Computers and Electronics in Agriculture 112 (2015) 20–27

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Predicting apple sugar content based on spectral characteristics of apple tree leaf in different phenological phases

Yao Zhang, Lihua Zheng *, Minzan Li, Xiaolei Deng, Ronghua Ji

Key Laboratory of Modern Precision Agriculture System Integration Research, Ministry of Education, China Agricultural University, Beijing 100083, China

ARTICLE INFO

Article history:
Received 15 April 2014
Received in revised form 2 January 2015
Accepted 5 January 2015
Available online 22 January 2015

Keywords:
Apple sugar content
Two-dimension correlational spectra
Spectral analysis
Apple tree leaf
SVM

ABSTRACT

Sugar degree is an important factor in determining the quality of apple. The sugar accumulation in apple fruit is closely related to fruit tree growth and development in different phases. In order to reveal the relationship between tree growth state and apple sugar content, the spectral information of apple tree leaves in different phenological phases was used to predict the fruit sugar degree. The visible and near infrared spectral reflectance of the leaves samples were measured by using a Shimadzu UV-2450 spectrograph, and the sugar content of each fruit sample growing near each leaves sample was collected and measured using laboratory methods. Then two dimensional correlation spectrum analysis was brought in, and the dynamic spectra in different phenological phases were obtained by using sugar contents as the perturbation quantity. Comprehensive observation on the spectral characteristics of leaf samples was conducted much accurately by analyzing two-dimensional correlation spectra of both synchronous and asynchronous. And then the effective spectral response bands of sugar contents and the contribution proportion to fruit sugar accumulation in different periods were investigated. And then, using the contribution proportion of each band as the single-period weighting factor, the fruit sugar sensitive wavebands were acquired. The fruit sugar content was forecasted using the sensitive bands in different phenological phases. After comparing and analyzing, it was found that the model based on parametric optimal solution of SVM showed good accuracy. The calibration R^2 of the model reached to 0.8934, the RMSEC was 0.4925 Brix, the validation R^2 reached to 0.8805, and its RMSEP was 0.4906 Brix. It reaches to a practical level and can be used to predict the sugar content in apple fruit.

1. Introduction

The sweet taste of apples depends on the sugar content, which is one of the key factors determining the quality of apples (Liu et al., 2005). Traditional detection methods used to adopt chemical testing with sampling method, which is complex, time-consuming, high cost and technical condition requested in analyzing process. Besides that, the traditional methods are difficult to achieve timely monitoring and hardly avoid destroying the samples. In recent years, near infrared spectroscopy technology as a non-destructive testing method is widely used in the determination of apple fruit sugar degree (Steinmetz et al., 1999; Peirs et al., 2000; Lu et al., 2000; Park et al., 2002). Wang and Han (2008) used genetic algorithms to optimize waveband and established partial least-squares regression model to predict apple sugar content, which reached higher precision comparing with the full spectra PLS model and the model built by the experiential spectra. Hybrid linear analysis was transplanted by Zhang et al. (2006) for assessing apple sugar degree and reached higher accuracy than the traditional models. Zhang et al. (2007) took advantage of NIR spectroscopy to detect the fruit sugar, the results showed that fruit sugar predicting model established by the wavelength of 600–1100 nm was practical and efficacious. Xia et al. (2010) proposed a novel local regression method combined with similarity evaluation infrared spectra. The results showed that PLS local regression method based on Bayesian statistics (B-PLS) could accurately predict the sugar content in apple and its performance was superior to that of partial least square global regression model (G-PLS) and partial least square local regression method based on Mahalanob (M-PLS).

In determining the sensitive wavebands of certain substances, compared with the traditional one-dimension spectroscopy analysis, the two-dimension correlation spectroscopy analysis shows more advantages. Firstly, it can improve the spectral resolution. Since the system of the two-dimension correlation spectroscopy analysis generates dynamic changes according to external disturbance excitation, the overlapping peak or the small peak which are covered formerly would make an extension and display clearly

* Corresponding author.
E-mail address: zhenglh@cau.edu.cn (L. Zheng).

http://dx.doi.org/10.1016/j.compag.2015.01.006
0168-1699/

in the third dimensional direction. Secondly, it can be used to solve the problem of the attribution of some certain puzzle peaks. Most absorption peaks have their reasonable attributions. In NIR spectrum region, because of the complex absorbance of composing frequency and double frequency, the spectrum bands overlap seriously. And this problem would be complicated to be solved if the one-dimension spectroscopy analysis was used. However, the two-dimensional correlation technology can be used to analyze the correlation relationship between wavebands and define the attribution of the spectral varieties (Lv, 2014). This method has greatly enhanced the spectra identification capability (Zhou et al., 2003). After it was extended into generalized two-dimensional correlation spectrum, it has been widely used to analyze data in many fields (Lei et al., 2008) such as polymer (Tang et al., 2007), protein (Ren et al., 2012), food (Qin et al., 2004) and medicine structure (Zuo et al., 2002).

The research results showed that no matter using the traditional methods or the NIR spectroscopy methods to detect the apple fruit sugar content, they were all focus on the apple fruit, which would lead a lack of dynamic monitoring fruit growth process and destruction to the apple fruits. Meanwhile, although many spectra analysis methods had been used, such as PLS, genetic algorithm optimization and hybrid linear analysis, the analysis with two-dimension spectra method has not been used.

