Preliminary study for the development of an algorithm for on-line analysis of the sugar content of intact fruits using near infrared spectroscopy

In-Geun Hwang and Sang-Ha Noh

Department of Agricultural Engineering, Seoul National University,103 Seodun-Dong, Kwonsun-Gu, Suwon 441-744, Korea.

Introduction

In recent years, research has been focused on the development of non-destructive measurement techniques for internal quality attributes of fruit such as firmness, maturity, sugar content, acidity, etc, using near infrared (NIR) spectroscopic methods. Non-destructive NIR methods have been applied to measure the internal quality in peaches and nectarines,[1] raisins[2] and kiwifruit.[3]

Moons *et al.*[4] investigated the possibilities for predicting the total soluble solid content (SSC), pH, titrable acidity and hardness of intact apples in the spectral range of 400 to 2400 nm. Davenel *et al.*[5] predicted sugar content in apples with an error of prediction of 6.8 g L^{-1} of sugar and a correlation coefficient of 0.96. Lammertyn *et al.*[6] established relations between the visible and NIR reflectance spectra(380–1650 nm) of Jonagold apples and quality characteristics such as the pH, SSC, the stiffness factor and other texture parameters and the impact of the data preprocessing on the prediction performance was investigated.

However, most of these experiments mentioned above have been carried out with slow-scan and laboratory-use spectrophotometers. Bellon *et al.*[7] developed an NIR instrument (800–1050 nm) coupled with diode arrays, camera sensor and optical fibres which could be used for on-line sorting of fruit by sugar content at the conditioning stations. Standard error of prediction (*SEP*) was 1.03°Brix. In Japan, two types of on-line sugar content sorters have been developed and commercialised. One is a reflection type which was developed by Mitsui Mining and Smelting Co. Ltd in 1989 and the other is a transmission type by Maki Co. Ltd in 1996. It is known that the latter can measure the acidity as well as sugar content. In Korea, Choi *et al.*[8] developed an on-line sugar sorter with a real-time spectrophotometer which was coupled with a PDA array detector and optical fibre, covering the spectral ranges of 630 to 1100 nm. *SEP* of the on-line sorter is 0.798°Brix at a sorting rate of two apples per second. It has been suggested that compensation for apple temperature should be accounted for in improving sugar content measurement.

The ultimate purpose of this study is to develop a sorting system for on-line sugar content measurement of Fuji apple, oriental pear, etc. with a real-time spectroscopic detector covering the wavelength range of 600 to 1100 nm. The scope of this study is to design a reflection probe coupled to the real-time detector to minimise the effect on the spectrum due to the relative change in distance between the probe and the object and to develop a robust calibration model including pre-processing methods which could reduce the noise effect caused by the fast scanning of the spectra.

Near Infrared Spectroscopy: Proceedings of the 9th International Conference

Materials and methods

Materials

Sugar solutions having different concentrations of fructose (2.5–10.0%), glucose (1.0–6.0%), sucrose (1.0–4.0%) and combined mixtures (4.5–20.0%) were made to investigate the absorption bands of their constituents. A total of 60 Fuji apples were purchased from a packing house in 1999, where the apples were stored for about seven months in cold storage, to test the performance of a new reflection probe designed for on-line purposes. Apple sizes were 79.4 to 106.7 mm in diameter and 47.5 to 95.0 mm in height.

Probe design

The first thing to be accounted for in designing a reflection probe for on-line analysis is to minimise impact of the relative distance change between the probe which is fixed at the side of a line-up conveyer and the fruit which is conveyed on a tray due to differences in fruit size. The second is that the detection area of the probe should be as large as possible so that a representative sugar content of each fruit could be measured since sugar content within a fruit is different from location to location. Third is that the light intensity which reaches the detector through the receptor should be greater than the detection limit.

A reflection probe (Figure 1) was made with eight optical fibre bundles for illumination and nine for collecting the reflected light so that the illumination area by light source bundles increased proportionally with the distance between the probe and the fruit and, also, the view area of each receptor bundle could be confined within the illuminated area. The diameter of the illuminator circle and the receptor ring were determined as 23.2 mm and 48.0 mm, respectively, on the basis of the small size of the Fuji apples. In order to confine the view area of the receptors within the illuminated area, regardless of the apple size, the optimum angle of the receptor fibre bundles to the illumination bundles was estimated to be about 14°Brix.

