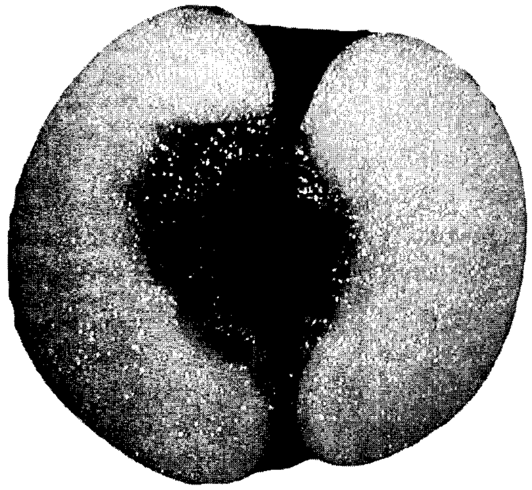


Fig. 1 Discoloration and rot at ovary.

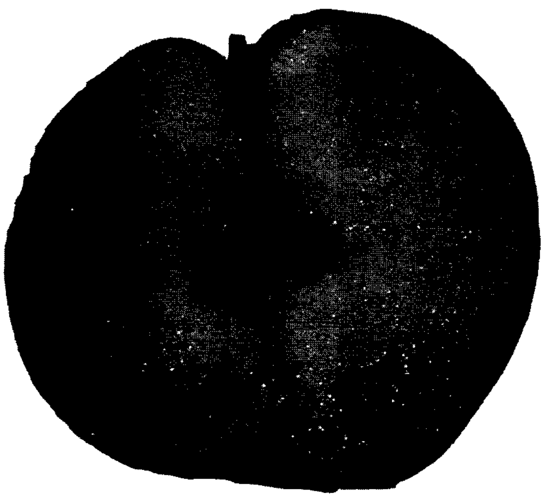


Fig. 2 Spongy disorder at flesh.

Materials and Methods

1. Instruments

To acquire the transmittance spectra of a pear, real-time spectroscopic sensing system was constructed. The system was composed of light source, movable sample trays to hold the fruits, collimating lenses to efficiently collect the transmittance ray, optical fiber, real-time spectrometer and the conveyor lines to transport sample trays. As a controller, an industrial computer (Pentium III) was used.

As a light source, four 200W tungsten-halogen lamps (Osram, Germany) were used. Total power of light source was 800W, which was sufficient for the penetration of light into fruits (Lammertyn et. al. 2000). It was designed that the lights were uniformly exposed to a fruit sample. The sample tray was made of light shield material.

A commercial real-time spectrometer (Ocean Optics, USA) with gratings (wavelength range of 500 to 1,100 nm) was used. The transmitted light through the fruits was transported to the spectrometer by 1,000 μm diameter of optical fiber. The spectrum data was sent from the spectrometer to A/D board (sampling frequency 2 MHz) of computer. Fig. 3 shows the view of on-line experimental apparatus acquiring the transmitted spectra and schematic of its operation.

