Nondestructive Detection of Internal Flavor in ‘Shatian’ Pomelo Fruit Based on Visible/Near Infrared Spectroscopy

Sai Xu

Institute of Quality Standard and Monitoring Technology for Agro-products of Guangdong Academy of Agricultural Sciences, Guangzhou 510640, China

Huazhong Lu

Guangdong Academy of Agricultural Sciences, Guangzhou 510640, China

Xu Wang

Institute of Quality Standard and Monitoring Technology for Agro-products of Guangdong Academy of Agricultural Sciences, Guangzhou 510640, China

Christopher M. Ference

U.S. Department of Agriculture, Forest Service, Forest Health Protection, Alexandria Field Office, 2500 Shreveport Highway, Pineville, LA 71360

Xin Liang and Guangjun Qiu

Institute of Quality Standard and Monitoring Technology for Agro-products of Guangdong Academy of Agricultural Sciences, Guangzhou 510640, China

Additional index words. acidity, intelligent detection, large fruit, modeling, soluble solid

Abstract. Visible/near-infrared (VIS/NIR) spectroscopy is a powerful tool for rapid, nondestructive fruit quality detection. This technology has been widely applied for quality detection of small thin-peel fruit, although less so for large thick-peel fruit because of the low signal-to-noise ratio of the spectral signal, resulting in a reduction of accuracy. More modeling work should be focused on solving this problem. This research explored a method of spectroscopy for the total soluble solid (TSS) content and acidity detection of ‘Shatian’ pomelo, which are two major parameters of fruit internal flavor. VIS/NIR spectral signal detection of 100 pomelo samples during storage was performed. Detection based on raw data, signal jitter, and scattered light noise removal, feature extraction, and deep learning were performed and combined with modeling detection to achieve an accurate step-by-step detection. Our results showed that 600 W is the optimal light intensity for detecting the internal flavor of pomelo. The TSS content of pomelo is optimally detected using Savitzky-Golay (SG) + multiplicative scatter correction (MSC) + genetic algorithm (GA) + principal component analysis (PCA) + convolutional neural network (CNN) + partial least squares regression (PLSR); however, acidity of pomelo is optimally detected using SG + MSC + GA + PLSR. With the optimal detection method, the coefficient of determination and root mean squared error (RMSE) of the validation set for TSS detection are 0.72 and 0.49, respectively; and for acidity detection are 0.55 and 0.10, respectively. Even though the accuracy is not high, the data are still acceptable and helpful in non-destructive quality grading of large quantities postharvest fruit. Therefore, our results demonstrated that VIS/NIR was feasible for detecting the TSS content and acidity of postharvest pomelo, and for providing a possible method for the nondestructive internal quality detection of other large thick-peel fruit.

Internal flavor is one of the most important factors affecting the commodity value of fruit, mainly involving TSS content and acidity (Wei et al., 2021). However, fruit internal flavor can be uneven as a result of being affected by a series of environmental parameters such as sunlight, temperature, water, fertilizer, and so on (Li et al., 2015). Also, internal quality is difficult to detect precisely based on visual appearance only. Thus, an efficient method for internal flavor detection is needed for the fruit industry.

Traditionally, fruit internal flavor is detected manually (Qiu and Wang, 2015) and by assay analysis (Cavalcante and Martins, 2005) methods. However, both of those methods are labor intensive and time-consuming, are only suitable for spot checking, and cannot meet the need for flavor detection for an entire harvest of fruit from an orchard. An intelligent method—machine vision technology (Gongal et al., 2018; Naik and Patel, 2017)—has been applied in the field of fruit quality detection for a long time. Although

machine vision saves labor and improves efficiency, it can only measure fruit appearance characteristics such as size, color, shape, volume, and so on, which does not provide enough related information to detect internal fruit flavor accurately. The electronic nose (Baietto and Wilson, 2015; Qiu and Wang, 2017) is an intelligent method for fruit quality detection based on volatile characteristics. However, its internal quality detection ability is usually unsatisfactory because the internal volatiles acquired are much weaker than the external volatiles when performing nondestructive detection.

