10.1177/0003702818805116

An Improved Background-Correction Algorithm for Raman Spectroscopy Based on the Wavelet Transform

Mingbo Chi,[1] Xinxin Han,[1,2] Yang Xu,[1,2] Yue Wang,[1] Fengfeng Shu,[1] Wenchao Zhou[1], and Yihui Wu[1*]

* Correspondence to: Yihui Wu, email: yihuiwu@ciomp.ac.cn

[1]State Key Laboratory of Applied Optics, Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China

[2]University of Chinese Academy of Sciences, Beijing 100039, China

Abstract

In the traditional background correction algorithm based on the wavelet transform, approximation coefficients considered as frequency responses of background signal are usually set to zero. However, there are many meaningless negative values generated in the background corrected spectrum because of the calibration errors of this algorithm. Intensities of some weak peaks even become negative and these peaks will disappear after the calibration of negative values. To solve these problems for the background correction of Raman spectrum, an improved intelligent algorithm which utilizes a suppression coefficient to modify approximation coefficients is proposed in this paper. A series of simulation analyses, as well as experimental investigations, are made to test the performance of this algorithm. It is proved that the usage of suppression coefficient could increase the background correction accuracy and decrease the number of meaningless negative values in the reconstructed spectra, which will prevent the disappearance of weak Raman peaks after the calibration of negative values and increase the sensitivity of Raman spectral analysis.

Keywords: Raman spectroscopy, background correction, signal processing, wavelet transform

Introduction

Raman spectroscopy has attracted increasing attention in bio-imaging in which biological

components can be differentiated due to the differences in the position of peaks in Raman spectra as well as their relative intensities.[1–4] However, Raman spectroscopy, especially for biological samples, is often obscured by the background due to fluorescence from organic molecules and contaminations.[5] The intensity of the background is usually much larger than that of the weak Raman signal. In order to simplify the interpretations of Raman spectrum, it is essential to implement proper background correction before spectral analysis.

Some experimental techniques can be used to remove the background signal. Excited between 220 and 250 nm, Raman emission occurs within a fluorescence-free region of the spectrum, eliminating the obscuration of weak Raman signals by fluorescence from target or surrounding materials.[6] In order to reduce the magnitude of fluorescence, Pelletier and Altkorn[7] used liquid-core optical fiber waveguides to collect Raman signals. However, these experimental methods are expensive and complicated to implement. Significant efforts have been directed towards removing the background by signal processing with high accuracy. Signal processing algorithms for background correction in Raman spectroscopy can be categorized into two groups. The first group is mainly based on polynomial fitting[8–9] and utilizes information about background shape, position, and signal-to-noise ratio (SNR) of the signal. The manual polynomial fitting is not so effective and its accuracy depends on the user's experience.[8] The performance of automatic polynomial fitting is poor when the SNR of the spectrum is low.[9] Moreover, a major limitation with this group of methods is that the noise in the signal makes de-noising and smoothing an inevitable process before background correction.

The second group of methods relies on the difference in frequency response of background compared with characteristic peaks. Wavelet transform (WT) [10–12] which can provide multiresolution analysis is a powerful tool for these methods. After wavelet transform, the signal can be decomposed into localized contributions characterized by approximation coefficients and detail coefficients. Each contribution corresponds to a single frequency component. The background can be seen as low-frequency components of the signal while the noise and spectral peaks correspond to high-frequency components. The pure

spectral signal can be reconstructed through wavelet inverse transform by setting the approximation coefficients which correspond to low-frequency components to zero. This method requires no prior knowledge about the spectrum and no mathematical assumption of the background distribution.[13–16] However, there are significant spectral intensity decrease and a lot of meaningless negative values in the reconstructed spectrum. Even intensities of some weak peaks become negative. Generally, negative values are substituted by zero in the reconstructed signal.[16] Unfortunately, we found that these weak peaks would disappear after background correction using this method. In the background correction for Raman spectrum, the intensity of which is usually much weaker than the background signal, an improved algorithm is needed to deal with the negative values and preserve the weak peaks.