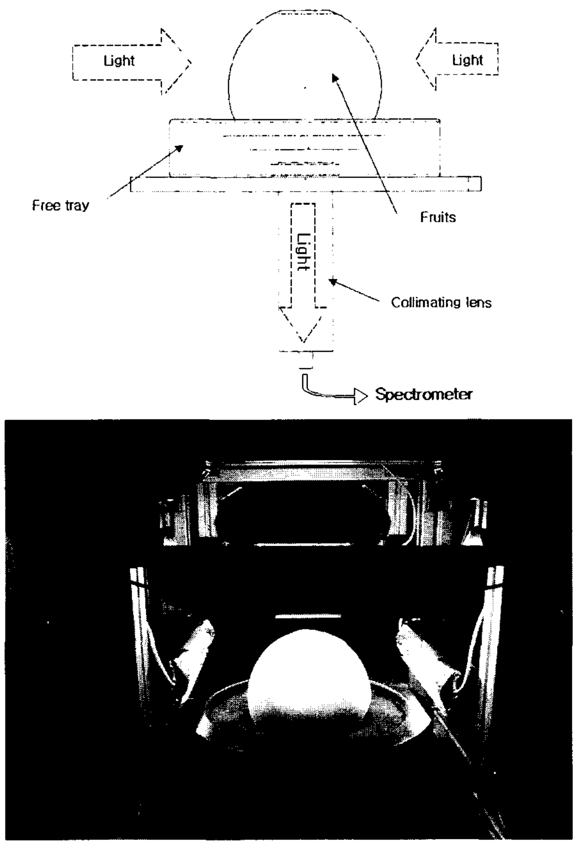


Fig. 3 View of experimental apparatus.

2. Materials

Shingo pear, which is one of famous species of Korean pear, was harvested in 2000 and stored for 6 months. All samples were cut in half after acquiring the spectrum data and examined the internal defects with the naked eye. In practice, the internal defects of pears such as browning, rot, and spongy disorder could not be classified independently because state of the defect was a kind of combination of all three defects.

Therefore, sample groups were divided into 2 classes (class 1: sound and class 2: internal defects). The number of the experimental samples is shown at Table 1.

Table 1. Number of samples (sound and internal defect)

Class	No. of Samples
1 (sound)	114
2 (browning + rot + spongy disorder)	70
Total	184

3. Characteristics of Spectra of Internal Defects

To compare the spectra characteristics of sound pears with defect ones, the standard deviation and coefficient of variation (CV) were used throughout the entire wavelength range of 500 to 1,100 nm. Since standard deviation(std) and variance are absolute values, it is difficult to compare the variation of data having large difference. CV can be used to compare the relative change of spectra considering the mean of each spectrum data set. To investigate the relative characteristics of the transmitted spectra, CV curve was used. CV value at each wavelength was computed using as:

$$CV_i = \frac{\sigma_i}{\bar{x}_i} \tag{1}$$

where σ_i = Standard deviation at the i_{th} wavelength

$\bar{x}_i$ = Mean of sensor intensity (counts) at the i_{th} wavelength

4. Qualitative Analysis

Spectroscopic data have many wavelength points (variables). The qualitative analysis using principal components analysis(PCA) was employed to analyze the spectroscopic data. PCA is useful for reduction of dimension (Jackson, 1991). PCA-MD (PCA-Mahalanobis distance), PCA-LDA (PCA-Linear discriminant analysis), artificial neural network (ANN), k-nearest neighboring (k-NN), and soft independence modeling of class analogy (SIMCA) were generally used to analyze spectroscopic data. Especially SIMCA and k-NN methods classify classes by the similarity among each class and its elements (Tominaga, 1999).

In this paper, SIMCA method was employed to detect the internal defects of pears. This is a kind of supervised discrimination method. Each class was regarded as an independent group and the characteristics of each class was extracted using PCA. The distance of the orthogonal projection among principal component model of each class and unknown data were calculated. Using these distances, the class of the unknown data was determined (Maesschalck, 1999). If spectrum matrix X (m×n, m: number of spectrum, n: wavelength) has multi-classes and the spectrum matrix of class K is X_K, it is represented using Eq.(2) after mean centering by PCA.

$$X_K = \bar{X}_K + T_K P_K' + E_K \tag{2}$$

where X_K =Spectrum matrix($m \times n$) for class K

T_K = Score matrix ($n \times k$) for class K

P_K = Principal component matrix ($k \times n$) for class K

Standard deviation of residuals of class K (s_0^K) was used for the comparison of the residual of the unknown data. This was calculated using Eq.(3).

$$s_0^K = \sqrt{\frac{\sum_{i=1}^{m_K}\sum_{j=1}^{n_K}(e_{ij}^K)^2}{(n_K - k_K)(m_K - k_K - 1)}} \tag{3}$$

where

s_0^K = Standard deviation of residuals of class K

e_{ij}^K = Residual of the i-th data at class K

m_K =No. of samples for class K

n_K =No. of variables for class K

k_K =No. of factors for class K

Under the assumption that residuals of class K follow the normal distribution, F-test was performed to calculate the limit of the Euclidean distance (S_{crit}) among each spectrum data within the PCA model.

