pigments in tea leaves based on Raman spectroscopy and calibration model transfer

Jianjun Zeng[1], Shuai Zhang[1], Wei Luo[1], Junjing Sha[2], Yifeng Huang[3], Xuemei Liu[3], Baishao Zhan[1], Hailiang Zhang[1,*] and Xiaoli Li[2,*]

[*]Correspondence: 2242526723@qq.com (HZ[1]); xiaolili@zju.edu.cn (XL[2])
[1]College of electrical and automation engineering, East China Jiaotong University, Nanchang 330013, China.
[2]College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310058, China.
Full list of author information is available at the end of the article

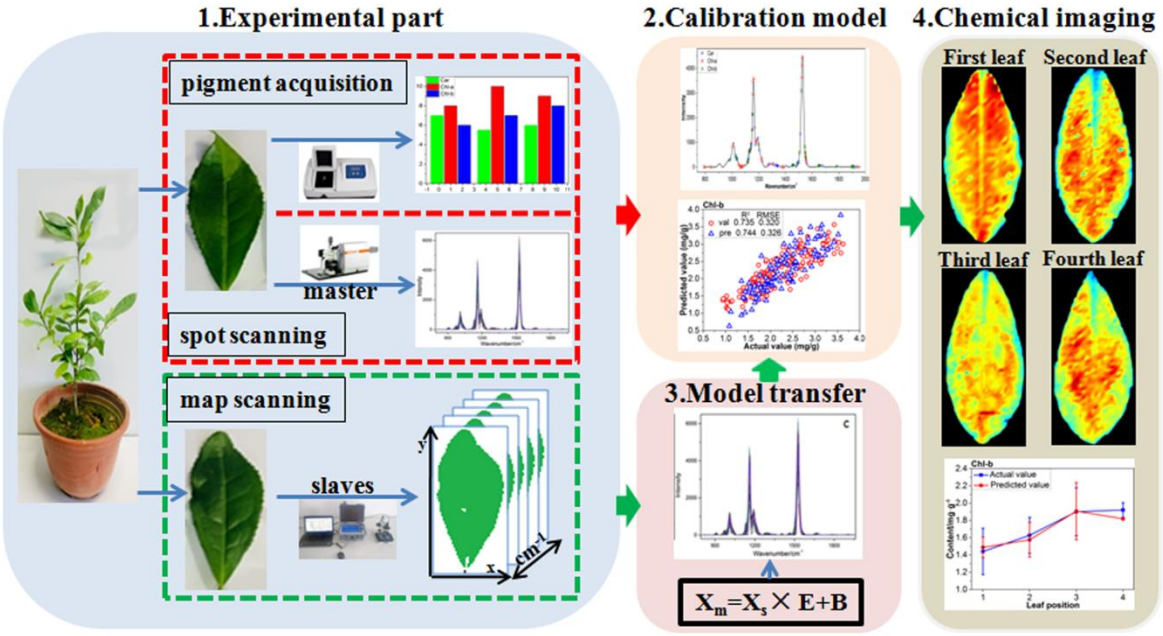
Abstract

Background: Photosynthetic pigments participating in the absorption, transformation and transfer of light energy play a very important role in plant growth. While, the spatial distribution of foliar pigments is an important indicator of environmental stress, such as pests, diseases and heavy metal stress.

Results: In this paper, in situ quantitative visualization of chlorophyll and carotenoid was realized by combining the Raman spectroscopy with calibration model transfer, and a laboratory Raman spectral model was successfully extended to a portable field spectral measurement. Firstly, a nondestructive and fast model for determination of chlorophyll and carotenoid in tea leaf was established based on confocal micro-Raman spectrometer in the laboratory. Then the spectral model was extended to a real-time foliar map scanning spectra of a field portable Raman spectrometer through calibration model transfer, and the spectral variation between the confocal micro-Raman spectrometer in the laboratory and the portable Raman spectrometer were effectively corrected by the direct standardization (DS) algorithm. The portable map scanning Raman spectra of the tea leaves after the model transfer were got into the established quantitative determination

model to predict the concentration of photosynthetic pigments at each pixel of the tea leaves. The predicted photosynthetic pigments concentration of each pixel was imaged to illustrate the distribution map of foliar pigments. Statistical analysis showed that the predicted pigment contents were highly correlated with the real contents.

Conclusions: It can be concluded that the Raman spectroscopy was applicable for in situ, non-destructive and rapid quantitative detecting and imaging of photosynthetic pigment concentration in tea leaves, and the spectral detection model established based on the laboratory Raman spectrometer can be applied to a portable field spectrometer for quantitatively imaging of the foliar pigments.



Keywords: photosynthetic pigments, concentration distribution imaging, feature extraction, model evaluation, quantitative analysis, Raman spectroscopy

Background

Tea is one of the world's three major beverages. The world's tea production exceeds 4 million tons per year and more than 2 billion people consume tea [1]. Tea, which contains high levels of antioxidants and can prevent many diseases including cardiovascular and cancer, has received more and more attention from people [2]. Photosynthesis is the determinant of productivity and the basis of plant growth and development, and an important source of carbon in plants. The photosynthesis

of tea leaves is closely related to the quality and yield of tea. Photosynthetic pigments including chlorophyll a (Chl-a), chlorophyll b (Chl-b), and carotenoids (Car) play a very important role in plant growth. Furthermore, the spatial distribution of foliar pigments is an important indicator of environmental stress, such as pests, diseases and heavy metal stress [3, 4]. In addition, the color and lustre of tea plays an important role in the consumption and production of tea, and the pigment concentration of tea is also an important factor affecting the color of tea. Therefore, developing a method for nondestructive detection and quantitative visualization of foliar photosynthetic pigment content is an important task for plant protection, cultivation and tea processing [5].

Most of the existing imaging studies on photosynthetic pigments concentration are based on visible-near infrared hyperspectral imaging technology. Previous studies have investigated crops such as cucumber leaf [6, 7], spinach [8], pepper leaf [9], tomato [10] and so on, and have achieved good results. Zhao et al. [11] used visible-near infrared hyperspectral technology to extract the corresponding feature parameters and built models based on 7 algorithms. The chlorophyll concentration was predicted by the models and the distribution map of chlorophyll concentration was drawn. Raman spectroscopy is a non-destructive analytical technique, which is based on the scattering interaction between light and chemical bond in materials. It can provide detailed information of chemical structure, phase and morphology, crystallinity and molecular interaction of samples. Compared to visible-near infrared hyper-spectrum, Raman spectroscopy has the advantages of high resolution, which can provide more spatial information for pest and heavy metal detection, and help to further study the mechanism of pests and heavy metal stress.