VIS/NIR spectroscopy has been applied increasingly to the internal flavor detection of fruit, such as apple (*Malus domestica*) (Mehinagic et al., 2003), orange (*Citrus sinensis*) (Song et al., 2020), peach (*Amygdalus persica* L.) (Cortés et al., 2017), and kiwi (*Actinidia chinensis* Planch) (Ma et al., 2021), and is often combined with a detection model for fast and accurate measurement. However, previous research focused mainly on small thin-peel fruit, with less attention paid to large thick-peel fruit. This is because information from small thin-peel fruit can be acquired more easily using either reflected or transmitted light, whereas the same information is difficult to obtain from large thick-peel fruit using reflected light and has a worse signal-to-noise ratio when using transmitted light. Thus, it is important to focus on modeling work to help improve the accuracy of large thick-peel fruit internal flavor detection, which has not been investigated as widely.

As a thick-peel fruit, it is difficult to determine the internal flavor of ‘Shatian’ pomelo [*Citrus maxima* (Burm) Merr.] using nondestructive detection based on VIS/NIR spectroscopy. ‘Shatian’ pomelo is larger than most other grapefruit pomelos, with a diameter of more than 15 cm, a peel thickness of ≈ 1.5 cm, and flesh wrapped by the mesocarp. The complex internal structure and large size increase the differences among samples. Thus, previous research reported that although VIS/NIR was feasible for the detection of the internal flavor of grapefruit (Li et al., 2019; Ncama et al., 2017), it was not feasible for detecting the internal flavor of ‘Shatian’ pomelo. Because of the difficulty of peeling, knowing the internal flavor of pomelo before peeling becomes more important.

To find a feasible method for ‘Shatian’ pomelo internal flavor detection, in this study we designed a VIS/NIR transmitted spectroscopy platform and applied it to pomelo fruit sampling, and optimized procedures for data preprocessing, feature extraction, and deep learning to improve the detection effect of pomelo internal flavor (TSS content and acidity). The research results provide a fast, intelligent, and nondestructive detection method for internal flavor quality detection of pomelo.

Materials and Methods

Pomelo fruit samples

All pomelo fruit samples were harvested from an orchard in Meizhou City, Guangdong

Fig. 1. Structure of the laboratory-developed visible/near-infrared sampling platform.

Province, China, and then shipped to our laboratory for detection in Guangzhou City within 24 h. There were 100 ‘Shatian’ pomelo fruit samples tested in total. The average $\pm$ SD pomelo transverse diameter, vertical diameter, and peel thickness were 18.10 ± 1.80 , 19.71 ± 2.50 , and 1.46 ± 0.21 cm.

VIS/NIR spectrum sampling

Our laboratory developed a VIS/NIR spectrum sampling platform, as shown in Fig. 1. To reduce external light, pomelo samples were measured in a dark box. The arc-shaped light set is on the right side, which includes eight 100-W halogen lamps. In consideration of the practical needs of an assembly line detection, a movable tray was applied to convey and stabilize each tested pomelo. The spectrum signal was transmitted through the pomelo from the right side to the left, was received by an optical fiber, and was translated into a digital signal using two spectrometers (QE PRO with wavelengths of 400–1100 nm and NIR QUEST with wavelengths of 900–1700 nm; Ocean Optics Inc., Dunedin, FL). Thus, the transmitted spectrum wavelengths between 400 and 1700 nm could be tested and recorded. For sampling, the pomelo fruit were placed on the tray, with spectrum signal wavelengths of 400 to 1050 and 1050 to 1700 nm acquired by the QE PRO and the NIR QUEST spectrometers, respectively. The presampling process was 1) save the dark current value D, 2) offset the dark current value (D of NIR QUEST plus the

difference between D from the QE PRO and D from the NIR QUEST at 1050 nm), and 3) save the reference value R (3.6-cm-thick spectral calibrated panel made of barium sulfate material). Last, with the pomelo sampling detector response value (P), the pomelo transmissivity was equal to $(P - D)/(R - D)$.