In this paper, an improved intelligent algorithm which can decrease the number of negative values and calibrate the intensity errors in the reconstructed spectrum is proposed. In this algorithm, original spectrum is decomposed based on WT and a suppression coefficient used to modify approximation coefficients is calculated based on spectral energy analysis. Simulated spectra with different kinds of background are used to validate the performance of the proposed algorithm. This algorithm is also successfully applied in the background correction of an experimental Raman spectrum of rat's liver cell. The results show that weak Raman peaks are well preserved in the removal of the background signal using this improved algorithm. Additionally, no prior knowledge about the spectral background distribution and smoothing step are needed in this background correction algorithm which makes this algorithm a powerful and intelligent tool for Raman spectroscopy.

Theory

Similar to the Fourier transform (FT) which uses sine and cosine as the basis function, WT utilizes a family of basis functions which have a zero-mean oscillation behavior. Considered as the convolution between the signal and a wavelet function, WT decomposes a signal into localized contributions (approximations and details) labeled by scale and shifting parameters. A mother wavelet could produce families of wavelets through Eq. 1:

$$\psi_{ab}(T) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) (1)$$

where $\psi(t)$ is the mother wavelet, a is the scale parameter, b is the shifting parameter or the position of the scale window, and $\psi_{ab}(t)$ is the scaled and shifted wavelet. The wavelet coefficient $W(a,b)$ which corresponds to the localized contributions of the signal $f(t)$ is defined as

$$W(a,b) = \int_{+\infty}^{-\infty} f(t)\psi_{a,b}(t)dt \quad (2)$$

As shown in Eq. 2, the wavelet coefficient can be considered as a weighted sum of the original signal at corresponding scale. The multiresolution wavelet transform algorithm by Mallat,[17] which is usually performed for discrete wavelet transform, makes the signal pass through a low-pass filter to implement large scale analysis and pass through a high-pass filter to implement small scale analysis as shown in Figure 1a. Then we can get a group of approximation coefficients (CAn) that represent the estimation of the signal at large scale and a group of detail coefficients (CDn) that contain the detail information of the signal. The group of approximation coefficients can be further decomposed to form a new group of approximation coefficients and a new group of detail coefficients. Then the original signal is decomposed into orthogonal components at different scales. The features of the signal are captured by detail coefficients at proper decomposition level, and the smooth components of the signal are mainly captured by approximation coefficients. Thus the wavelet transform can be used to separate the pure spectral signal from the background in Raman spectrum and implement the background correction. Generally, the approximation coefficients at proper decomposition level are directly set to zero (c=0) and then the spectrum is reconstructed just using detail coefficients to remove the background signal in the spectrum. However, in this approach, the intensity of the reconstructed spectrum will decrease and there are a lot of negative values generated in the reconstructed spectrum.

In our opinion, just setting approximation coefficients to zero will lead to the over-treatment of the original signal. The approximation coefficient is calculated via the inner product of the original signal and the scale function, and it can be seen as a weighted estimation of the whole signal at the large scale which is determined by the decomposition

level. Though the detail feature of the spectrum is smoothed out when the background signal is reconstructed just using the approximation coefficients, it does not mean that the detail features of the signal have no contribution to this reconstructed signal as well as the approximation coefficients. As shown in Figure1b, the simulated spectrum with a curved simulated background is decomposed based on wavelet transform and then the reconstructed background signal can be obtained by just using the approximation coefficients in the wavelet inverse transform. The intensity of the reconstructed background is larger than that of the original simulated background though the two curves are similar. This proves that the detail features of the original spectrum indeed contribute to the approximation coefficient and part of the approximation coefficient is essential for the reconstruction of the pure spectral signal. In addition, though the detail features of the spectrum will be reserved, the background correction by discarding the approximation coefficients and just using the detail coefficients to reconstruct the spectrum through wavelet inverse transform will cause an overestimation of the background and the intensity at each pixel of the reconstructed spectrum will decrease. A simulated illustration is made as shown in Figure 2. All the simulation and experiment computation in the current study were carried out utilizing Matlab (The MathWorks Inc.) for Microsoft Windows, v.7.9 (R2009b). In this study, the simulated spectrum is the sum of background, Raman signals and random noise. The pure Raman signal $p(x)$ is composed of six narrow Gaussian peaks, which can be mathematically described as follows:

$$p(x) = 40e^{-\frac{(x-200)^2}{2\times20^2}} + 5e^{-\frac{(x-500)^2}{2\times8^2}} + 27e^{-\frac{(x-800)^2}{2\times30^2}} + 100e^{-\frac{(x-1300)^2}{2\times8^2}} + 3e^{-\frac{(x-1500)^2}{2\times12^2}} + 50e^{-\frac{(x-1800)^2}{2\times8^2}} \quad (3)$$