There are a lot of studies on the use of Raman spectroscopy for chemical imaging, but few studies have been done on the applicability of Raman quantitative detection model. Schulz et al. [12] used a NIR-FT-Raman spectrometer to detect carotenoid in various fruits and vegetables such as carrot, tomato, and nectarine, and found Raman spectral fingerprint peaks of carotenoid in plants. The characteristic peaks were used to image the carotenoid distribution inside the plants. Qin et al. [13] developed a bench-top point-scan Raman chemical imaging system to detect and visualize the

internal distribution of lycopene in postharvest tomato, and established a Raman chemical image to visualize the spatial distribution of lycopene at different stages of maturity. Yang et al. [14] used a custom row-scan Raman hyperspectral imaging system to detect and display the main chemical components of maize seeds. The characteristic peaks associated with corn starch, mixture of oil and starch, zeaxanthin, lignin and oil were found. Each single band image corresponding to the characteristic band successfully represents the spatial distribution of chemical components in seeds.

In addition to the above qualitative analysis, Raman spectroscopy also has many applications in quantitative detection. Baranska et al. [15] detected lycopene and carotene in tomato and its products by FT Raman spectroscopy, and its modeling effect was better than that of near infrared spectrum but slightly lower than that of infrared spectrum. Bhosale et al. [16] and Dane et al. [17] detected the carotenoid concentration in kinds of fruits and vegetables (such as tomato, carrot, mango, etc.) using Raman spectroscopy. The result showed that there was a high correlation between the Raman spectrum signal intensity and the carotenoid concentration, and the correlation coefficient (R) was up to 0.9618. The above researches shows that Raman spectroscopy has great potential for quantitative detection of pigments, but the applicability of the quantitative model of Raman spectroscopy, especially the field applicability of the laboratory model, has not been studied so far.

In practical applications of spectral determination model, it is often encountered that a multivariate calibration model developed based on one instrument (Master) cannot be used on another instrument (Slave) of the same type as the Master. Or there will be big biases in the prediction result. The poor adaptability of this spectral model greatly limits the application prospect of spectral detection technology, and the spectral models that took a lot of time and money to build in the laboratories are difficult to be used in field or production practice. A strategy of model transfer has been frequently adopted to solve this problem. The essence of model transfer is to overcome the inconsistency between the measured signals (or spectra) of samples on different instruments [18]. Confocal micro-Raman spectrometer has high resolution and precision, and can capture more useful

spatial information, but this instrument is too expensive and heavy to be used in field or in vivo plant detection. Portable Raman spectrometer is compact and convenient, and can be used for in vivo or field detection, but it has the disadvantage of low resolution. Linking the two through model transfer not only can be applied in actual production, but also make the model more accurate. Ji et al. [19] used a direct standardization (DS) model transfer algorithm to remove the environmental factors from the field spectrum so as to effectively estimate the soil properties. Wang et al. [20] proposed a model transfer algorithm based on genetic algorithm, which makes the partial least squares model of aviation fuel density successfully transferred from one instrument to another, and the accuracy of model prediction result is close to the calibration model and higher than the DS model. Good results have been obtained in the previous literatures, but the effect of model transfer method on the variation correction of Raman spectrometers has not been reported. Here we are committed to building a high-precision Raman spectral pigment measurement model based on high-performance instrument in the laboratory, and study the application of this spectral model to the portable field measurement equipment.

In order to solve the above problems, this paper aims to study the Raman spectral characteristics of tea leaf and establish a in situ quantitative analysis model between the concentration of photosynthetic pigment and its confocal micro-Raman spectra in tea leaf. On this basis, the portable Raman mapping data after model transfer were brought into the established model, the concentration of each pixel was predicted, and the chemical imaging was carried out to obtain the distribution map of the concentration of photosynthetic pigment in tea leaf at different position.

Materials and Methods

Materials and instruments

Material preparation

The tea variety is longjing 43 (*Camelliasinensis(L.) O.Kuntze*). In the summer of 2012, longjing 43 seedlings were obtained from the tea garden in Fuyang, Hangzhou, China. Tea seedlings were planted in pots in the agricultural internet of things exhibition center of college of biosystems engineering and food science, Zhejiang University, receiving natural light and artificial watering.

Material 1: in total of 315 leaves including the first four leaves in shoot (as shown in Fig. s1) were plucked for Raman spectral collection, then all the leaves were stored in the refrigerator within 4 ℃ immediately. Following, approximately 0.1 g of weighed leaf, excluding central vein was taken for reference measurement of concentration of chlorophyll and carotenoid according to the Reference [8] based on the ultraviolet spectrophotometer with the unit $mg \cdot g^{-1}$. In detail, 10 mL pigment extraction solution (95% alcohol solution) was added to the cut and ground sample, and stored in a dark room for about 24 h.

Material 2: in total of 16 leaves including 4 leaves in each position (as shown in Fig. s1) were collected for different four tea plants. And the photosynthetic pigment concentration of the 16 leaves was measured in the same method as in the previous section.

Instrument

Master spot scanning: a laser confocal micro-Raman spectrometer (Renishaw, United Kingdom/Via-Reflex 532/XYZ) was used for collecting Raman spectra of the Material 1. Specific parameters were as follow: the excitation wavelength is 532 nm; the spectral collection range is 579-3062 Raman shift/cm^{-1} with a spectral resolution of 0.2 nm, the laser intensity is 50 mW, the exposure time is 1 s, and the objective lens multiple is 5. Each sample was collected at three points from top to base, and the average spectrum was used as the representative Raman spectrum of the sample. So, a total of 315 spectra were obtained.

Slave map scanning: a portable Raman spectrometer (Ocean optics QE Pro, United States) was used for spectral collection of Material 2. Specific parameters were as follow: the excitation wavelength is 532 nm, the spectral collection range is 77-2146 Raman shift/cm^{-1}, the laser intensity is

100 mW, the exposure time is 3 s, and the average time is 2 s. The 16 leaves were map scanned by the Raman spectrometer with horizontal and vertical spatial resolution of 1 mm.

In this experiment, the software WIRE 3.3 (Renishaw, United Kingdom) was used to collect and extract Raman spectral data. The pictures in this paper were drawn by Origin 9.0 (Originlab, United States) and Photoshop CS6 (Adobe, United States), and all the preprocessing and modeling methods were performed in Matlab 2013b (MathWorks, United States).

Spectral analysis methods

Spectral pretreatment

The obtained spectral information contains not only the chemical structure information of the sample, but also many background and noise signals from the interference source such as the instrument itself and the experimental operating environment. Therefore, in order to eliminate the influence of the extraneous and interfering signals on the sample signal, the original data can be preprocessed [21]. In this study, five data preprocessing methods were applied including multiplicative scatter correction (MSC), wavelet transform (WT), standard normal variate (SNV), rolling-circle filter (RCF) and adaptive iteratively reweighted penalized least squares (airPLS).

The full-band Raman spectrum contains a large amount of redundant information and noise [22]. These interference signal not only affect the prediction performance of the model, but also are not conducive to further detecting the Raman spectral response mechanism of chlorophyll. So, the competitive adaptive re-weighted algorithm (CARS) was used to extract the effective band for spectral measurement of the photosynthetic pigments. Based on the effective bands from CARS, the computational complexity of the spectral modeling can be reduced. Furthermore, the Raman spectral response mechanism of photosynthetic pigment may be discovered based on the assignment of these characteristic bands.