After repeated adjustment, the optimal distance from the light set to the pomelo was set to 25 cm, and the optimal distance from the pomelo sample to the receiving fiber was set as 2 cm (the shorter distance can avoid stray light more efficiently), the optimal integral time of the 400 to 1050- and 1050 to 1700-nm wavelengths was set as 300 and 1000 ms, respectively. To maximize the sampling signal, there were no other attachments between the receiving fiber and the sample, such as integrating sphere and so on. The spectral signal would be too weak if the light intensity (the number of working lights) was not enough; however, the pomelo peel would be burned if the light intensity was too strong. To guarantee the symmetry of the arc-shaped light, light intensities of 200, 400, 600, and 800 W were tested by turning on lights sequentially from the middle to the sides. We found that 200 W returns a relatively weak spectral signal whereas 800 W results in the pomelo peel burning within 1 s. Both 400 and 600 W returned enough spectral signal transmissivity and did not burn the pomelo peel within a few seconds. Further comparison is needed to find the optimal light intensity using more samples. The raw spectral signals of 100 pomelo samples tested under 400 and 600 W are shown in Fig. 2A and B, respectively. The raw data showed that 600 W had a better signal-to-noise ratio than 400 W, a higher transmissivity, and a stronger clustering performance. Further research is needed to ensure the optimal light intensity for modeling detection.

Internal flavor assessment

TSS and acidity (citric acid) assessment was conducted after VIS/NIR spectrum acquisition by using a digital pocket refractometer (PAL-BX/ACID1; ATAGO Co. Ltd., Tokyo, Japan). For TSS assessment, pomelo samples were peeled to get fruit flesh, which was then crushed and homogenized, and the juice was filtered through gauze. Two drops of juices were taken to measure the TSS content directly. For acidity assessment, 1 mL juice was put in a glass beaker and diluted

with 50 mL distilled water, after which two drops of the diluted juice were used for the acidity measurement. Each sample was measured three times, and the TSS content and acidity for that sample was recorded as the average of these three values. Between each measurement, the refractometer was calibrated with distilled water. The average $\pm$ SD of TSS content and acidity were $15.48 \pm 1.16\%$ and 1.25 ± 0.14 g/L, respectively. The TSS content and acidity of 100 pomelo samples had a relatively symmetric distribution, as is shown in Fig. 3A and B, respectively.

Data analysis method

In our study, PLSR was applied to build up the detection model to check the internal flavor detection ability of spectral data. SG was applied to remove the signal jitter in the spectral curve. MSC was applied to remove the scattered light noise in the spectral curve. GA was applied to select useful features in the spectral curve and remove redundancy. CNN was applied to conduct deep learning feature extraction of the spectral curve. PCA was applied for the dimension reduction of VIS/NIR Spectrum data for the CNN input.

Partial least squares regression. PLSR (Shetty and Gislum, 2011) is a multivariate data analysis technique that generalizes and combines features from PCA and multiple linear regression. It is useful in predicting a set of dependent variables from a large set of independent collinear variables. PLSR has been particularly successful in developing multivariate calibration models, as it uses the concentration information (y) in determining how regression factors are computed from the spectral data matrix (X), thereby reducing the impact of irrelevant X variations in the calibration model. Therefore, PLSR is ideal for multivariate calibration of spectroscopic data. The coefficient of determination (R^2) is the key parameter for evaluating the correlation between the predicted value and the actual value. The range of R^2 is from zero to one, where a greater R^2 value equals better predictive ability (a stronger relationship between the predicted value and the actual value). In addition, RMSE is another method for detection method evaluation; the closer the RSME value is to zero, the better the method’s prediction (a smaller error between the predicted value and the actual value).

