

# Rapid Identification and Characterization of Recovered Edible Oil, Based on Raman and Near-Infrared Spectroscopy

Yang Chen, Qingsong Luo, Jie Wang and Xiao Zheng\*

School of Mechanical Engineering, Wuhan Polytechnic University, Wuhan, Hubei of China 430023

\*Corresponding author

**Abstract**—A variety of recovered-edible-oil identification models were established by using Raman combined with near-infrared spectroscopy (NIR). Eight types of 156 edible vegetable oil samples were collected to acquire their Raman and NIR spectra. The spectral data were processed for modeling. The preprocessing methods for the Raman spectra included the moving average method (11 points), adaptive iterative reweighted-penalty least squares method, and the normalization method based on the intensity of the characteristic peak at 1454 cm<sup>-1</sup> (MA11-airPLS-Nor). The preprocessing method for the NIR spectra was the standard normal variable transformation algorithm combined with a detrending technique (SNV\_DT). The Raman and NIR spectra were fused at the feature level by using independently the serial fusion and wavelet fusion approaches. The results showed that with the serial-fusion- and wavelet-fusion-based models, the identification of recovered oils can be achieved very rapidly. Furthermore, the comprehensive performances of the models based on fused Raman and NIR data were better than those of models based on separate Raman or NIR data.

**Keywords**—vegetable oil; spectral fusion; Raman spectroscopy; near-infrared spectroscopy

## I. INTRODUCTION

Vegetable oils provide essential nutrients for our daily diet[1]. With the growth of the edible oil market in China, some unethical providers are willing to mix recovered oil with genuine fresh oil to increase their profit. This behavior harms the health and interests of consumers. Therefore, the identification of edible oils is vital to maintaining order in the edible oil market. Traditional physical and chemical detection methods of identifying specific oils have disadvantages such as high cost, long evaluation times, and complex test operation procedures[2-4]. Spectral analysis technology has shown great potential in this application but relatively few viable detection approaches have been reported. Thus, it is of great significance to develop a spectral analysis method that will allow the rapid identification and characterization of recovered oils[5].

## II. EXPERIMENTAL

### A. Samples

A total of 156 edible oil samples, including soybean oil, peanut oil, rapeseed oil, rice oil, corn oil, sunflower oil, tea-seed oil and olive oil, were obtained from supermarkets, laboratories, and factories in Wuhan City. A total of 72 recovered oil samples were acquired from restaurants and

home-made oils. The output classification of genuine edible oil samples and recovered oil samples was set to be 1 and 0, respectively. The type and number of these samples are shown in Table 1.

TABLE I. TYPE AND NUMBER OF OIL SAMPLES

Type	Output of classification	Number of samples	Type	Output of classification	Number of samples
soybeanoil	0	25	sunfloweroil	0	14
rapeseedoil	0	33	riceoil	0	19
Tea-seedoil	0	27	cornoil	0	16
peanutoil	0	24	oliveoil	0	8
recoveredoil	1	72	—	—	—

### B. Instruments and Software

The Raman spectra were acquired with a RamTraceer-200 laser Raman spectrometer (OptoTrace Technologies, Inc., China). The laser wavelength was 785 nm, the resolution was ≤ 8 cm<sup>-1</sup>, the wavenumber studied was in the range of 250–2340 cm<sup>-1</sup> and the maximum laser power was 320 mW. The NIR spectra were acquired by using a laser near-infrared vegetable-oil detector developed by the authors' research group. The spectra were scanned 32 times in the range of 1350–1800 nm. The resolution was 3.5 cm<sup>-1</sup>, the wavelength repeatability was 0.01 nm, and the signal-to-noise ratio (250 ms, RMS) was higher than 5500: 1. The instrument was maneuverable and could be connected to the motherboard of an industrial computer, which facilitated the acquisition of the spectra.

### C. Acquisition of Spectra

The spectral acquisition procedure is presented as follows: Under the conditions of constant room temperature (25°C), humidity, and illumination intensity, an oil sample was heated to 60°C using a thermostatic water bath. The preheated oil sample was transferred into a 2-mm cuvette and then was placed in the water bath for heating for 1 min. Then, the spectrum was acquired. After that, the sample was replaced twice with the same oil for repeated spectral scanning. The three spectra thus acquired were averaged to derive the target spectrum for data preprocessing and modeling. The original Raman and NIR spectra acquired (after averaging) are exhibited in Figures 1 and 2.