Modeling and evaluation methods

Partial least squares regression (PLSR) was adopted to establish a quantitatively relationship between the concentration of photosynthetic pigments and Raman spectra of leaf. The PLSR has the advantages of simplicity, accuracy, convenience and wide applicability. It is the most commonly used and most effective multivariate statistical method in chemometric modeling analysis [23].The performance of the PLSR model was evaluated by the following indicators including coefficient of determination (R^2), root mean square error (RMSE) [24]. In detail, the R^2_C, R^2_{CV}, R^2_P and $RMSE_C$, $RMSE_{CV}$, $RMSE_P$ represent the determination coefficient and root mean square error of calibration, cross validation and prediction, respectively.

Considering the subsequent model transfer, the spectral bands of the two instruments need to be unified. The common band range of the two spectrometers is 579-2146 cm^{-1}, but there are obvious high-frequency noise signals at the both ends. So the spectral range of 792-1961 cm^{-1} was selected for modeling.

Model transfer

In order to test the prediction result of the model, the sample after the scanning should measure 3 photosynthetic pigments concentration, and compare the actual value with the predicted value. Scanning takes a long time, and if it is detected in vitro, it will inevitably affect the accuracy of photosynthetic pigment concentration. Although the master also has a scanning function, its structure is complicated and impossible to perform living body detection of tea leaves. The slave is small, portable, simple in structure, and can perform scanning without removing the tea leaves. Therefore, the slave is used to scan four leaves of different leaf positions.

The direct standardization (DS) method was adopted to improve the adaptability of the spectral model, which is a multivariate full-spectrum model transfer algorithm [25]. The advantage is that the principle is simple, the difference between the standard spectral data and the spectral data to be corrected can be compared, and each wavenumber is sequentially corrected by the full band to realize the transfer of the model [20].

The common spectral range of 579-2146 cm^{-1} of the two Raman spectrometers was selected, and to remove the high-frequency noise signals at both ends, the range of 792-1961 cm^{-1} was selected for further model transfer. Since the spectral resolutions of the two spectrometers are different, interpolation processing was performed to form a uniform number of spectral variables. It is important for model transfer to choose representative samples to define the differences between the master and slave instruments. For the slave instrument, the average spectrum of all pixels for each leaf (minus the background) map scanning was taken as the representative spectrum of the leaf, so a total of 16 spectral profiles were obtained. And 4 representative spectra were selected by the Kennard-Stone (KS) algorithm from the 16 spectra. While, for the master instrument, 4 spectra were also selected by KS algorithm from the 315 spectra of master spot scanning. Model transfer was performed by selecting 4 spectra (including all of the leaf positions) from the master and slave instrument.

The flow of the DS algorithm is as follow [26, 27]. The spectral matrices of master and slave are X_m, X_s, respectively. Both X_m and X_s have size m × p, where m represents the number of representative transfer spectra (4 in this case) and p represents the number of wavenumbers.

$$X_m = X_s \times E + B, \quad (1)$$

Where E is the transfer matrix with size p × p of unknown parameter, which accounts for the variation in both X_m and X_s, and B is the background correction matrix [19].

The spectrum S_s of sample to be tested measured on the slave can be used for analysis after conversion:

$$S_{s,std} = S_s \times E + B, \quad (2)$$

The spectral data $S_{s,std}$ represents the corrected slave spectra through transferred by the DS algorithm, which will be transmitted to the photosynthetic pigments determination model developed by the master. And the prediction accuracy of the model for the slave sample is the performance of model transfer.

Results and Discussion

Establishment of quantitative determination model

Sample division

The 315 samples from master spot scanning were divided into a training set and a test set based on 2:1 ratio. First, the samples were arranged in ascending order according to their concentrations of Chl-a, Chl-b and Car, and each three is one set in turn. Then, the second sample in each set is divided into the test set, and the rest are set as the training set. So, 210 training set samples and 105 test set samples were obtained. The statistical information of the sample sets were shown in Table 1.

Table 1. The statistical information of the sample set.

		Max[1] (mg·g^{-1})	**Min[2] (mg·g^{-1})**	**Mean ± SD[3]**	**Num[4]**
Car	Training set	1.4917	0.1901	0.8461 ± 0.2282	210
	Test set	1.4786	0.2557	0.8465 ± 0.2267	105
Chl-a	Training set	10.4111	1.877	5.6513 ± 1.4116	210
	Test set	10.4111	2.3599	5.6559 ± 1.4191	105
Chl-b	Training set	5.7826	0.6288	2.2675 ± 0.6804	210
	Test set	5.7826	0.9380	2.2785 ± 0.7128	105

[1] Max, maximum value; [2] Min, minimum value; [3] SD, standard deviation; [4] Num, number of samples.

Raman spectral quantitative determination of photosynthetic pigments in tea leaf

As shown in Table 2, different pretreatment methods produced different results referring to the values of R^2 and RMSE, indicating that pretreatment has a great influence on the performance of the model. In regard of the Car, model 5 based on the WT preprocessing method is obviously better than model 1 based the original data. In detail, the R^2_P of the model 1 increased from 0.614 to 0.713 of the model 5, and $RMSE_P$ of the model 1 decreased from 0.140 to 0.108 of the model 5. For Chl-a and Chl-b, model 10 and model 16 respectively obtained the best results based on the optimal pretreatment method of RCF. Comparing with the model 7 based on the original data, the R^2_P of model 10 increased from 0.597 to 0.800, and $RMSE_P$ of model 10 decreased from 0.900 to 0.599. While,

the R^2_P and $RMSE_P$ of model 16 were respectively 0.734 and 0.330, which were obviously better than the relevant parameters (0.718 and 0.342) of model 13. Furthermore, the difference among calibration, validation and prediction of the model 5, 10 and 16 was also relatively small, which indicates that the stability of these models is improved through pretreatment.

Table 2. PLSR modeling results of different pretreatment methods.

	Model	Pretreatment	Calibration		Validation		Prediction	
			$RMSE_C$	**R^2_C**	**$RMSE_{CV}$**	**R^2_{CV}**	**$RMSE_P$**	**R^2_P**
Car	1	Origin	0.082	0.843	0.107	0.732	0.140	0.614
	2	MSC	0.080	0.847	0.106	0.738	0.134	0.650
	3	SNV	0.080	0.848	0.105	0.744	0.135	0.642
	4	RCF	0.122	0.653	0.142	0.532	0.131	0.662
	5	WT	0.086	0.826	0.106	0.737	0.108	0.713
	6	airPLS	0.123	0.646	0.140	0.549	0.199	0.227
Chl-a	7	Origin	0.463	0.870	0.577	0.798	0.900	0.597
	8	MSC	0.557	0.812	0.692	0.712	0.867	0.626
	9	SNV	0.834	0.650	0.898	0.600	0.886	0.609
	10	RCF	0.535	0.827	0.653	0.745	0.599	0.800
	11	WT	0.095	0.806	0.116	0.715	0.108	0.721
	12	airPLS	0.622	0.773	0.820	0.615	1.078	0.421
Chl-b	13	Origin	0.240	0.854	0.300	0.774	0.342	0.718
	14	MSC	0.284	0.784	0.339	0.695	0.366	0.677
	15	SNV	0.282	0.789	0.343	0.690	0.353	0.701
	16	RCF	0.295	0.810	0.347	0.717	0.330	0.734
	17	WT	0.299	0.772	0.349	0.690	0.354	0.698
	18	airPLS	0.247	0.844	0.340	0.707	0.385	0.644