Savitzky-Golay. Calculating derivatives of spectral data by the SG (Zimmermann and Kohler, 2013) numerical algorithm is often used as a preliminary preprocessing step to resolve overlapping signals, enhance signal properties, and suppress unwanted spectral features that arise from nonideal instrument and sample properties. The key parameters of SG contain the values of order (the order of polynomial for smoothing) and frame length (length of the smooth of each step). The parameter setting is decided by repeated testing.

Multiplicative scatter correction. MSC (Chen et al., 2015) is typically used to compensate for light scattering effects and changes in path length. MSC minimizes these

Received for publication 14 July 2021. Accepted for publication 18 July 2021.

Published online 21 October 2021.

This research was supported by the Research and Development Program in Key Areas of Guangdong Province (grant no. 2018B0202240001), the National Natural Science Foundation of China (grant no. 31901404), the Guangzhou Science and Technology Planning Program (grant no. 201904010199), the New Developing Subject Construction Program of Guangdong Academy of Agricultural Science (grant no. 202134T), the Presidential Foundation of Guangdong Academy of Agricultural Science (grant no. 201920), the Presidential Foundation of Guangdong Academy of Agricultural Science (grant no. 202034), and the Jinying Talent Training Program of Guangdong Academy of Agricultural Science (R2020PY-JX020).

H.L. is the corresponding author. E-mail: huazlu@scau.edu.cn.

This is an open access article distributed under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

Fig. 2. Raw spectrum of 100 pomelo samples (A) under 400 W and (B) under 600 W.

deviations by fitting a linear model between a reference spectrum and other spectra of the dataset using the linear least squares method. The reference spectrum is often chosen as the average of all spectra in the dataset. Following the application of MSC, all spectra appear to have the same absorbance level.

Genetic algorithm. GA (Jarvis and Goodacre, 2005) is a randomized search algorithm that has been developed in an effort to imitate the mechanics of natural selection and natural genetics. GA is commonly used to generate high-quality solutions to optimization, and to search problems by relying on biologically inspired operators such as mutation, crossover, and selection. Thus, for GA operation, the initial population number, the crossover probability, the mutation probability, and the number of iterations are all key parameters that influence the feature selection method.

Convolutional neural networks. CNN (Chen et al., 2016) has a strong ability for feature mining via convolution and pooling operations. The sparse connectivity and weight-sharing of CNN allow for a better extraction of the abstract features and a certain degree of training parameter number simplifications. The infrastructure of a typical CNN contains an input layer, a convolution layer, a pooling layer, and a full connected layer. The input layer must be a matrix to satisfy the requirements of the convolution and pooling operations. The convolutional layers apply a convolution operation to the input while using feature maps, and then pass the result to the next layer. Each feature connects with the input layer by weight/filter. The pooling layers reduce the size of the data by combining the outputs of neuron clusters at one layer into a single neuron in the next layer. The fully connected layers connect every neuron in one layer to every neuron in another layer, and they are used for the rasterization of data after pooling. In accordance with previous research (Jian et al., 2018),

feature vector (X) was converted to feature matrix (XX^T) to fit the requirements of CNN input. PCA (Lichtert and Verbeeck, 2013) was applied for the dimension reduction of VIS/NIR spectrum data. After repeated testing, it was determined that too many or too few principal components decreased the detection accuracy, and that using only the first 30 principal components was the appropriate method for further analysis.

Software. All data analyses were performed using Matlab R2017a software (MathWorks Inc., Natick, MA).