White Gaussian noise was added to pure Raman signal $p(x)$ using the Matlab function AWGN($p(x)$, b), where b is the SNR of the Raman signal with random noise. The SNR is defined as the ratio of signal power to the noise power, often expressed in dB as shown in Eq. 4.

$$SNR_{dB} = 10\log_{10}\frac{P_{\text{signal}}}{P_{\text{noise}}} \quad (4)$$

P_{signal} is average power of the pure Raman signal and P_{noise} is average power of noise signal.

Curved background $b_c(x)$ can be mathematically described as follows:

$$b_c(x) = 150e^{-\frac{(x-1100)^2}{2\times1200^2}}+100 \quad (5)$$

Figure 2a shows the simulated pure Raman spectrum consisted of Gaussian peaks and random noise and its SNR is 150 dB. Figure 2b shows the simulated spectrum with curved background. When the spectrum is reconstructed just using the detail coefficients through wavelet inverse transform, a new baseline consisted of negative values will appear and the intensities of some weak peaks also become negative as shown in Figure 2c. As the negative value is meaningless in spectroscopy, it is usually substituted by zero. As shown in Figure 2d, the second and fifth peaks with low intensity disappear after the calibration of negative values in the reconstructed spectrum, which will cause the loss of useful spectral information and the sensitivity of subsequent spectral analysis.

In order to remove the negative values caused by the overestimate of the background, we introduce a new coefficient called suppression coefficient. The suppression coefficient is calculated based on spectral energy analysis. To accurately remove the background, it is necessary to keep part of the approximation coefficients contributed by the detail features of the pure spectrum in the background correction. Because of the orthogonality of the wavelet basis,[10,18] the quadratic sum of the coefficients in the wavelet transform vector can be used to estimate the energy of the original signal as shown in Eq. 6.

$$\sum_{a,b}|W(a,b)|^2 = \mathrm{A}\int dt|f(t)|^2 \quad (6)$$

A is a constant. The quadratic sum of the detail coefficients in the vector can be also used to estimate the energy of the pure spectral signal. Then the proportion of the energy of the pure spectral signal in the original signal can be obtained by the ratio of the above two quadratic sums. In order to increase the calibration accuracy, part of the approximation coefficients contributed by the pure spectral signal should be reserved when correct the background based on WT. Therefore, a suppression coefficient defined as the square root of the proportion of the energy of the pure spectral signal in the original signal is used in the improved algorithm. As shown in Figure 1a, the original spectrum is decomposed by wavelet transform at proper decomposition level. Then multiply each of the approximation coefficients by the calculated

suppression coefficient and the pure spectral signal will be reconstructed through wavelet inverse transform with high accuracy. The detailed procedure follows the steps described below.

(i) Choose a wavelet and decomposition level L. Compute the wavelet decomposition of the original signal at Level L and get the vector of approximation coefficients (CA_i) and the vector of detail coefficients (CD_j).

(ii) Calculate the suppression coefficient c using Eq. 7.

$$c = \sqrt{\frac{\sum_j CD_j^2}{\sum_i CA_i^2 + \sum_j CA_j^2}} \quad (7)$$

(iii) Multiply each of the approximation coefficients by the suppression coefficient. Reconstruct the spectral signal using the modified approximation coefficients and the original detail coefficients. In addition, some detail coefficients which correspond to the noise and are less than a proper threshold can be discarded after the calculation of the suppression coefficient,[15,19] and then the de-noising and background correction of the original signal can be implemented at the same time just through one wavelet inverse transform.