Selecting of characteristic bands for quantitative determination

Raman spectroscopy provides a wide range of spectral information. In this research, there are 1005 and 1044 spectral variables from the master and the slave instruments respectively. And, there are still remaining 448 spectral variables after intercepting the common wavenumbers and removing the two ends of the spectrum seriously disturbed by noise. The Raman spectra contain not only biological, physiological and structural information related to detection objects, but also redundant information [22]. In order to explore the mechanism of the detection of photosynthetic pigment in tea leaves by Raman spectroscopy, a large number of redundant and interference information were excluded. Furthermore, selecting a small number of effective band can shorten the modeling time and improve the accuracy of the model. The CARS was adopted to extract the effective band for spectral measurement of the photosynthetic pigments based on the spectral data pretreated by the WT and RCF pretreatment, and the models based on the characteristic bands were established, and the modeling results were shown in Table 3. It can be found that the RCF was better than the WT pretreatment for all the three pigments, and the models based on these characteristic bands were better compared with the full-band models (as shown in Table 2).

Table 3. PLSR modeling results based on characteristic band.

	Model	Number of characteristic bands	Pretreatment	Calibration		Validation		Prediction	
				RMSEC	**R^2_C**	**RMSECV**	**R^2_{CV}**	**RMSEP**	**R^2_P**
Car	19	40	RCF	0.109	0.739	0.124	0.669	0.101	0.769
	20	40	WT	0.125	0.702	0.133	0.661	0.123	0.706
Chl-a	21	37	RCF	0.586	0.793	0.623	0.767	0.517	0.852
	22	37	WT	0.603	0.782	0.777	0.753	0.623	0.835
Chl-b	23	32	RCF	0.298	0.769	0.320	0.735	0.326	0.744
	24	32	WT	0.343	0.746	0.443	0.708	0.389	0.726

The spectral profiles before and after the RCF pretreatment were shown in Fig. 1, it can be found that the RCF method eliminated the fluorescence background and increased the signal-to-noise ratio

of spectra, this may be the reason why the RCF pretreatment can improve the performance of the spectral determination models. Oh et al. [28] used RCF pretreatment in real-time estimation of glucose concentration in algae by Raman spectroscopy, and the result was also improved.

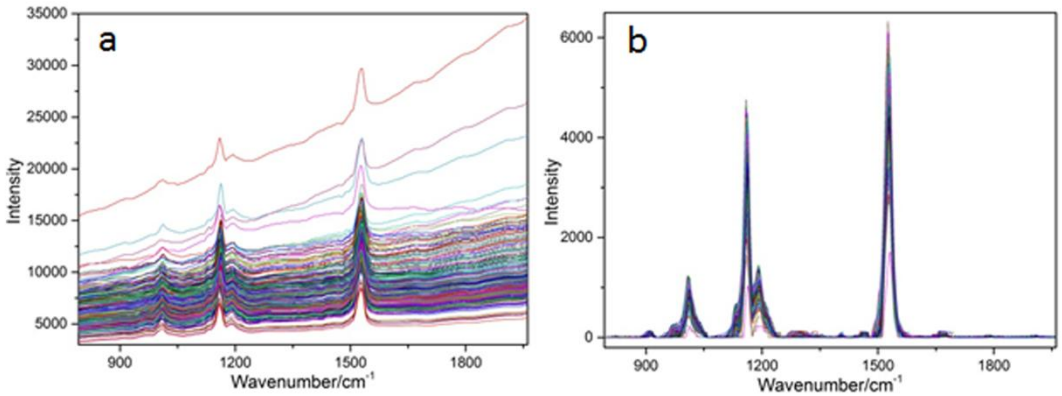


Fig. 1. Raman spectral profiles of tea leaf samples from master instrument. (a) original spectra; (b) spectra processed by the RCF.

Scatter diagram of prediction values and real values of the models (model 19, 21 and 23) for training and test samples were shown in Fig. 2. It can be found that the models based on the characteristic wavenumbers had achieved better result than the model based on the full-band model (as shown in Table 2). In addition, low dimensional input variables of characteristic wavenumbers greatly reduce the complexity of the model and improve the calculation speed of the model.

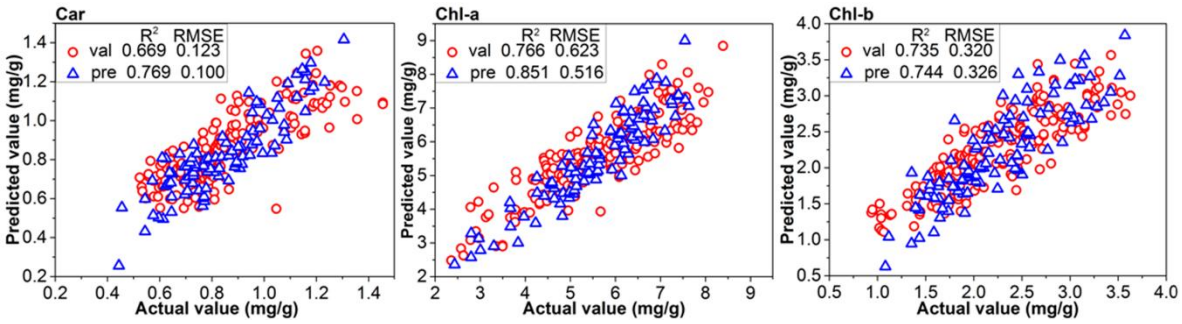


Fig. 2. PLSR model result based on the characteristic wavenumbers.

As the Raman spectroscopy can reflect the fingerprints information of the composition and structure of substances, an assignment of these characteristic wavenumbers was implemented to further explore the substance basis of quantitative determination of pigment by Raman spectroscopy.

The characteristic bands corresponding to each model were showed in Fig. 3. There were three distinct peaks in the figure, including the rocking vibration in the CH_3 plane at 1008 cm^{-1}, the C-C stretching vibration at 1159 cm^{-1}, and C=C stretching vibration at 1528 cm^{-1} which are the characteristic peaks of photosynthetic pigment [29, 30]. And the assignment of these characteristic wavenumbers was shown in Table 4, it can be seen that most of the wavenumbers were related to photosynthetic pigment, which explains the reason why models based on characteristic wavenumbers obtained good results.