Spectral measurements

Spectra of sugar solutions were measured with a slow-scan spectrophotometer (NIRsystems 6500) at the range of 600 to 1100 nm (wavelength increment 2.0 nm) in transmission mode. A 10 mm quartz cuvette was used to measure the transmittance spectra. Spectra of Fuji apples were taken with the designed probe and the real-time detector (Figure 1) which consisted of a slit (600 μm), grating (400 l mm^{-1}) and PDA array detector (MS125, Oriel Inc.). Spectra rate was set up at 5 Hz. Each apple was put on the centre of 65 mm-radius round tray which was placed 80 mm apart from the front surface of the probe. In this arrangement the distance between the probe and the equator side of apple varied from about 26 to 41 mm depending on the size of apple samples.

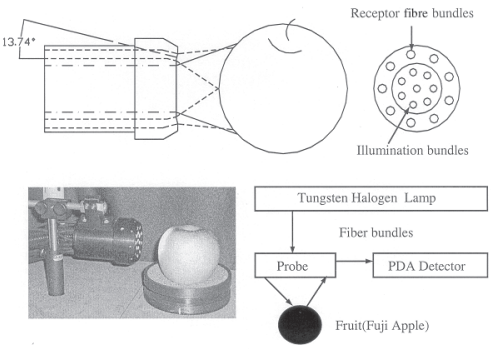

Figure 1. The designed reflection probe and block diagram for real time spectrum measurement.

SSC measurement

The apple juice expressed from a piece of equator side apple where the spectrum was taken was collected directly into a digital refractometer

(PR-1, Atago) to measure the SSC (°Brix) with an accuracy of ± 0.1% in the range from 0.0 to 55.0°Brix.

Processing of the spectroscopic data

To test the influence of the pre-processing on the prediction performance of the calibration models, two types of pre-processing were used. First, the raw spectral data were smoothed by two methods of Hanning window and Moving average, respectively, as expressed in Equations (1) and (2), but the difference between them is to use weight values of one in Moving average and to use weight values similar to Gaussian distribution in Hanning window. The feature of these methods is not to decrease the resolution of the original spectral data.

$$MA_i = \sum_{j=-m}^{m} A_{i+j} \tag{1}$$

$$HA_i = \sum_{j=-m}^{m} W_j A_{i+j} \tag{2}$$

where, MA_i, HA_i : absorbance value at the i'th wavelength smoothed by moving average and Hanning window, respectively
$A_{i=j}$: absorbance at the ($i = j$)'th wavelength
$2m = 1$: number of points in one segment
W_j : Weight value = $0.5*\{1 = \cos[\pi*j/(m = 1)]\}/(m = 1)$

The second type of preprocessing was the use of the second differentiation with different gap size to correct the additive and multiplicative effects in the spectra. Numerical values for the second derivative spectra were computed by using central difference. Principal component regression (PCR) was used to build the prediction models with the pre-processed data together with the SSC data.

Results and discussion

Absorption bands of sugar solution

Figures 2 and 3 present typical absorbance spectra, the second derivative spectra and correlation coefficients between the second derivative values and the concentrations of sugar solution and Fuji apple, respectively. It is noted that the original and the 2nd derivative spectra and correlation curves in both figures are very similar to each other in terms of shape, number of peaks and peak wavelengths. The peaks at 625, 750, 780, 840, 965 nm in the 2nd derivative spectra for sugar solution should be absorption bands of water and those at 775, 840 and 960 nm in apple should also be due to absorption by water, if we consider the fact that water content in sugar solu-

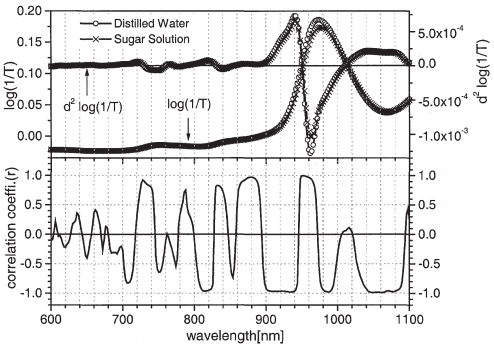

Figure 2. Spectra of log(1/*T*) and d² log(1/*T*) of sugar solution and distilled water, and correlation plot between the 2nd derivatives and sugar concentration(preprocessing : Hanning window, smoothing segment = 13 points(26 nm), gap size = 1 point (2 nm).

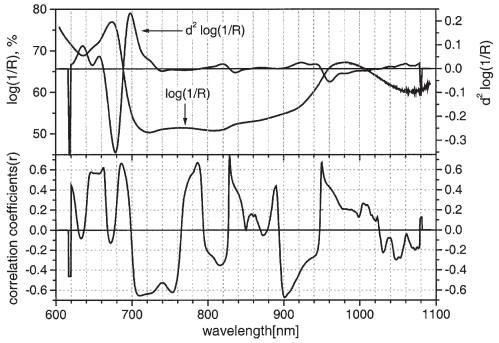

Figure 3. Spectra of log(1/*R*) and d² log(1/*R*) of Fuji apple measured with the probe designed and correlation plot between the 2nd derivatives and SSC [preprocessing : Hanning window, smoothing segment = 55 points (26.2 nm), gap size = 4 points (1.9 nm)].

tion and apple is greater than 80%. These absorption bands of apple coincided well with those of water.