The improved background correction algorithm is implemented to deal with the simulated spectrum with curved background shown in Figure 2b. The approximation coefficients after wavelet transform are multiplied by the calculated suppression coefficient as shown in Eq. 7. Then the reconstructed spectrum by the inverse wavelet transform is shown in Figure 2e and the corresponding negative value calibrated spectrum is shown in Figure 2f.Compared with the background corrected spectrum shown in Figure 2c, the usage of the suppression coefficient can increase the reconstruction accuracy of the peak intensity and prevent the reconstructed spectrum from the affection of the meaningless negative values. In the reconstructed spectrum which discards the approximation coefficients, more than 80% of the spectral data becomes negative whereas in the spectrum reconstructed by the improved algorithm, the percentage of negative values is less than 20%. In addition, the second and

fifth weak peaks are well preserved after the calibration of the negative values as shown in Figure 2f.

In the traditional algorithm which discards the approximation coefficients, the influence of the pure spectral peaks on the approximation coefficients of the wavelet transform is neglected, which is the main cause of the reconstructed errors and the appearance of negative values. Thus the more peaks the pure spectrum contains, the larger errors will be generated in the reconstructed spectrum obtained with the algorithm which discards the approximation coefficients. In the improved algorithm, the essential information for the reconstruction of the pure spectral signal contained in the approximation coefficients is well estimated and reserved with the help of the suppression coefficient. Thus, the pure spectral signal has less effect on the reconstructed errors of the improved algorithm for a given background. To make a contrast experiment, the above algorithms are implemented on a series of simulated spectra consisted of the same flat background and different number of Gaussian peaks respectively. The differences between the reconstructed spectrum y_t' and the simulated pure spectral signal y_t are measured through the computation of the root mean square (RMS) values of the reconstructed errors, which are defined in Eq. 8. T is the length of y_t' and y_t.

$$\mathrm{RMS} = \sqrt{\frac{\sum_{t=1}^{T}\left(y_t' - y_t\right)^2}{T}} \quad (8)$$

The computation results are shown in Figure 3. As we can see, the reconstructed error obtained by just discarding the approximation coefficients increases linearly with the number of the peaks while the reconstructed error of the improved algorithm hardly changes with the number of peaks. So it is impossible to calibrate the reconstructed errors by adding a positive constant to the Raman spectrum obtained with the traditional algorithm. Moreover, the correction accuracy of the improved method is hardly affected by the pure spectral signal, so this improved algorithm can be used as an intelligent background correction tool for different kinds of Raman spectra.

To investigate the robustness of the improved algorithm, another three kinds of background signal (flat, linear, and sigmoidal background) are added to the Raman spectrum

shown in Figure 2a, and the SNR of the Raman spectrum is also changed by adding white Gaussian noise with different power. The simulated original spectrum and the background corrected spectrum obtained with the improved algorithm are shown in Figure 4. The flat background in Figure 4a and d can be described using a constant while the linear background signal $b_L(x)$ in Figure 4b and e is mathematically described as follows:

$$b_L(x) = 0.04x + 100 \quad (9)$$

The sigmoidal background signal $b_S(x)$ in Figure 4c and f is also mathematically defined in Eq. 10:

$$b_s(x) = \frac{200}{1+e^{-0.006(x-1000)}} + 200 \quad (10)$$

As shown in the contrast experiments, the three kinds of background are all well eliminated while the second and fifth peaks with low intensity are also well preserved in the reconstructed spectra. This indicates that the improved background correction algorithm proposed in this paper can deal with different kinds of background signal and no prior knowledge about the background distribution is needed in the background correction process. In addition, this improved background correction algorithm can achieve good performance both in high SNR situation and low SNR situation. The RMS values of the errors between the reconstructed spectra and the corresponding simulated Raman spectra are calculated and they are almost the same when the simulated Raman spectra consist of the same background but different SNRs. This is mainly because that the improved algorithm is based on wavelet transform and relies on the difference in frequency response of background compared with characteristic peaks and noises. The power increase or decrease of white Gaussian noise will not change its frequency response. So the improved algorithm is immune to the effects of noise level.

To make a further contrast experiment, we implemented both the improved wavelet transform algorithm in this paper and the iterative polynomial fitting algorithm[9] on a group of simulated Raman spectra with the same curved background and different SNR. The simulated pure Raman signal is defined in Eq. 3 and the baseline or the background signal in the simulation is mathematically defined in Eq. 5. White Gaussian noise signal with different

power is added to the simulated Raman signal through MATLAB function AWGN(). The experiment results are shown in Figure 5. As shown in Figure 5a, the RMS values of the reconstructed errors using the improved wavelet transform algorithm hardly change with the SNR of the simulated Raman spectrum while the iterative polynomial algorithm performs not so well in low SNR situations. In addition, the processing speed of the improved wavelet transform algorithm is faster. As shown in Figure 5b, the average running time of the improved WT algorithm is about a quarter of that of the iterative polynomial algorithm when the SNR of the simulated Raman spectrum is 190 dB and both of the two algorithms have almost the same RMS value. Therefore, the improved wavelet transform algorithm does not need to consider the effects of noise and takes less computing time, which is favored by real-time applications.