$$s_{crit} = \sqrt{F_{crit}(s_0^K)^2} \tag{4}$$

where S_{crit} = Limit of the Euclidean distance among each spectrum data

F_{crit} = F value at confidence limit α

If the unknown spectrum data x^{new} was predicted whether it belongs to class K, the distance of principal component coordinates was calculated using Eq.(5) to Eq.(8).

$$t^{new} = (x^{new} - \bar{x}^K)\, P_k \tag{5}$$

$$\hat{x}^{new} = \bar{x}^K + t^{new}\, P_k \tag{6}$$

$$e_u^{new} = x^{new} - \hat{x}^{new} \tag{7}$$

$$s^{new} = \sqrt{\frac{\sum_{i=1}^{n_K} (e_{ui}^{new})^2}{n_K - k_K}}$$

where

x^{new} = Unknown spectrum

$\bar{x}^K$ = Mean spectrum of class K

$\hat{x}^{new}$ = Reconstructed spectrum of unknown spectrum by PCA analysis

e_u^{new} = Residuals between unknown spectrum and reconstructed spectrum

t^{new} = Score vector of unknown spectrum

s^{new} = Standard deviation of residuals

Whether the unknown data belongs to class K or not was determined by F-test. This F value was calculated using Eq.(9).

$$F^{new} = \frac{(s^{new})^2}{(s_0^K)^2} \tag{9}$$

F^{new} value was used for the similarity test with class K by F-test. And Q and T^2 were used for the consideration of the variation degree within class K (Jackson, 1991). Algorithms were coded with Matlab software (version 5.3, MathWorks, USA) and classification model was developed. Window-based operating software for on-line nondestructive detecting system was programmed with Microsoft Visual C++ 6.0. Fig. 4 is the flowchart of the above procedure.

Correct Classification Rate (CCR) was used for the evaluation of the classification model and was defined as following.

$$\mathrm{CCR}(\%) = \frac{(\text{number of correctly classified samples})}{(\text{number of total samples})} \times 100 \tag{10}$$

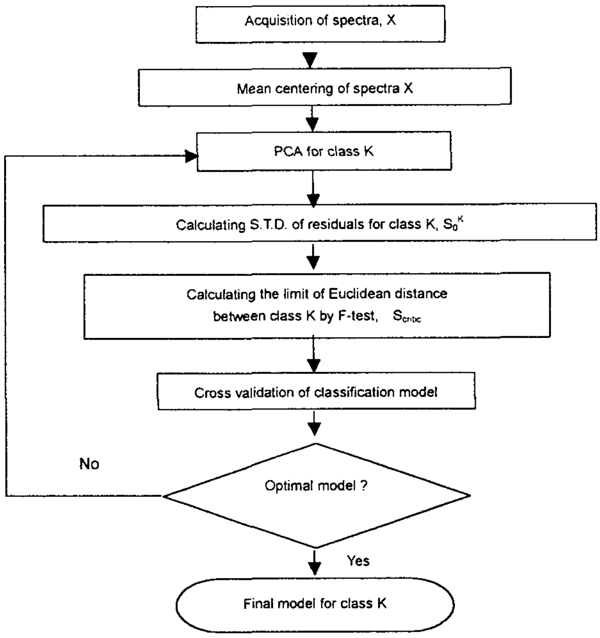


Fig. 4 Flowchart of SIMCA method.

Results and Discussion

1. Characteristics of Internal Defects Spectra

Fig. 5 and Fig. 6 represent the spectra of the sound and defect samples respectively. They were the intensities (counts) of the transmitted light through the pears at wavelength range of 500 to 1,100 nm. Standard deviation and CV spectrum for two data sets are illustrated in Fig. 7. It could be seen that the CV curve of defect samples has more variation than that of sound samples throughout the entire wavelengths.