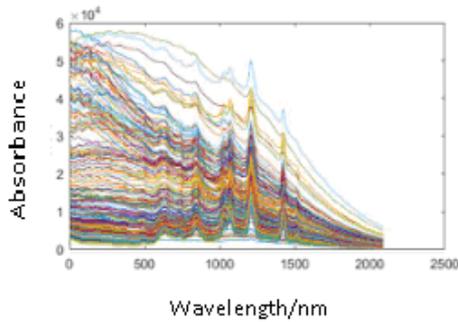


FIGURE I. RAMAN SPECTRA OF THE RECOVERED AND GENUINE VEGETABLE-OIL SAMPLES

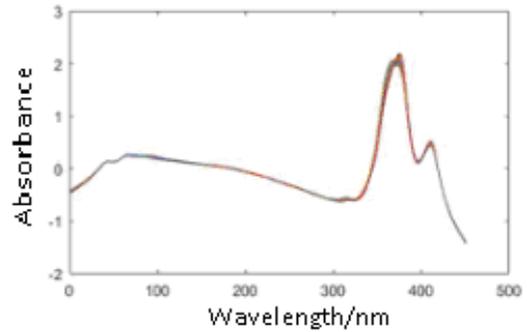


FIGURE IV. THE PREPROCESSED NIR SPECTRA

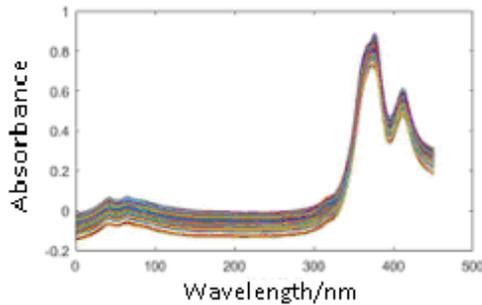


FIGURE II. NIR SPECTRA OF THE RECOVERED AND GENUINE VEGETABLE-OIL SAMPLES

D. Preprocessing of Spectral Data

The spectral instrument was susceptible to external factors such as room noise. Therefore, before the spectral analysis was undertaken, the spectra should be preprocessed to reduce or eliminate the influence of these factors during the establishment of the models. In this work, the Raman spectra were preprocessed separately using the moving average method (11 points), the adaptive iterative reweighted-penalty least square method, and the normalization method, based on the intensity of the characteristic peak at 1454  $\text{cm}^{-1}$  (MA11-airPLS-Nor)[6]. The preprocessed Raman spectra are shown in Figure 3. Additionally, the NIR spectra were preprocessed following a standard normal variable transformation algorithm combined with a detrending technique (SNV\_DT)[7]. The preprocessed NIR spectra are shown in Figure 4.

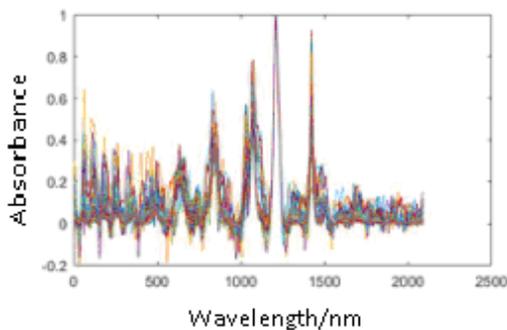


FIGURE III. THE PREPROCESSED RAMAN SPECTRA

During the present investigation, the Raman and NIR spectra were fused separately at the feature level with serial fusion and wavelet fusion approaches. In detail, the serial fusion transformed the Raman and NIR spectra at the feature level into the same coordinate system. Then, the dimensions of a feature were reduced by using feature extraction algorithms such as CARS, SPA, BiPLS and so forth. On the other hand, in the feature-based wavelet transform fusion method, on the basis of the evident features of both images, the wavelet coefficients at each location were selected according to the following fusion procedure: first, the high-resolution image ‘A’ and low-resolution multi-spectral image ‘B’ were transformed separately following tower-type wavelet transform, to derive the low-frequency approximate components a and b and high-frequency detailed component c. The three components were fused following fusion principles to derive the fused spectra [6 and 7], as shown in Figures 5 and 6. The fused image then was rebuilt through an inverse wavelet transform.