Results and Discussion

Detection based on raw data. The PLSR detection ability of pomelo TSS content based on raw spectral data under different light intensities (400 and 600 W) is shown in Fig. 4A and B, respectively. The R^2 value of TSS detection under 600 W was 0.45, which was better than 400 W. The PLSR detection ability of pomelo acidity based on raw spectral data under different light intensities (400 and 600 W) is shown in Fig. 1C and D, respectively. The R^2 value of acidity detection under 600 W was 0.34, which was better than 400 W. Thus, the optimal light intensity of internal flavor of pomelo detection based on spectroscopy was 600 W. The internal flavor of small thin-peel fruit can usually be accurately detected directly based on raw spectral data (Nicolai et al., 2007) because of the favorable signal-to-noise ratio. However, for large thick fruit such as pomelo, both TSS and titratable acidity detection results proved that a certain amount of TSS-/acidity-related information can be gleaned from the raw spectral data, but further spectral data optimization is needed as a result of the still unsatisfactory detection ability.

SG and MSC for data preprocessing. To acquire data from a larger range of wavelengths of pomelo spectral signals, QE PRO

and NIR QUEST were used jointly to cover a range of 400 to 1700 nm. The boundary between the QE PRO spectrometer signal and the NIR QUEST spectrometer signal was set at 1050 nm, where an obvious signal jitter in raw spectral sampling data was seen (Fig. 5). Nevertheless, the spectral curve seems relatively smooth (Fig. 2), although some small and unobvious signal jitters are unavoidable for sensor technology. To remove these signal jitters, third-order 11-point SG was applied, which had a strong ability to remove the signal jitter at the boundary and to adjust the data curve finely at other places (Fig. 5). In addition, based on transmitted spectroscopy, scattered light is unavoidable, especially for large fruit detection, because the light has more divergent space. Thus, MSC was applied after SG processing to eliminate the spectrum translation caused by scattered light.

The TSS and acidity detection ability of pomelo after SG and MSC data preprocessing are shown in Fig. 6A and B, respectively. The TSS and acidity detection ability were both improved. The R^2 values of the validation set of TSS and acidity PLSR detection were 0.55 and 0.44, respectively. Further detection ability improvement is still needed.

GA for feature extraction. The full spectral data usually contain both junk data and useful data (Goormaghtigh et al., 2009). To remove junk data and keep useful data as precisely as possible, GA was applied for feature extraction. After repeated testing, the parameters of GA for the pomelo spectral data in our study were an initial population number of 60, a crossover probability of 0.5, a mutation probability equal to 0.01, and 120 iterations. A total of 337 features were extracted for TSS detection modeling; 568 features were extracted for acidity detection modeling. The feature extraction results of TSS and acidity detection are shown in Fig. 7A and C, respectively, and their PLSR detection results are shown in Fig. 7B and D, respectively. After

Fig. 3. Flavor parameter distribution of 100 pomelo samples. (A) Total soluble solid (TSS) content. (B) Acidity.

Fig. 4. Flavor parameter detection of pomelo by raw data and partial least squares regression. (A, B) Total soluble solid (TSS) content and (C, D) acidity under (A, C) 400 W and (B, D) 600 W. RMSE = root mean squared error.

Fig. 5. Signal jitter removal at 1050 nm of the raw spectral data. SG = Savitzky-Golay.

Fig. 6. Flavor parameter detection of pomelo by Savitzky-Golay, multiplicative scatter correction, and partial least squares regression. (A) Total soluble solid (TSS) content. (B) Acidity. RMSE = root mean squared error.

feature extraction, the R^2 value of the validation set of TSS and acidity PLSR detection were improved to 0.63 and 0.55, respectively. Thus, it is feasible to use spectroscopy to detect the TSS and acidity of pomelo. However, deep learning may help in mining more useful features to improve accuracy further.

Deep learning detection. Deep learning was applied to improve the detection accuracy of TSS content and acidity of pomelo

further. A deep learning method (CNN) was applied for a more exhaustive examination of SG + MSC + GA processed data. Because CNN has a high computation burden, after reference to previous research (Liu et al., 2010), PCA was applied to reduce the data dimension before conducting CNN, and the first 20 PCA features were extracted and transferred to matrix format (from X to XX^T) for CNN operation for deep mining of

features. After repeated runs, the optimal feature extraction network parameters of CNN for TSS and acidity detection were determined, and they are shown in Table 1.