Wavelength range of 500 to 700 nm is the region of VIS (visible ray) and hence the variation was influenced by the color change of the inside of pears. Wavelength range of 700 to 1,100 nm is the region of NIR (near infrared ray) and change of the internal constituents (water contents, sugar contents, etc.) caused by the internal defects occurred in this region. Though the internal constituent itself was not analyzed and the quantitative change of spectra could not be found, the characteristic of samples with internal defects was apparently different from sound samples (CV curve in Fig. 7). Based on this result, it was concluded that the classification between sound and internal defect samples was possible.

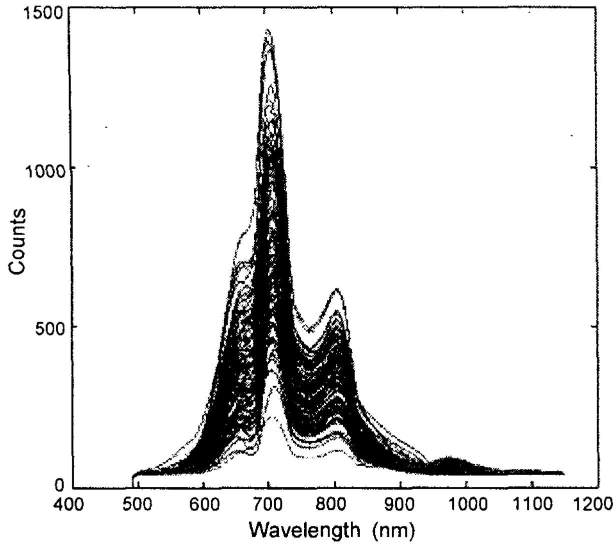


Fig. 5 Transmitted spectra of sound samples.

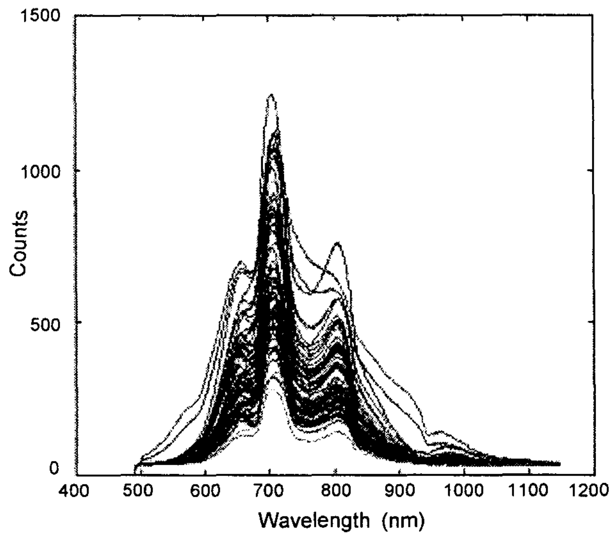


Fig. 6 Transmitted spectra of defect samples.

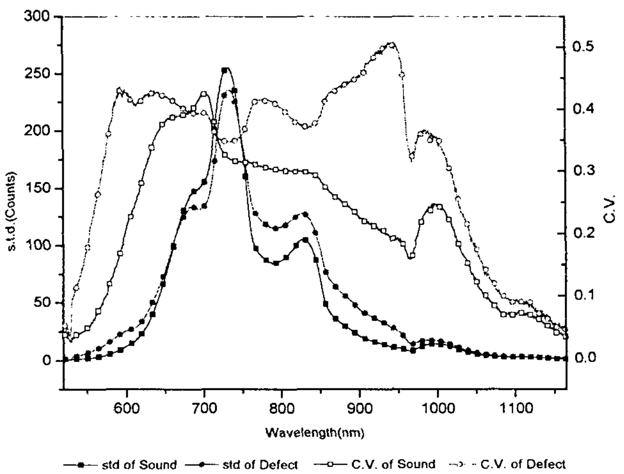
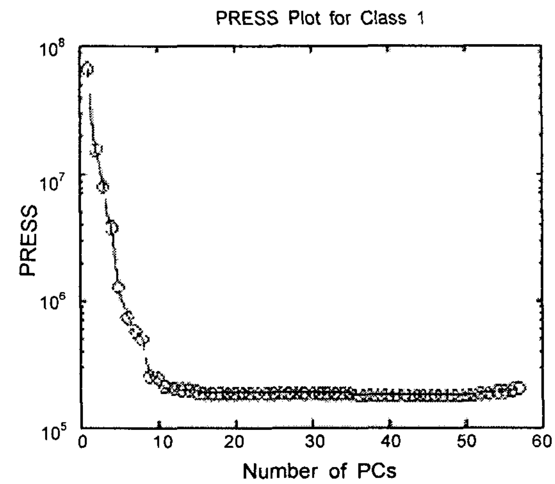