In order to choose the proper decomposition level and wavelet type, a series of simulated experiments are made. In these simulated experiments, the improved algorithm is used to deal with the simulated spectra consisted of Raman signal shown in Figure 2a and different kinds of typical background (curved, flat, linear, or sigmoidal background) by changing decomposition levels and wavelet types. The RMS values of the reconstructed errors are calculated and used to evaluate the performance of the algorithm. The relationship between the RMS values and decomposition level as well as wavelet type is shown in Figure 6. The experiment results show that the improved algorithm in this paper can obtain good performance in the background correction for all the four kinds of background when the original spectra are decomposed with coif5 wavelet at level 9.

Results

The Raman spectrum of rat liver cell was collected using the Raman spectral Imaging system built in our lab. The excitation radiation, produced by a He–Ne laser (532 nm), was directed by an Olympus microscope and focused on the sample by a 50X microscope objective. Rayleigh scattering was rejected by a Notch filter and the Raman signal was focused on the slit of the spectrometer. Then the input signal was dispersed by a grating and detected by a

scientific CCD detector. The resolution of the spectrometer is about 2 cm^{-1} and the Raman spectra covering the 600 cm^{-1} to 1100 cm^{-1} spectral range were recorded. The specimen is driven by a scanning stage with high precision to realize the Raman imaging in a 60 μm × 60 μm area.

The original experimental Raman spectrum is shown in Figure 7 as well as the background corrected spectra obtained with the algorithm which discards the approximation coefficients and the improved algorithm which utilizes the suppression coefficient. As we can see, there is a significant fluorescence background in the original Raman spectrum, which will cause a decrease of the SNR of the Raman signal. Then the original spectrum is decomposed by wavelet transform at level 9 using the coif5 wavelet. In the reconstructed spectrum obtained through wavelet inverse transform when the approximation coefficients are set to zero, about 80% of the data in the reconstructed spectrum is negative and the intensities of some weak spectral peaks (644 cm^{-1}, 719 cm^{-1}, 829 cm^{-1}, 853 cm^{-1}, and 936 cm^{-1}) become negative. The above peaks disappear after the meaningless negative values are substituted with zero as shown in Figure 7a. On the other hand, the suppression coefficient calculated via Eq. 7 is used to modify the approximation coefficients. After the wavelet inverse transform, the background is well corrected and less than 15% of the data become negative in the reconstructed spectrum. None of the Raman peaks disappear after the calibration of the negative values as shown in Figure 7b. The experimental results are consistent with the above simulation results and the quality of background correction is significantly improved by the usage of the suppression coefficient.

As phenylalanine[20] is characterized by the spectral contribution at 1003 cm^{-1} and tyrosine is characterized by the spectral contribution at 644 cm^{-1}, Raman image for the distribution of phenylalanine and tyrosine in liver cells can be constructed using the spectral intensity at the above wavenumbers (1003 cm^{-1} and 644 cm^{-1}) of each scan point after the scan of the specimen. The traditional background correction algorithm which discards the approximation coefficients and the improved algorithm which utilizes the suppression coefficient are implemented on the spectrum of each scan point respectively to make a

contrast experiment. The meaningless negative values in the reconstructed spectra are all substituted by zero. Figure 8 shows the comparison of the Raman images constructed using the spectral data processed by the traditional algorithm which discards the approximation coefficients and the improved algorithm utilizing the suppression coefficient. For the spectral contributions at 1003 cm^{-1} which have relative high intensities and high SNR, the images for the distribution of phenylalanine which are constructed with the spectral data processed with the above two algorithms are similar but there is a significant intensity increase generated by the usage of the suppression coefficient. However, for the spectral contributions at 644 cm^{-1} which have weak intensities and low SNR, there are almost no Raman signals in the image constructed by spectral data processed with the traditional algorithm while the distribution of Raman signal at 644 cm^{-1} can be still recognized in the image constructed by the spectral data processed with improved algorithm. Thus it is well convinced that the usage of the suppression coefficient can prevent the decrease of the Raman peak intensity and the loss of weak Raman signal in the background correction, and then increase the accuracy and sensitivity of the Raman spectral analysis.