Pomelo TSS and acidity deep learning detection results (SG + MSC + GA + PCA + CNN + PLSR) are shown in Fig. 8. The R^2 value of the validation set for TSS detection increased from 0.63 to 0.72. However, the R^2 value of the validation set for acidity detection

Fig. 7. Flavor parameter detection of pomelo by Savitzky-Golay, multiplicative scatter correction, genetic algorithm (GA), and partial least squares regression. (A, B) Total soluble solid (TSS) content. (C, D) Acidity. (A, C) GA feature selection. (B, D) Detection result. RMSE = root mean squared error.

Table 1. Parameter setting for convolutional neural network (CNN) feature extraction.

Table 1: CNN network structure for each chemical constituent (CNN) based on dataset								
	CNN network structure						Learning rate	Iterations
	Layer 1			Layer 2				
	FMN	FMS	PS	FMN	FMS	PS		
TSS	6	3	2	4	2	2	0.015	300
Acidity	7	3	2	3	2	2	0.013	400

FMN = number of feature maps; FMS = feature map size; PS = pooling layer size; TSS = total soluble solids content.

decreased from 0.55 to 0.03. Previous research also found that deep learning can improve detection ability, but also has the risk of decreasing detection accuracy (Xu et al., 2019). Mostly, it depends on how much useful information is contained in the database. Both useful and junk features could be increased/enlarged by deep learning. Thus, the TSS content of pomelo is optimally detected by SG + MSC + GA + PCA + CNN + PLSR; however, the acidity of pomelo is optimally detected by SG + MSC + GA + PLSR. Although the accuracy is not high, the data are still acceptable and helpful for nondestructive quality grading of large quantities of postharvest fruit. In addition, TSS content and acidity (composed primarily of citric acid) can be detected by

spectroscopy because VIS/NIR is affected by the stretched vibration overtones and combination modes of hydrogen-containing groups (X-H) including O-H, N-H, C-H, and S-H (Ozaki, 2012). The reason that spectroscopy-based detection ability for pomelo TSS content is better than for acidity is because the content of TSS is much greater than that for acidity.

Conclusion

The applicability of VIS/NIR spectroscopy for the nondestructive detection of internal flavor (TSS content and acidity) of postharvest pomelo was studied in our research. The results showed that 600 W was the optimal light intensity for detecting the

internal flavor of pomelo. Detection based on spectroscopy SG was applied to remove the signal jitter noise, particularly at 1050 nm, of the pomelo spectral curve. MSC was applied to remove the scattered light noise in the spectral curve. Based on raw data, the R^2 value of PLSR detection results for pomelo TSS content and acidity were 0.45 and 0.34, respectively. After SG and GA processing, the R^2 value of PLSR detection of pomelo TSS content and acidity increased to 0.55 and 0.44, respectively. Then, GA was applied for feature extraction. The R^2 value of SG + MSC + GA + PLSR detection results of TSS content and acidity of pomelo increased to 0.63 and 0.55, respectively. Deep learning was applied for further mining of more useful features and to improve detection accuracy. The R^2 value of SG + MSC + GA + CNN + PCA + PLSR detection results of TSS content increased to 0.72. However, the R^2 value of SG + MSC + GA + CNN + PCA + PLSR detection results of acidity decreased to 0.03. Thus, the TSS content of pomelo is optimally detected by SG + MSC + GA + PCA + CNN + PLSR; however, the acidity of pomelo is optimally detected by SG + MSC + GA + PLSR.

Fig. 8. Flavor parameters detection of pomelo by Savitzky-Golay, multiplicative scatter correction, genetic algorithm, convolutional neural network, principal component analysis, and partial least squares regression. (A) Total soluble solid (TSS) content. (B) Acidity. RMSE = root mean squared error.