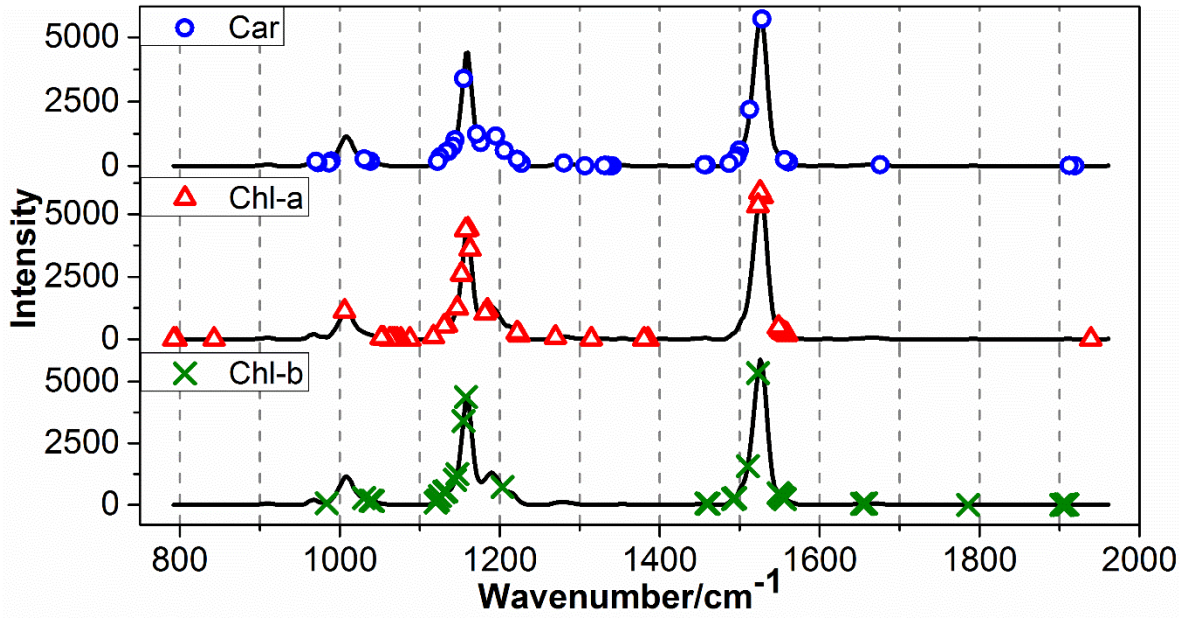


Fig. 3. Characteristic wavenumbers selected based on CARS algorithm.

Table 4. Chemical assignment of Raman characteristic wavenumbers.

Wavenumbers/cm^{-1}	Assignment	References
788/800/1046/1073/1114/1222/1553	Chlorophyll a	[31]
971	β-carotene [τ(C_{11}-C_{12})]	[32]
988	Chlorophyll	[33]
1026	β-carotene [ϱ(9Me), ν(C_8–C_9)]	[32]
1068/1155/1165/1265/1392/1530/1555	Chlorophyll a	[34]
1128/1160/1210/1380/1465/1523/1567/1644	Chlorophyll b	[34]
1133	β-carotene	[35]

1137	β-carotene [ν(C_{10}–C_{11})]	[32]
1144	Chlorophyll [ν(CN), δ(CNC)]	[33]
1145	Chlorophyll a [ν(C_aN), δ(C_aNC_a)]	[32]
1149	β-carotene [14–15, 15H, 15=15′]	[35]
1157	β-carotene [ν(C_{14}–C_{15}), δ(C_{15}–H)]	[32]
1172	β-carotene [5–4, 18r, 6–7]	[35]
1186	Chlorophyll a [ν($C_m$$C_{10}$), δ ($C_b$H)]	[32]
1187	β-carotene [10H, 11H, 8–9]	[35]
1191	β-carotene [δ(C_{10}–H), δ(C_{11}–H)]	[32]
1216	β-carotene [ν(C_{12}–C_{13}), δ(C_{14}–H)]	[32]
1225	Chlorophyll [δ(CH), δ(CH_2)]	[33]
1226	β-carotene [12–13, 14H, 15=15′]	[35]
1281	β-carotene [15H, 14H, CCC15b]	[35]
1310	β-carotene [12H, 11=12, 15=15′]	[35]
1322	β-carotene [δ(C_{12}–H), ν(C_{11}-C_{12})]	[32]
1347	β-carotene [12H, 11=12, 15=15′]	[35]
1450	β-carotene [δas(9Me), δas(13Me)]	[32]
1485	β-carotene [13=14, 11=12, 12H]	[35]
1518	Lutein [C═C stretching vibration]	[36]
1524	Chlorophyll a [ν($C_b$$C_b$), ν($C_a$$C_b$)]	[32]
1528	Carotenoids [ν1(C═C)]	[33]
1542	Chlorophyll a [ν($C_b$$C_b$)]	[32]
1552	Chlorophyll a [ν($C_a$$C_b$), ν($C_b$$C_b$)]	[32]
1562	β-carotene [11=12, 15H, 12H]	[35]

285

In addition, several characteristic wavenumbers extracted in this study were also related to protein (1651 cm^{-1}) and nucleic acid (1665 cm^{-1}) [36], etc., this may be due to that the concentration of photosynthetic pigments is the percentage of the amount of pigment to the mass of dry matter in tea leaves, in other words, the quantity of other dry matter in tea will also affect the percentage of pigment, so the characteristic peaks of other dry matter in the tea will also be selected. Furthermore, the wavelength selection algorithm based on data mining may also select some bands without specific component assignment as a benchmark for data processing.

Calibration model transfer

Direct standardization of spectral data from the master and slave instruments

The direct standardization (DS) method was adopted to standardize the Raman spectral responses from the slave instrument. It can be seen from the Fig. 1(a) and Fig. 4(a) that the spectral data measured by the two instruments all had distinct fluorescent background, and the trend of the fluorescent background was different due to the different instruments. The spectra after removing the fluorescent background by using RCF was respectively shown in Fig. 1(b) and Fig. 4(b), it can be found that the RCF pretreatment greatly improves the signal to noise ratio of the spectra, which is conducive to the subsequent analysis. Slave spectral after DS was shown in Fig. 4(c). Comparing with Fig. 1(b), it can be found that the spectra of slave instrument (as shown in Fig. 4(c)) after DS was similar with that of master instrument, indicating that the spectral variation between the master and slave Raman spectrometer can be effectively eliminated.

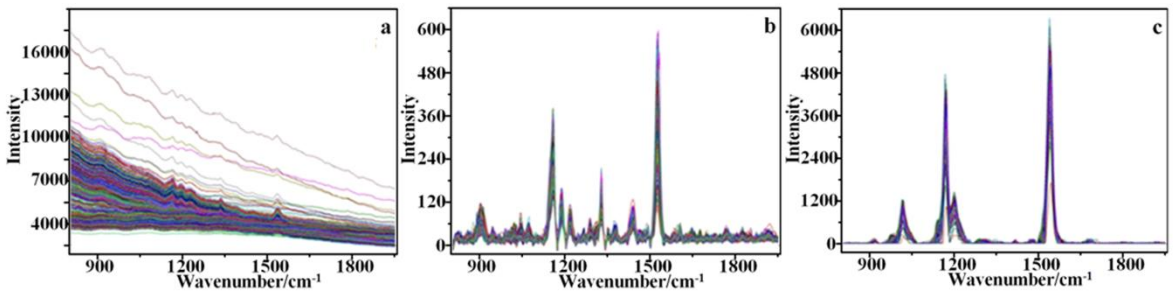


Fig. 4. (a) Raman spectral profiles of tea leaf samples from slave instrument. (a) original spectra; (b) spectra processed by the RCF; (c) spectra after DS.