Conclusion

Benefited from the multiscale analysis implemented by the discrete wavelet transform, the essential information contained in the approximation coefficients for the reconstruction of the pure spectrum is well reserved by the usage of the suppression coefficient to modify the approximation coefficients in the background correction of Raman spectrum. It is proved by both the simulation and experimental results that there is a significant decrease in the reconstructed errors of the Raman peak intensities and the number of meaningless negative values in the background corrected spectrum obtained with improved algorithm which utilizes the suppression coefficient, and the Raman peak with weak intensity is well preserved which increases the sensitivity and accuracy of the Raman spectral analysis. Therefore, this improved algorithm can be used as an efficient and powerful tool for Raman spectral analysis and imaging.

Funding

This work was supported by the National Natural Science Foundation of China (Grant numbers 61605198, 6172781); the science and technology development project of Jilin Province, China (grant number 20180520230JH).

References

1. K. Hamasha, Q.I. Mohaidat, R.A. Putnam, R.C. Woodman, et al. “Sensitive and Specific Discrimination of Pathogenic and Nonpathogenic Escherichia coli Using Raman Spectroscopy—A Comparison of Two Multivariate Analysis Techniques". Biomed. Opt. Express. 2013. 4(4): 481–489.
2. A.F. Chrimes, A.A. Kayani, K. Khoshmanesh, P.R. Stoddart, et al. “Dielectrophoresis-Raman Spectroscopy System for Analyzing Suspended Nanoparticles". Lab Chip. 2011. 11(5): 921–928.
3. J. Qi, W. Shih. "Performance of Line-Scan Raman Microscopy for High-Throughput Chemical Imaging of Cell Population". Appl. Opt. 2014. 53(13): 2881–2885.
4. C.S. Liao, M.N. Slipchenko, P. Wang, J. Li, et al. "Microsecond Scale Vibrational Spectroscopic Imaging by Multiplex Stimulated Raman Scattering Microscopy". Light: Sci. Appl. 2015. 4: 265–272.
5. J. Kostamovaara, J. Tenhunen, M. Kogler, I. Nissinen, et al. “Fluorescence Suppression in Raman Spectroscopy Using a Time-Gated CMOS SPAD". Opt. Express. 2013. 21(25): 31632–31645.
6. R. Bhartia, W.F. Hug, R.D. Reid. "Improved Sensing Using Simultaneous Deep UV Raman and Fluorescence Detection". Proc. SPIE 9073, Chemical, Biological, Radiological, Nuclear, and Explosives (CBRNE) Sensing XV, 90730I. Baltimore, Maryland; 29 May 2014. doi: 10.1117/12.2053069.
7. M.J. Pelletier, R. Altkorn. "Efficient Elimination of Fluorescence Background from Raman Spectra Collected in a Liquid Core Optical Fiber". Appl. Spectrosc. 2000. 54(12):

1837–1841.

8. A. Jirasek, G. Schulze, M. Yu, M. Blades, et al. "Accuracy and Precision of Manual Baseline Determination". Appl. Spectrosc. 2004. 58(12): 1488–1499.

9. J. Zhao, H. Lui, D. McLean, H. Zeng. "Automated Autofluorescence Background Subtraction Algorithm for Biomedical Raman Spectroscopy". Appl. Spectrosc. 2007. 61(11): 1225–1232.

10. I. Daubechies. “The Wavelet Transform, Time-Frequency Localization and Signal Analysis". IEEE Trans. Inf. Theory. 1990. 36(5): 961–1005.

11. E. Bacry, J.F. Muzy, A. Arneodo. “Singularity Spectrum of Fractal Signals from Wavelet Analysis: Exact Results". J. Stat. Phys. 1993. 70(3–4): 635–674.