The wavelength bands having high negative correlations with sugar content appeared at 812, 852, 900–942, 988 and 1038–1085 nm for sugar solutions and at 810–825, 902 and 1058 nm for apple in the NIR range, respectively. As shown in Table 1, most of these peaks coincided with absorption bands of carbohydrates reported by Williams and Norris.[9]

Feasibility of on-line (real-time) sugar content measurement

Influence of Pre-processing

SSC measurements of Fuji apple varied between 9.4 and 18.2°Brix. The spectrum and the SSC were measured at four different sides on one apple and preprocessing (Hanning window and Moving average) and statistical analysis (PCR) were performed on 120 data set for calibration and 120 for validation. Figure 4 presented the effect of the gap size for the 2nd derivatives on the *SEP* at various segment size for smoothing. The gap size did affect the *SEP* values in both preprocessing methods of the Hanning window and the Moving average and variation in *SEP,* due to the smoothing size, were relatively smaller when the raw data were smoothed by Hanning window than by the Moving average. From Figure 4, it is also concluded that stabilised *SEP* values could be obtained if the raw data are smoothed by Hanning window of segment

Table 1. Peak bands in absorption spectra, the second derivative spectra and correlation coefficients (between sugar content or SSC and the second derivatives) for liquid sugar and Fuji apple

Items		Sample	Peak bands (nm)		Absorpsorption bands of water and carbohydrates[9]
			600–800	800–1100	
Raw spectra		Sugar Fuji apple	750 675,** 775	850 975** 850 975**	830–840 : H_2O^* & Carb. 870–890 : Carbo.* 900-930 : Carbo.* 958 : H_2O^{**} 970–990 : Carbo.** & H_2O 1010–1030 : Carbo. & H_2O 1053 : Carbo.
2nd derivative spectra		Sugar Fuji apple	625, 750, 780 680**, 775	840 965** 840 960*	
Correlation coefficients	(-)	Sugar Fuji apple	706- 725-- 758-- 792	812-- 852 900-942-- 988-- 1038–1085-- 810–825- 902-- 1058-	
	(+)	Sugar Fuji apple	730++ 789 665+ 690+	832+ 880++ 950++ 835 885+ 950++ 1017+	

*Indicative of the relative strength of an absorber with ** being the strongest
– and + indicative of the relative strength of the correlation with — and ++ being the strongest negative and positive correlation, respectively

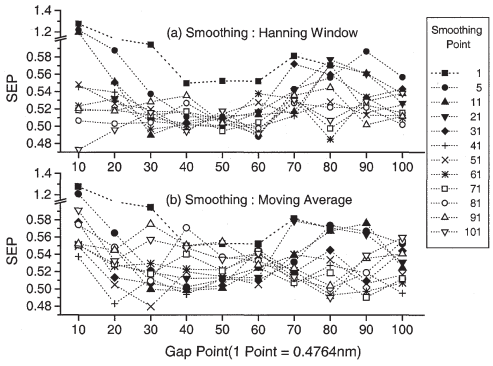

Figure 4. Influence of pre-processing conditions on *SEP* values of PCR model.

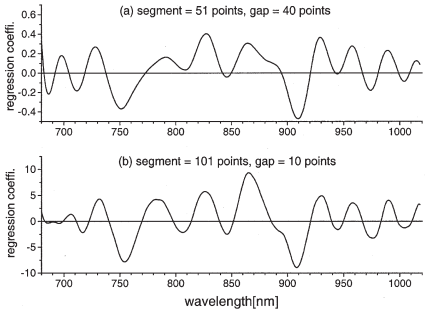

Figure 5. Regression coefficient vectors of PCR models to predict SSC of Fuji apple at different pre-processing conditions(smoothing : Hanning window).

size greater than five points (about 2.4 nm) and then the 2nd derivatives were taken with the gap size range of 30 (14.3 nm) to 100 points (47.6 nm).

Robustness of calibration model

The quality of the calibration is quantified by the *SEC*, *SEP* and *r* between the measured and predicted models. A robust model should have low *SEC* and *SEP* values, a high correlation coefficient, a small difference between the *SEC* and *SEP* and a small number of latent variables.