Imaging photosynthetic pigment in tea leaf based on model transfer

Through the above analysis, the quantitative relationship between photosynthetic pigment concentration in fresh leaves and there're master Raman spectroscopy had been verified, and the quantitative determination models of chlorophyll and carotenoid concentration based on characteristic wavenumbers had been established. In order to realize the in situ and non-destructive imaging of chlorophyll and carotenoid concentration in fresh leaves of tea, the slave Raman spectra of material 2 after DS were transported into the established model 19, 21, 23, respectively in pixel-wise order, so the photosynthetic pigment concentration of each pixel in tea leaf was predicted. The predicted photosynthetic pigment concentration was imaged, and the image was subjected to filter filtering to obtain distribution maps of photosynthetic pigments as shown in Fig. 5.

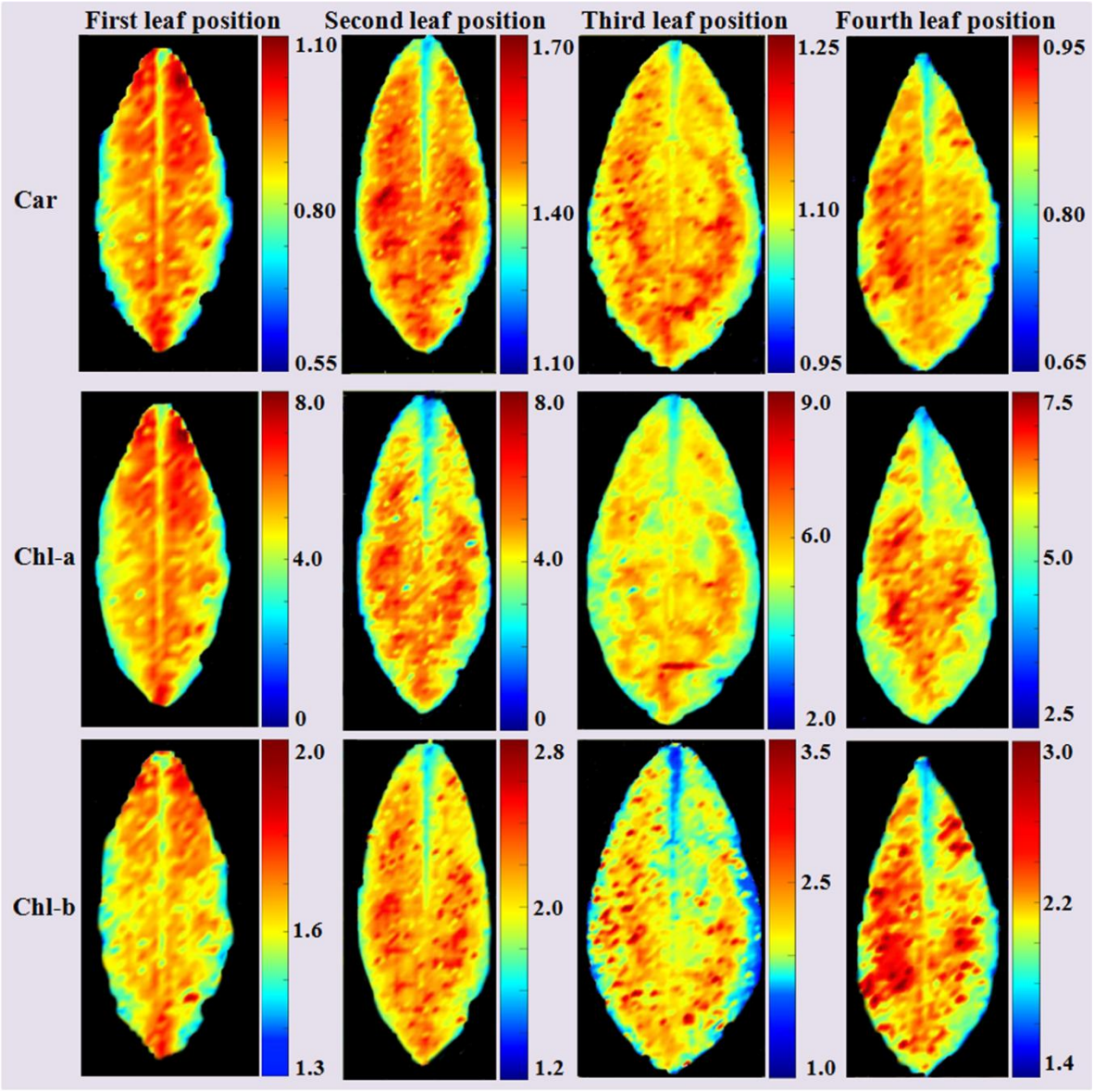


Fig. 5. Photosynthetic pigment concentration distribution map.

By imaging the photosynthetic pigments concentration, it can be found that the pigments concentration in the central vein and margin of the leaf is significantly lower than that in other region, which is related to the maximum efficiency of photosynthesis. This is consistent with the finding of Zhao et al. [11]. In zhao's paper, $RMSE_C$, R^2_C, $RMSE_P$ and R^2_P of Chl-b model are 9.918, 0.711, 8.601, and 0.693, respectively, which is obviously worse than the result of our research. Therefore, the above results show that it is feasible to predict the concentration of photosynthetic pigments based on Raman spectroscopy. Furthermore, the spectral calibration model constructed in

the laboratory (master Raman spectrometer) can be used to measure the distribution of foliar pigment with portable instruments (slave) in the field through model transfer.

Evaluation of performance of the calibration model transfer

After the map scanning spectra of the slave spectrometer were corrected, the spectrum at each pixel was brought into the quantitative determination model to predict the photosynthetic pigment concentration at that pixel. Then the photosynthetic pigment concentration of the foliar pixels was averaged to represent the pigment concentration of the leaf. Furthermore the predicted average value of pigment concentration was compared with the actual value to evaluate the performance of this calibration model transfer, in detail, the R^2 and RMSE were shown in Fig. s2. It can be found that the predicted value of the model for the foliar map scanning spectra is highly correlated with the actual value, which indicates that the pigment determination model based on the master instrument can predict the spectrum of the slave instrument after calibration model transfer. The imaging of foliar pigments results and the correlation analysis proved that the model transfer of the two spectrometers had achieved good results, and this method is feasible.