In this research, the apple sugar content was served as study object. In order to determine the contributions to sugar accumulation in the bagged fruit in different physiological phases, the leaf spectra characteristics in different period were investigated to trail the fruit sugar content information. Using the sugar content as perturbation, the dynamic spectrum of apple leaf was obtained. By means of two-dimensional correlation spectroscopy analysis, the most sensitive wavebands were found to assess the change of sugar concentration. Combining the sugar accumulation contributions in different phenophase, the sugar content in apple was inverted and therefore the sugar accumulation with the trees growing was revealed.

2. Materials and methods

2.1. Experiment

The experiments were conducted in the apple orchard located in Beijing XiangTang culture village (E116.2154°, N40.1439°) in flowering period (F.P. May 13), Shoot-growing Stage (S.G.S. May 29), fruit Setting Period (F.S.P. June 12), Branch shooting period (B.S.P. July 12), bud differentiation stage (B.D.S. July 20) and defoliation period (D.P. October) separately. The apple breed is red Fuji, and the soil type belongs to sandy loam. 10 apple trees were selected randomly from different regions. Then a main branch of each tree was selected and three representative parts (base part, middle part and top part) of every bough were marked, and these were the target area for all the subsequent experiments. About 8–10 leaves were collected as one leaves sample in each marked area in six periods respectively and its VIS/NIR spectra reflectance were measured. The corresponding apple was picked as one fruit sample in the same area in defoliation period and its sugar content was detected using the laboratory chemical analysis method. In this research, 20 samples were selected to establish the sugar content prediction model and 10 samples were used to validate the model.

2.2. Spectrum determination

The spectral reflectance was detected by using the spectrometer (UV-2450, Shimadzu Co. Ltd., Japan) (Li et al., 2011). Its spectral range is 300–900 nm, its spectral resolution is 1 nm. Before data collecting, the spectrometer correction was conducted using the standard whiteboard (reflectivity of whiteboard is 1). Each sample was measured for 3 times and mean value was calculated as its relative reflectance.

2.3. Apple sugar content detection

Apple sugar contents were measured by referring to the standard of the physical and chemical general principles of China food hygiene inspection method (GB/T5009.1-2003). In order to represent the sugar content information more accurately, about 1 cm^3 of pulp at light side and dark side of each apple were took and squeezed. The juice was then measured by the handheld refractometer (PR-101α, Atago Co. Ltd., Japan). Then the average sugar content of each apple was calculated (Chen and Liu, 2012).

2.4. Experiment data processing method

2.4.1. Two-dimensional correlation spectrum calculating

This research using sugar contents as perturbation factor and the dynamic spectrum of apple leaf was then obtained. The results of two-dimensional correlation analysis were conducted taking advantage of Shige software (Shigeaki Morita, Kwansei-Gakuin University, Japan). Two-dimensional chromatogram contour layer was set to 8.

Two-dimensional correlation analyzing is aimed at studying the variability of dynamic spectrum $\tilde{y}(v,t)$ (t is outer interference). The definition of $\tilde{y}(v,t)$ is shown as formula (2) (Shen et al., 2005). When there is outer interference, it is represented by the original spectrum subtracting the reference spectrum. When there is no outer interference, it is zero.

$$\tilde{y}(v,t) = \begin{cases} y(v,t) - \bar{y}(v) & T_{\min} < t < T_{\max} \\ 0 & else \end{cases} \tag{1}$$

Two-dimensional correlation intensity of $X(v_1, v_2)$ is obtained by analyzing the correlation between the corresponding variations of $\tilde{y}(v,t)$ caused by two independent variables of v_1 and v_2 respectively. It is shown as formula (2) (Wang et al., 2009).

$$X(v_1, v_2) = \phi(v_1, v_2) + i\psi(v_1, v_2) \tag{2}$$

In which, $\phi(v_1, v_2)$ is the intensity of 2D correlation synchronization, and it represents the similarity of the intensity variance caused by two independent optical variables with external disturbance of t. $\psi(v_1, v_2)$ is the intensity of 2-D correlation asynchronization, which represents the difference of the intensity variance (Shen et al., 2005).

In synchronization spectrogram, if the peak at (v_1, v_2) is positive, it means that the variance of spectra intensity of v_1 and v_2 changing at the same direction, otherwise, they will change in opposite directions. When $v_1 > v_2$, if the product of the peaks at (v_1, v_2) of both the synchronization and asynchronous spectrogram is positive, it means that v_1 changes prior to v_2, otherwise, v_1 changes after v_2 (Wang et al., 2009).

2.4.2. Support vector machine (SVM) regression

Support vector machine (SVM) (Vapink, 2000) has some advantages in prediction. However there is no specific theory at selecting the parameters (penalty parameter of C and kernel function parameter of g) in SVM studying and training, and which affects the prediction accuracy and efficiency of SVM. Artificial parameters setting are always adopted on the basis of specific issue. The better parameter combinations were obtained after repeating selecting and comparing, which lead to inefficiency and blindness. This research adopted the cross validation method to realize the optical parameter selection in SVM. Firstly, penalty parameter of C and