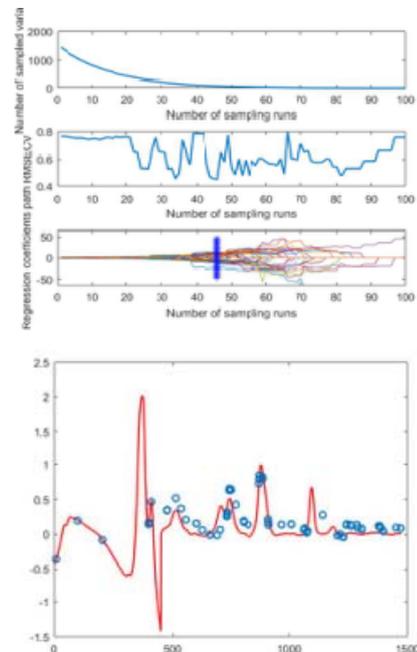


FIGURE V. CARS-OPTIMIZED SERIAL-FUSION SPECTRA

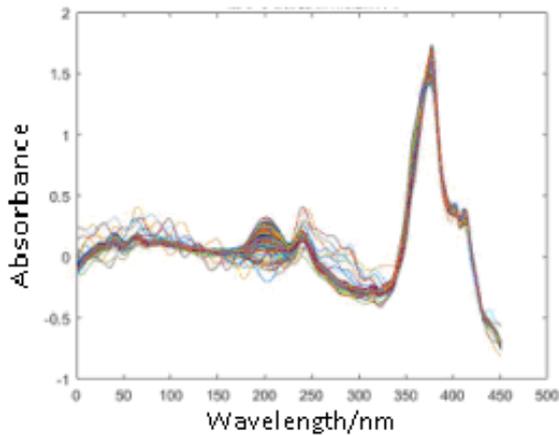


FIGURE VI. WAVELET-FUSION SPECTRA OF RECOVERED AND GENUINE VEGETABLE-OIL SAMPLES

E. Modeling Method

Support vector machine classification (SVC) is a common method used in qualitative identification analysis[8]. Linear modeling methods cannot extract spectral information effectively, which affects the accuracy of the models. In contrast, SVC can extract spectral information efficiently and therefore improve model accuracy. In addition, it has been proved experimentally that the penalty factor C and parameter g of the kernel function of the radial basis function (RBF) are the key factors that control the performance of SVC.

By using the SVC method, an NIR-Raman-DWT-SVC model was established with the wavelet-fusion-based Raman and NIR spectral data as input variables; an NIR-Raman-CARS-SVC model was established with the serial-fusion-based Raman and NIR spectral data as input variables. The prediction results of fusion-based models, the Raman-SVC-based model, and the NIR-SVC-based model are shown in Table 2.

TABLE II. PARAMETERS AND PREDICTION ACCURACY OF THE MODELS

Model	Parameter	Accuracy/%
Raman-SVC	(C,g)=(4,0.25)	100.00
NIR-SVC	(C,g)=(1,16)	98.24
NIR-Raman-DWT-SVC	(C,g)=(16,2)	100.00
NIR-Raman-CARS-SVC	(C,g)=(8,2)	100.00

Further, the grid search method (GS) was employed to optimize the parameters of the NIR-Raman-DWT-SVC and the NIR-Raman-CARS-SVC models. The parameter optimization and prediction results from these models are shown in Figures 7 and 8.