Fig. 7 Characteristics of sound and defect spectra.

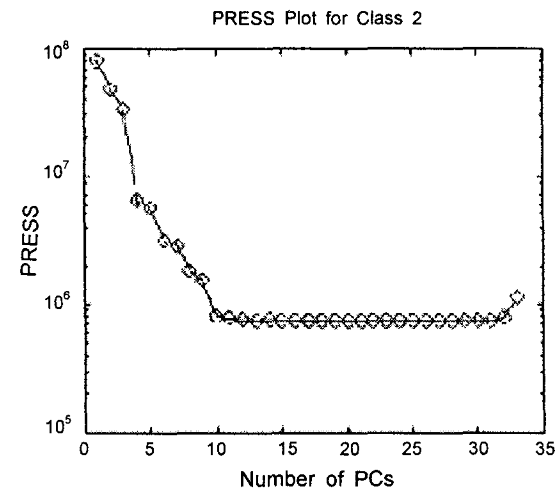
2. Classification of Internal Defects

According to the above results, we could find the non-similarity between sound and defect samples. Total number of 184 samples was used for the classification experiments and was divided into two classes (class 1: sound, class 2: internal defects). First, 92 samples were used to develop the classification model based on SIMCA. And as unknown data, the other samples were used to validate the developed model.

Predicted residual error of sum of square (PRESS) curve was obtained from PCA model for each class (Fig. 8 (a) and (b)). From this curve, the optimal principal component number was determined using the minimum PRESS method. The optimal principal component numbers of class 1 and class 2 were 16 and 11 respectively.



(a)



(b)

Fig. 8 PRESS curve for sound(a) and internal defect(b) samples.

Q-T^2 plot was used to examine the degree of the non similarity between class 1 and class 2 (Fig. 9 (a) and (b)). Limits of Q and T^2 were also calculated within the confidence limit of 95%. These limits were applied to the Q-T^2 plot and it was found that two classes have apparently different characteristics. Table 2 shows the limits of Q and T^2 and numbers of optimal PC.