12. B.K. Alsberg, A.M. Woodward, D.B. Kell. “An Introduction to Wavelet Transforms for Chemometricians: A Time-Frequency Approach". Chemom. Intell. Lab. Syst. 1997. 37(2): 215–239.

13. Z.M. Zhang, S. Chen, Y.Z. Liang, Z.X. Liu, et al. "An Intelligent Background-Correction Algorithm for Highly Fluorescent Samples in Raman Spectroscopy". J. Raman Spectrosc. 2009. 41(6): 659–669.

14. P.M. Ramos, I. Ruisanchez. “Noise and Background Removal in Raman Spectra of Ancient Pigments Using Wavelet Transform” J. Raman Spectrosc. 2005. 36(9): 848–856.

15. B.F. Liu, Y. Sera, N. Matsubara, K. Otsuka, et al. “Signal Denoising and Baseline Correction by Discrete Wavelet Transform for Microchip Capillary Electrophoresis". Electrophoresis. 2003. 24(18): 3260–3265.

16. Y. Hu, T. Jiang, A. Shen, W. Li, et al. “A Background Elimination Method Based on Wavelet Transform for Raman Spectra". Chemom. Intell. Lab. Syst. 2006. 85(1): 94–101.

17. S. Mallat. A Wavelet Tour of Signal Processing. San Diego: Academic Press, 1999.

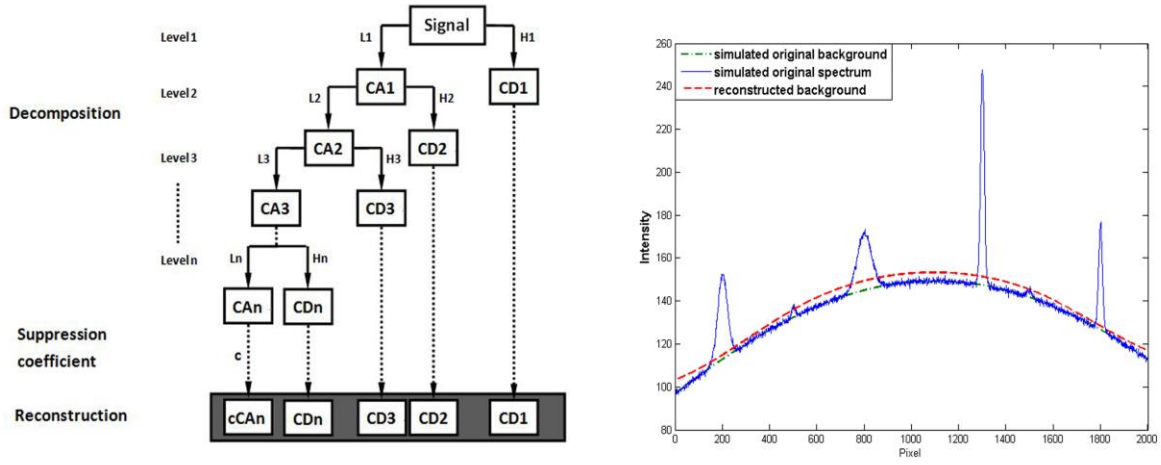
18. F. Zhao, J. Wang, A. Wang. “An Improved Spectral Background Subtraction Method Based on Wavelet Energy". Appl. Spectrosc. 2016. 70(12): 1994–2004.

19. D.L. Donoho. "De-Noising by Soft-Thresholding". IEEE Trans. Inf. Theory. 1995. 41(3): 613–627.

20. C. Krafft, M. Diderhoshan, P. Recknagel, M. Miljkovic, et al. “Crisp and Soft Multivariate Methods Visualize Individual Cell Nuclei in Raman Images of Liver Tissue Sections". Vib. Spectrosc. 2011. 55(1): 90–100.

Captions

Figure 1. (a) The Schematic description of wavelet transform and the improved algorithm. CAn is approximation coefficients, CD1–CDn are detail coefficients and c is the calculated suppression coefficient. (b) The comparison between the background which is reconstructed just using the approximation coefficients (CAn) and the original simulated background.



(a) (b)

Figure 2. (a) The simulated pure Raman spectrum which is composed of six Gaussian peaks and random noise. (b) The simulated original spectrum with curved background. (c) The background corrected spectrum reconstructed by just using the detail coefficients in the wavelet inverse transform. (d) The negative values calibrated spectrum after background correction by just using the detail coefficients in the wavelet inverse transform. (e) The background corrected spectrum using the suppression coefficient. (f) The negative value corrected spectrum after background correction using the suppression coefficient.