Although the accuracy is not high, the data are still acceptable and helpful for the nondestructive quality grading of large quantities of postharvest fruit. The nondestructive detection method developed in our study is expected to be used to develop a real-time spectral detection device for TSS content and acidity grading of stored pomelo before marketing, and also to provide reference to internal flavor detection of other large thick-peel fruits.

Literature Cited

- Baietto, M. and A.D. Wilson. 2015. Electronic-nose applications for fruit identification, ripeness and quality grading. *Sensors (Basel)* 15:899–931, doi: 10.3390/s150100899.
- Cavalcante, Í.H.L. and A.B.G. Martins. 2005. Physical and chemical characterization of dovalis fruits. *Intl. J. Fruit Sci.* 5:39–46, doi: 10.1300/J492v05n04_05.
- Chen, H., Q. Song, G. Tang, Q. Feng, and L. Liang. 2015. The combined optimization of Savitzky-Golay smoothing and multiplicative scatter correction for FT-NIR PLS models. *ISRN Spectrosc.* 2013:1–9, doi: 10.1155/2013/642190.
- Chen, Y., H. Jiang, C. Li, X. Jia, and P. Ghamisi. 2016. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Trans. Geosci. Remote Sens.* 54:6232–6251, doi: 10.1109/TGRS.2016.2584107.
- Cortés, V., S. Cubero, N. Aleixos, J. Blasco, and P. Talens. 2017. Sweet and nonsweet taste discrimination of nectarines using visible and near-infrared spectroscopy. *Postharvest Biol. Technol.* 133: 113–120, doi: 10.1016/j.postharvbio.2017.07.015.
- Gongal, A., M. Karkee, and S. Amatya. 2018. Apple fruit size estimation using a 3D machine vision system. *Info. Process. Agr.* 5:498–503, doi: 10.1016/j.inpa.2018.06.002.
- Goormaghtigh, E., R. Gasper, A. Bénard, A. Goldsztein, and V. Raussens. 2009. Protein secondary structure content in solution, films and tissues: Redundancy and complementarity of the information content in circular dichroism, transmission and ATR FTIR spectra. *Biochim. Biophys. Acta Proteins Proteomics* 1794:1332–1343, doi: 10.1016/j.bbapap.2009.06.007.
- Jarvis, R.M. and R. Goodacre. 2005. Genetic algorithm optimization for pre-processing and variable selection of spectroscopic data. *Bioinformatics* 21:860–868, doi: 10.1093/bioinformatics/bti102.
- Jian, D., B. Hu, Y. Liu, C. Wei, G. Zhang, X. Tang, and B. Hu. 2018. Study on quality identification of macadamia nut based on convolutional neural networks and spectral features. *Spectrosc. Spectral Anal.* 38:1514–1519, doi: 10.3964/j.issn.1000-0593(2018)05-1514-06.
- Li, X., Y. Liu, A. Ouyang, X. Sun, J. Hu, X. Jiang, and Y. Ouyang. 2019. A general model for judging and brix detection of grapefruit variety based on near infrared. *Chinese J. Luminescence* 40:808–814, doi: 10.3788/fjxb20194006.0808.
- Li, Y., M.Q. Han, F. Lin, Y. Ten, J. Lin, D.H. Zhu, P. Guo, Y.B. Weng, and L.S. Chen. 2015. Soil chemical properties, ‘Guanximiyou’ pummelo leaf mineral nutrient status and fruit quality in the southern region of Fujian province, China. *J. Soil Sci. Plant Nutr.* 15:263–269, doi: 10.4067/S0718-95162015005000029.
- Lichtert, S. and J. Verbeeck. 2013. Statistical consequences of applying a PCA noise filter on EELS spectrum images. *Ultramicroscopy* 125:35–42, doi: 10.1016/j.ultramic.2012.10.001.