Calibration models having various pre-processing conditions were compared in Table 2. Models 2, 3 and 4 have relatively low *SEC* and *SEP* values, small differences between them and small number of

Table Table 2. Effect of preprocessing on the prediction performance of the SSC model.

Model	Method	Preprocessing	Latent variables	Calibration		Validations		
				SEC	Corr.(*r*)	*SEP*	Bias	Corr.(*r*)
1	PCR	H-S11-G40*	27	0.438	0.975	0.506	–0.019	0.963
2	PCR	H-S51-G40	11	0.485	0.964	0.496	–0.061	0.965
3	PCR	H-S101-G40	11	0.452	0.968	0.493	–0.011	0.964
4	PCR	H-S101-G10	14	0.459	0.968	0.473	–0.034	0.967
5	PCR	H-S101-G90	9	0.499	0.961	0.529	–0.030	0.958
6	PCR	M-S11-G40	21	0.451	0.972	0.492	0.004	0.964
7	PCR	M-S51-G40	19	0.433	0.973	0.489	–0.011	0.965
8	PCR	M-S101-G40	23	0.430	0.975	0.520	0.006	0.961
9	PCR	M-S101-G10	17	0.479	0.966	0.505	0.001	0.962
10	PCR	M-S101-G90	22	0.460	0.971	0.536	–0.015	0.957

* H, M, S and G stand for Hanning window, Moving average, smoothing size and gap size, respectively. Numbers followed by S or G are number of smoothing or gap points. One point is equivalent to 0.4764 nm in wavelength

latent variables. Also, the regression coefficient vectors (Figure 5) which are involved in the models 2 and 4 indicate that the peak points and their corresponding wavelengths almost coincided in both models and, furthermore, most of the peak wavelength bands also appear near the peak bands at which high correlations between the concentrations of sugar solution and the 2nd derivatives exist (Figure 2 and Table 1). As a result, it is concluded that these models have robustness and the probe, coupled with the real time PDA detector, could be used for on-line sugar content measurement.

Conclusions

To develop an on-line sugar content measurement system for Fuji apple, oriental pear, etc. with non-destructive NIR techniques a reflection probe, which was coupled to a commercialised real-time spectroscopic detector (600–1100 nm), was designed as a preliminary step. The performance of the probe was satisfactory since robust PCR models having *SEP* values in the range of 0.47 to 0.57 could be obtained at various pre-processing conditions of the raw spectra, even though the difference between the *SEC* and *SEP* values and the number of latent variables, which indicate robustness of model, were influenced by the pre-processing conditions. Hanning window indicated relatively lower values and smaller variations in *SEP* than the Moving average at various gap size levels for the 2nd differentiation. A robust model, having *SEP* of 0.473 and 14 latent variables, was obtained at the pre-processing condition of segment size, 101 points (48.1 nm, Hanning window) and the gap size, 10 points (4.8 nm). The coefficient vector involved in this model could be well interpreted with the absorption bands of liquid sugar.

Further work is required to check the performance of the prediction models with Fuji apples including other fruit at various conveying speeds.

References

1. D.C. Slaughter, *Transactions of the ASAE* **38(2),** (1995).
2. C.C. Huxoll, H.R. Bolin and B.E. Mackey, *J. Food Sci.* **60(1),** (1995).
3. V.A. McGlone and S. Kawano, *J. Postharvest Biology and Technology* **13,** (1998).
4. E.A. Moons, P. Dardenne and M. Sindic, in *Proc., Sensors for Non-destructive Testing International Conf.: Nondestructive Visible and NIR Spectroscopy for the Determination of Internal Quality in Apple*, Orlando, Florida, February. Ithaca, p.112 (1997).
5. A. Davenel, M. Crochon, J. Pourcin, P. Verlawue and M.C. Devaux, *Presentation at Agricultural Engineering International Conference : On line infrared measurements in food processing.* p.7 (1988).
6. J. Lammertyn, B. Nicolai, K. Ooms, V. De Smedt and J. De Baerdemaeker*, Transactions of the ASAE* **41(4),** (1998).
7. V. Bellon, J.L. Vigneau and M. Leclercq, *Appl. Spectrosc.* **47(7),** 1079 (1993).
8. K.H. Choi and K.J. Lee*, International Seminar on Technology Development for Nondestructive Quality Evaluation of Agricultural Products : Nondestructive technology to measure sugar content in fruits using NIR spectroscopy.* Agricultural Mechanisation Research Institute, Suwon, Korea, p.91 (1998).
9. P. Williams and K. Norris, *Qualitative applications of near-infrared reflectance spectroscopy in the Near-Infrared Technology in the Agricultural and Food Industries*, American Association of Cereal Chemistry, St. Paul, MN, USA, p.241 (1987)