The mean and variance of the actual and predicted value of photosynthetic pigment in four leaves of the same leaf location were calculated, as shown in Fig. 6. As can be seen from the Fig. 6, the trend of the actual value of photosynthetic pigment increase firstly and then decrease, this is consistent with the finding of Vicente et al. [37]. When at first leaf position, the photosynthesis ability of the leaf is weak, the concentration of photosynthetic pigment is low. With the increase of leaf age, the leaves develop gradually, the photosynthetic pigment concentration increase, and the photosynthesis rate reach the highest value; while the third leaf position enter the aging stage, the photosynthetic enzyme expression ability decrease, and the photosynthetic pigment concentration decrease [9]. The result show that the model established in this paper also has a prospect in the study of the leaf position and leaf age.

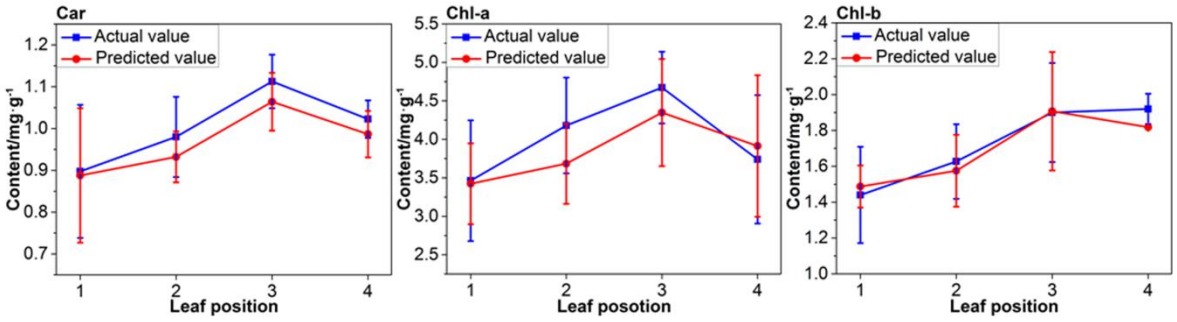


Fig. 6. Line chart of actual and predicted values of photosynthetic pigment concentration.

Conclusions

In the study, the potential of Raman spectroscopy for in situ, non-destructive and rapid quantitative detection and imaging of photosynthetic pigment concentration in tea leaves was proved. Based on the Raman spectral pretreatment method combined with the CARS characteristic bands selection, the quantitative determination models of chlorophyll and carotenoid concentration were established by regression analysis. By comparison, it can be found that the best pretreatment RCF was most suitable to eliminate the fluorescence interference and other noise in Raman spectrum. And the Raman spectral characteristic bands for pigments detection selected by CARS have been proved to be the Raman-active molecular vibration of pigment components.

In addition, model transfer method was applied to the master and slave spectrometers in order to obtain a model that can be available both in vivo or in the field and with high prediction accuracy. This is the first attempt to establish a connection between the two types of instruments, and achieved good results. The foliar map scanning spectra after DS was brought to the pigments determination model established based on master instrument, and the concentration of photosynthetic pigment of each pixel in tea leaves could be predicted. Calculating the R^2 between the predicted value and the actual value, the range was in the range of 0.752-0.866. The predicted chlorophyll and carotenoid concentration of each pixel were imaged to obtain the distribution map of photosynthetic pigment in tea leaves, illustrating that the concentration of photosynthetic pigment in the central vein and leaf margin was lower than other parts was obtained, which is consistent with other studies. Therefore, the technology of Raman spectroscopy can be used to imaging the concentration of photosynthetic

pigment in different leaf positions, which provided a technological basis for the effective detection of the growth and nutrient distribution of tea tree, as well as diseases, pests and heavy metal stress.

It is worth noting that we have successfully improved the applicability of the Raman spectral model for determination of photosynthetic pigments in tea leaf. Through the calibration model transfer, the tea pigment spectral detection model based on the laboratory spectrometer was successfully applied to the portable quantitative detection of leaf pigment in the field. The model transfer method can effectively eliminate the spectral variation between the master and slave Raman spectrometers and improve the applicability of the spectral model, which will greatly promote the application process of the nondestructive and fast spectra measurement technique.

Abbreviations

DS: direct standardization; Chl-a: chlorophyll a; Chl-b: chlorophyll b; Car: carotenoids; MSC: multiplicative scatter correction; WT: wavelet transform; SNV: standard normal variate; RCF: rolling-circle filter; airPLS: adaptive iteratively reweighted penalized least squares; CARS: competitive adaptive re-weighted algorithm; PLSR: partial least squares regression; R^2: coefficient of determination; RMSE: root mean square error; KS: Kennard-Stone.

Acknowledgements

No applicable

Authors' Contribution

XL* and HZ designed the overall scheme, performed the experiment of Raman spectra acquirement, and advised on the initial design of and oversaw the manuscript. JZ and SZ determined the concentration of photosynthetic pigment, WL and JS made major contributions in data processing. YH and XL3 wrote the manuscript. BZ advised on schematic drawings. All authors read and approved the final manuscript.

Funding

This research was funded by the National natural science foundation of China (61565005, 41867020, 31771676), Jiangxi provincial department of science and technology project (20181BBF60024, 20181BBF68010, GJJ190315), the National key R & D plan project (Project No: 2018YFD0700501-01), and Zhejiang Province Public Technology Research Program (Project No: 2017C02027).

Availability of data and materials

All the seed material and raw imaging data can be obtained from the authors upon request.

Ethics approval and consent to participate

No applicable

Consent for publication

All authors consent for the publication.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]College of electrical and automation engineering, East China Jiaotong University, Nanchang 330013, China. [2]College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310058, China. [3]College of civil engineering and architecture, East China Jiaotong University, Nanchang 330013, China.

References

1. Sang S, Lambert JD, Ho CT, Yang CS. The chemistry and biotransformation of tea constituents. Pharmacol Res. 2011;64:87-99.
2. Ošt' Ádalová M, Tremlová B, Straka I, Pokorná J, Čáslavková P. Evaluation of significant pigments in green teas of different origin. Potravinarstvo Sci J Food Ind. 2014;8:1-8.
3. Shi JY, Zou XB, Zhao JW, Wang KL, Chen ZW, Huang XW, et al. Nondestructive diagnostics of nitrogen deficiency by cucumber leaf chlorophyll distribution map based on near infrared hyperspectral imaging. Sci Hortic. 2012;138:190-7.
4. Liu YL, Li W, Wu GF, Xu XG. Feasibility of estimating heavy metal contaminations in floodplain soils using laboratory-based hyperspectral data—a case study along le'an river, china. Geo-spatial Information Sci. 2011;14:10-6.
5. Tserevelakis GJ, Tsagkaraki M, Zacharakis G. Hybrid photoacoustic and optical imaging of pigments in vegetative tissues. J Microsc. 2016;263:300-6.