- Liu, Y., X. Sun, and A. Ouyang. 2010. Nondestructive measurement of soluble solid content of navel orange fruit by visible-NIR spectrometric technique with PLSR and PCA-BPNN. *Lebensm. Wiss. Technol.* 43:602–607, doi: 10.1016/j.lwt.2009.10.008.
- Ma, T., Y. Xia, T. Inagaki, and S. Tsuchikawa. 2021. Non-destructive and fast method of mapping the distribution of the soluble solids content and pH in kiwifruit using object rotation near-infrared hyperspectral imaging approach. *Postharvest Biol. Technol.* 174:111440, doi: 10.1016/j.postharvbio.2020.111440.
- Mehinagic, E., G. Royer, D. Bertrand, R. Symoneaux, F. Laurens, and F. Jourjon. 2003. Relationship between sensory analysis, penetrometry and visible-NIR spectroscopy of apples belonging to different cultivars. *Food Qual. Prefer.* 14:473–484, doi: 10.1016/S0950-3293(03)00012-0.
- Naik, S. and B. Patel. 2017. Machine vision based fruit classification and grading: A review. *Intl. J. Comput. Appl.* 170:22–34, doi: 10.5120/ijca2017914937.
- Ncama, K., U.L. Opara, S.Z. Tesfay, O.A. Fawole, and L.S. Magwaza. 2017. Application of Vis/NIR spectroscopy for predicting sweetness and flavour parameters of ‘Valencia’ orange (*Citrus sinensis*) and ‘Star Ruby’ grapefruit (*Citrus × paradisi* Macfad). *J. Food Eng.* 193:86–94, doi: 10.1016/j.jfoodeng.2016.08.015.
- Nicolai, B.M., K. Beullens, E. Bobelyn, A. Peirs, W. Saeys, K.I. Theron, and J. Lammertyn. 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biol. Technol.* 46:99–118, doi: 10.1016/j.postharvbio.2007.06.024.
- Ozaki, Y. 2012. Near-infrared spectroscopy: Its versatility in analytical chemistry. *Anal. Sci. Intl. J. Japan Soc. Anal. Chem.* 28:545–563, doi: 10.2116/analsci.28.545.
- Qiu, S. and J. Wang. 2015. Application of sensory evaluation, HS-SPME GC-MS, E-nose, and E-tongue for quality detection in citrus fruits. *J. Food Sci.* 80:S2296–S2304, doi: 10.1111/1750-3841.13012.
- Qiu, S. and J. Wang. 2017. The prediction of food additives in the fruit juice based on electronic nose with chemometrics. *Food Chem.* 230: 208–214, doi: 10.1016/j.foodchem.2017.03.011.
- Shetty, N. and R. Gislum. 2011. Quantification of fructan concentration in grasses using NIR spectroscopy and PLSR. *Field Crops Res.* 120:31–37, doi: 10.1016/j.fcr.2010.08.008.
- Song, J., G. Li, X. Yang, X. Liu, and L. Xie. 2020. Rapid analysis of soluble solid content in navel orange based on visible-near infrared spectroscopy combined with a swarm intelligence optimization method. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* 228:117815, doi: 10.1016/j.saa.2019.117815.
- Wei, H., C. He, S. Zhang, H. Xiong, H. Ni, and Q. Li. 2021. Effects of four storage conditions on the sugar content, acidity, and flavor of ‘Guanxi’ honey pomelo. *J. Food Process. Preserv.* 45:E15088, doi: 10.1111/jfpp.15088.
- Xu, S., X. Sun, H. Lu, and Q. Zhang. 2019. Detection of type, blended ratio, and mixed ratio of Pu’er tea by using electronic nose and visible/near infrared spectrometer. *Sensors (Basel)* 19:2359, doi: 10.3390/s19102359.
- Zimmermann, B. and A. Kohler. 2013. Optimizing Savitzky-Golay parameters for improving spectral resolution and quantification in infrared spectroscopy. *Appl. Spectrosc.* 67:892–902, doi: 10.1366/12-06723.