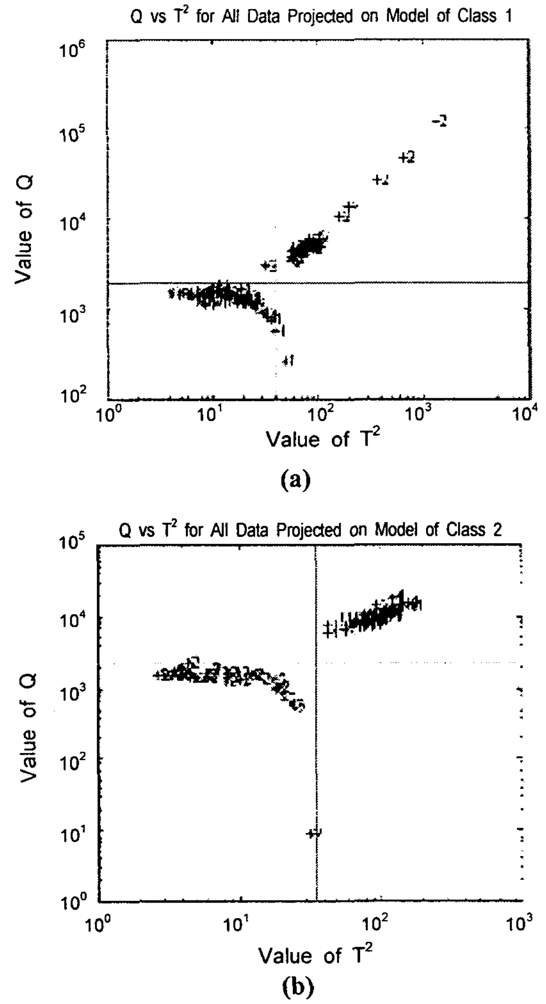


Fig. 9 Q-T^2 plot of sound(a) and internal defect(b) samples.

The developed SIMCA model can classify the samples with internal defects from the sound ones. For 92 unknown samples, 55 samples of 56 sound samples were classified as sound (CCR=98.2%) and 33 of 36 defect samples were classified as defect (CCR=93.9%). Hence, average value of CCR was 96.1% (Table 3).

Conclusions

This research was performed to investigate the capability of detecting the internal defects in agricultural products using VIS/NIR transmittance spectroscopy. Korean pears were harvested and stored in the storage house for about 6 months. The transmittance spectrum acquisition system for pears was constructed. Since the intensity of the spectrum output was dependent on the intensity of light sources, the adequate control of the light was required. Real-time spectrometer having the wavelength range of 500 to 1100 nm (visible and near infrared ray) was adopted.

To investigate the characteristics of the spectrum change between the sound and the defect samples, the curves of standard deviation and coefficient of variation were analyzed. It was concluded that the color change (discoloration) of samples with the internal defect had occurred at around 500 to 700 nm wavelength range and the internal constituents change (water, sugar contents) had occurred at around 700 to 1,100 nm wavelength range. This variation of the internal defects was useful for the development of the classification model.

Among the various methods of the qualitative analysis, the SIMCA method was chosen. With 184 samples, 92 samples were randomly selected to develop the calibration model. Validation was done with other 92 unknown samples. Since the CCR of validation showed the average value of 96.1%, it was concluded that the developed system and model were

Table 2. Statistics of developed SIMCA model

Class	No. of PC	Q_{limit}	T^2_{limit}
1(sound)	16	1890.5	41.1
2(defects)	11	2278.5	35.3

Table 3. Results of model validation of samples

Actual / By model	Sound	Internal defects	Average CCR (%)
Class 1(sound)	55	3	
Class 2(internal defects)	1	33	
CCR (%)	98.2	93.9	96.1

sufficient for commercial use.

Acknowledgments

This research was accomplished by financial supports of Ministry of Agriculture and Forestry, Republic of Korea (1998~2001). We wish to thank to a party concerned in this project.

References

Hwang, I. G. and S. H. Noh. 1999. Preliminary study for development of an algorithm for on-line sugar content of intact fruits using NIR spectroscopy. Proceedings of the 9th international conference on NIR spectroscopy, 379-384.

Hwang, I. G. 2000. Development of on-line apple (Fuji) sorting system by the soluble solid and acid contents using VIS/NIR spectroscopy. Ph.D. thesis of Seoul National University.

Jackson, J. E. 1991. A user's guide to principal components. John-Willey & Sons, Inc.

Lammertyn, J., A. Peris, J. De Baerdemaeker and B. Nicolai. 2000. Light penetration properties of NIR radiation in fruit with respect to non-destructive quality assessment. Postharvest Biology and Technology, 8:121-132.

Maesschalck, R. De and A. Candolfi. 1999. Decision criteria for soft independent modeling of class analogy applied to near infrared data. Chemometrics and Intelligent Laboratory Systems, 47:65-77.

Ryu, D. S. 2001. Development of nondestructive internal quality evaluation system for citrus using VIS/NIR transmittance spectroscopy. Ph.D. thesis of Seoul National University.

Tominaga, Y. 1999. Comparative study of class data analysis with PCA-LDA, SIMCA, PLS, ANNs and k-NN. Chemometrics and Intelligent Laboratory Systems, 49:105-115.