6. Zou XB, Shi JY, Hao LM, Zhao JW, Mao HP, Chen ZW, et al. In vivo noninvasive detection of chlorophyll distribution in cucumber (Cucumis sativus) leaves by indices based on hyperspectral imaging. Anal Chim Acta. 2011;706:105-12.
7. Zhao YR, Li X, Yu KQ, Cheng F, He Y. Hyperspectral imaging for determining pigment contents in cucumber leaves in response to angular leaf spot disease. Sci Rep. 2016;6.
8. Zhang C, Wang Q, Liu F, He Y, Xiao Y. Rapid and non-destructive measurement of spinach pigments content during storage using hyperspectral imaging with chemometrics. Measurement. 2017;97:149-55.
9. Yu KQ, Zhao YR, Zhu FL, Li XL, He YD. Mapping of chlorophyll and SPAD distribution in pepper leaves during leaf senescence using visible and near-infrared hyperspectral imaging. Trans ASABE. 2016;59: 13-24.
10. Polder G, van der Heijden GWAM, van der Voet H, Young IT. Measuring surface distribution of carotenes and chlorophyll in ripening tomatoes using imaging spectrometry. Postharvest Biol Tec. 2004;34:117-29.
11. Zhao JW, Wang KL, Ou YQ, Chen QS. Measurement of chlorophyll content and distribution in tea plant's leaf using hyperspectral imaging technique. Spectrosc Spec Anal. 2011;31:512-5.
12. Schulz H, Baranska M, Baranski R. Potential of NIR-FT-Raman spectroscopy in natural carotenoid analysis. Biopolymers. 2010;77:212-21.
13. Qin J, Chao K, Kim MS. Investigation of Raman chemical imaging for detection of lycopene changes in tomatoes during postharvest ripening. J Food Eng. 2011;107:277-88.
14. Yang G, Wang Q, Liu C, Wang X, Fan S, Huang W. Rapid and visual detection of the main chemical compositions in maize seeds based on Raman hyperspectral imaging. Spectrochim Acta Part A. 2018;200: 186-94.
15. Baranska M, Schütze W, Schulz H. Determination of lycopene and beta-carotene content in tomato fruits and related products: comparison of FT-Raman, ATR-IR, and NIR spectroscopy. Anal Chem. 2006;78:8456-61.
16. Bhosale P, Ermakov IV, Ermakova MR, Gellermann W, Bernstein PS. Resonance Raman quantification of nutritionally important carotenoids in fruits, vegetables, and their juices in comparison to high-pressure liquid chromatography analysis. J Agric Food Chem. 2004;52:3281-5.

17. Dane B, Darko D, Svjetlana L, Ksenija M, Charlotte VT, Buijnsters JG, et al. Correlation oftrans-lycopene measurements by the HPLC method with the optothermal and photoacoustic signals and the color readings of fresh tomato homogenates. Food Biophys. 2010;5:24-33.

18. Kalivas JH, Ottaway J. Feasibility study for transforming spectral and instrumental artifacts for multivariate calibration maintenance. Appl Spectrosc. 2015;69:407-16.

19. Ji W, Rossel RAV, Shi Z. Accounting for the effects of water and the environment on proximally sensed VIS–NIR soil spectra and their calibrations. Eur J Soil Sci. 2015;66:555–65.

20. Wang JX, Qu J, Li H, Han X, Xu G. Application of GA-DS to calibration transfer of aviation fuel density in near infrared spectroscopy. Liquid Fuels Tech. 2012;30:1975-80.

21. Sebben JA, Da JSE, Ranzan L, Fernandes NDM, Trierweiler LF, Trierweiler JO. Development of a quantitative approach using Raman spectroscopy for carotenoids determination in processed sweet potato. Food Chem. 2018;245:1224-31.

22. Wu M, Chen L, Huang X, Zheng Z, Qiu B, Guo L, et al. Rapid authentication of Pseudostellariaheterophylla (Taizishen) from different regions by Raman spectroscopy coupled with chemometric methods. J Lumin. 2018;202:239-45.

23. Sun W, Zhang X, Sun X, Sun Y, Cen Y. Predicting nickel concentration in soil using reflectance spectroscopy associated with organic matter and clay minerals. Geoderma. 2018;327:25-35.

24. Li XL, Jin JJ, Sun CJ, Ye DP, Liu YF. Simultaneous determination of six main types of lipid-soluble pigments in green tea by visible and near-infrared spectroscopy. Food chem. 2019;270:236-42.

25. Alam TM, Alam MK, Mcintyre SK, Volk DE, Neerathilingam M, Luxon BA. Investigation of chemometric instrumental transfer methods for high-resolution NMR. Anal Chem. 2009;81:4433-43.

26. Bouveresse E, Massart DL. Standardisation of near-infrared spectrometric instruments: a review. Vib Spectrosc. 1996;11:3-15.

27. Ziegel ER. A user-friendly guide to multivariate calibration and classification. Technometrics. 2004;46: 108-10.

28. Oh SK, Yoo SJ, Jeong DH, Lee JM. Real-time estimation of glucose concentration in algae cultivation system using Raman spectroscopy. Bioresource Technol. 2013;142:131-7.

29. Baranska M, Schulz H, Rosch P, Strehle MA, Popp J. Identification of secondary metabolites in medicinal and spice plants by NIR-FT-Raman microspectroscopic mapping. Analyst. 2004;129:926-30.

30. Schrader B, Schulz H, Klump HH. Non-destructive NIR-FT-Raman analyses of plants. J Mol Struct. 1999;509:201-12.

31. Zhou CL, Diers JR, Bocian DF. Qy -Excitation resonance Raman spectra of chlorophyll a and related complexes. Normal mode characteristics of the low-frequency vibrations. J Phys Chem B. 1997;101: 9635-44.

32. Picaud T, Le Moigne C, de Gracia AG, Desbois A. Soret-Excited Raman spectroscopy of the spinach cytochrome b6f complex. Structures of the b- and c-type hemes, chlorophyll a, and β-carotene. Biochemistry. 2001;40:7309-17.

33. Vitek P, Novotna K, Hodanova P, Rapantova B, Klem K. Detection of herbicide effects on pigment composition and PSII photochemistry in Helianthus annuus by Raman spectroscopy and chlorophyll a fluorescence. Spectrochim acta A. 2017;170:234-41.

34. Cai ZL, Zeng HP, Chen M, Larkum AWD. Raman spectroscopy of chlorophyll d from Acaryochloris marina. BBA-Bioenergetics. 2002;1556:89-91.

35. Vrettos JS, Stewart DH, de Paula JC, Brudvig GW. Low-temperature optical and resonance Raman spectra of a carotenoid cation radical in photosystem II. J Phys Chem B. 1999;103:6403-6.

36. Zhang YY, Gao WJ, Cui CJ, Zhang ZZ, He LL, Zheng JK, et al. Development of a method to evaluate the tenderness of fresh tea leaves based on rapid, in-situ Raman spectroscopy scanning for carotenoids. Food Chem. 2020;308.

37. Vicente R, Morcuende R, Babiano J. Differences in Rubisco and chlorophyll content among tissues and growth stages in two tomato (LycopersiconesculentumMill.) varieties. Agron Res. 2011;9:501-7.