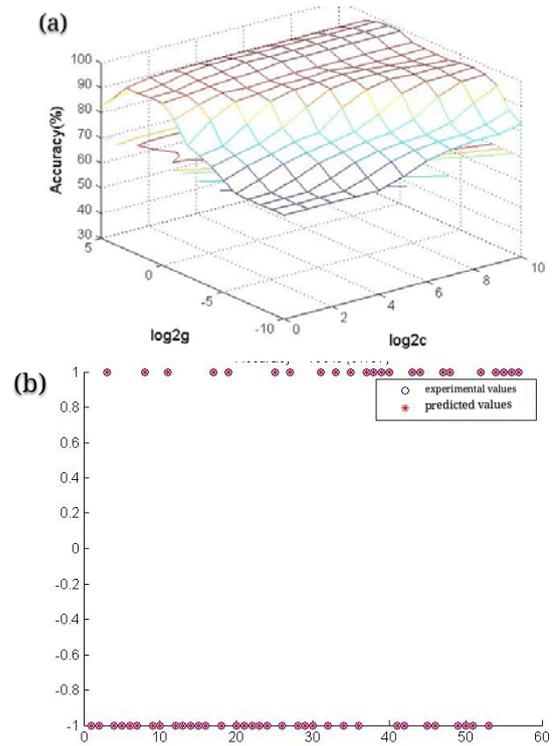


FIGURE VII. PARAMETERS OPTIMIZATION (A) AND PREDICTION RESULTS (B) OF THE NIR-RAMAN-DWT-SVC MODEL

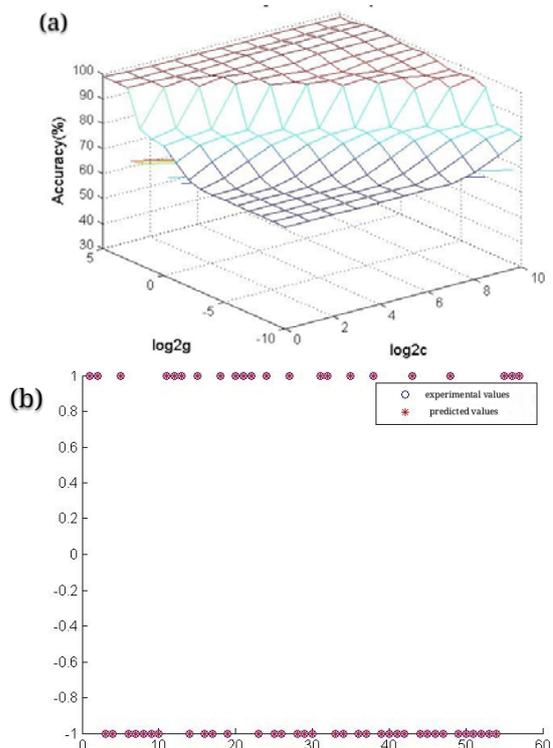


FIGURE VIII. PARAMETER OPTIMIZATION (A) AND PREDICTION RESULTS (B) OF THE NIR-RAMAN-CARS-SVC MODEL

### III. RESULTS AND DISCUSSION

Collectively, the data summarized in Table 2, Figure 7, and Figure 8 indicate that, by using a fusion-based NIR-Raman-DWT-SVC and NIR-Raman-CARS-SVC model, it is possible to obtain rapid and reliable identification and characterization details from recovered oil. The performance of the NIR-based SVC model was better than that of that based on Raman spectroscopy alone. This was attributable to the higher stability of the NIR spectra. Compared with the NIR-based SVC model, the prediction accuracy of the fusion-based NIR-Raman-CARS-SVC model was increased by 1.76%. The prediction accuracy of both fusion-based models approached 100%. Additionally, the parameters ( $C$ ,  $g$ ) of the NIR-Raman-CARS-SVC model reached a minimum, showing that serial fusion is effective in improving the performance of the model. These spectra-fusion-based detection models have high accuracy, prominent parameters, great generalization ability, and high stability.

### IV. CONCLUSIONS

Targeting the rapid identification of recovered oil, Raman and NIR spectroscopy were combined with stoichiometry, in order to establish models for the identification of recovered oil. This study focused on the serial and wavelet fusion of Raman and NIR spectra for the building and performance of the characterization models. In this way, near-immediate and reliable identification of recovered oils was realized. The performances of the fusion-based models were better than those of the individual Raman- or NIR-based models. The experimental results showed that the combination technology and fusion method proposed in this paper can facilitate prompt and valuable characterization of recovered oils.

### ACKNOWLEDGMENTS

The Authors would like to express appreciation for Grants: Innovation and Transformation of Grain Science and Technology of Hubei Province (20165104); Key Science and Technology Project of Wuhan City (2013010501010147), which supported this investigation.

### REFERENCES

- [1] Wang, R., "Present situation of edible vegetable oil consumption in China", *Journal of Heilongjiang Grain*, 2017, (5): 11-13.
- [2] Lu, H., Bi, Y., Zhao, Z., "Status of oil detection technology", *China Oils and Fats*, 2009, 34, (4): 1-6.
- [3] Yang, L., Yu, H., "Theory and Application of Multi-Source Information Fusion 2", Beijing University of Posts and Telecommunications Press, 2006.
- [4] Wang, Y., Li, S., "Multi-sensor Information Fusion and Its Application: A Survey", *Control and Decision*, 2001, 16, (5): 518-522.
- [5] Qi, S., "Quality determination of citrus fruit using electronic nose, electronic tongue, and fusion system", Zhejiang University, 2016.
- [6] Vol. N., "Machine Vision for Color Inspection of Potatoes and Apples", *Transactions of the ASAE*, 1995, 38, (5): 1555-1561.
- [7] Wen, Z., Tao, Y., "Building a rule-based machine-vision system for defect inspection on apple sorting and packing lines", *Expert Systems with Applications*, 1999, 16, (3): 307-313.
- [8] Li, C., Heinemann, P., Sherry, R., "Neural network and Bayesian network fusion models to fuse electronic nose and surface acoustic wave sensor data for apple defect detection", *Sensors & Actuators B Chemical*, 2007, 125, (1): 301-310.