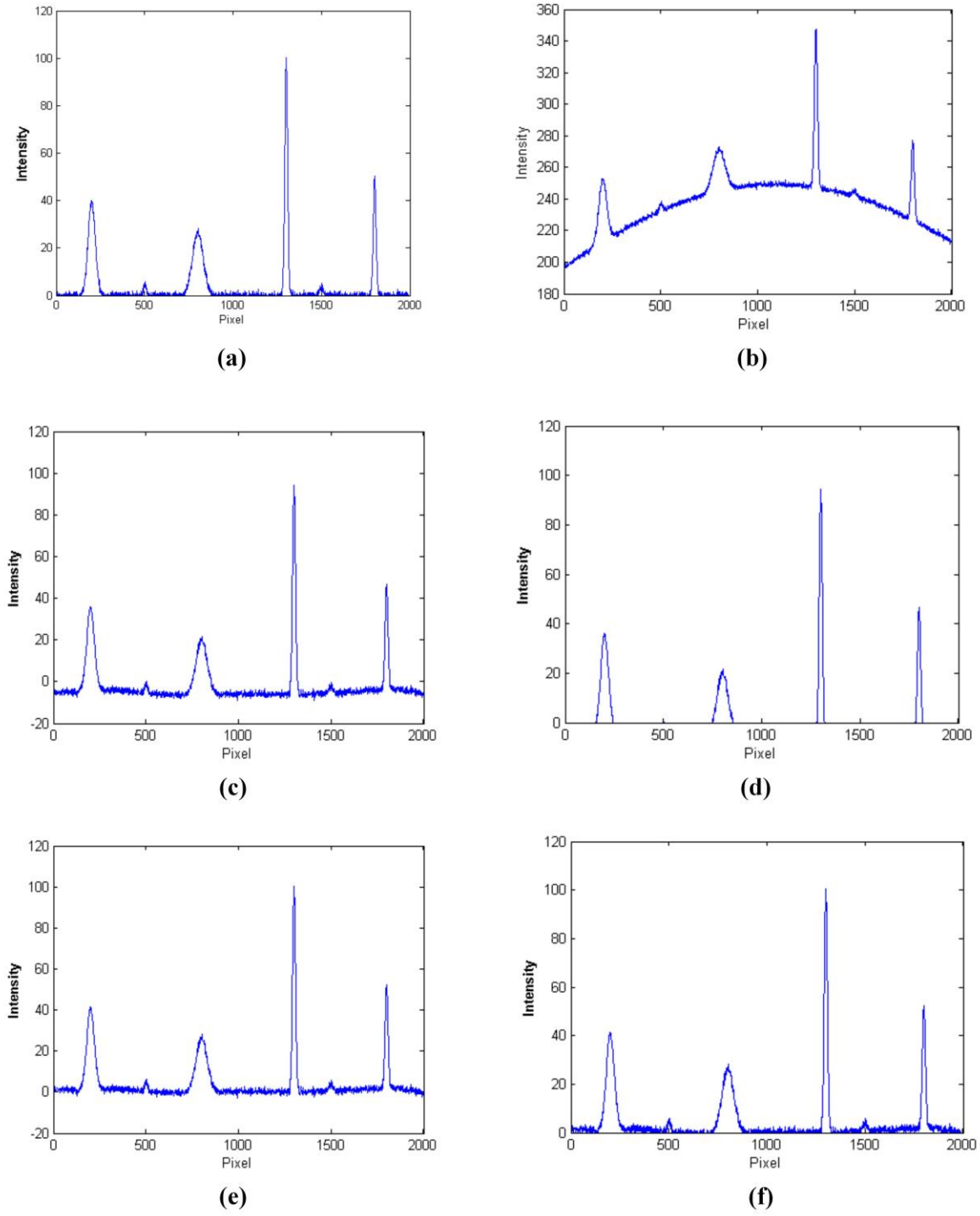


Figure 3. The relationship between the RMS of the reconstructed errors and the number of peaks when using the two kinds of background correction algorithm.

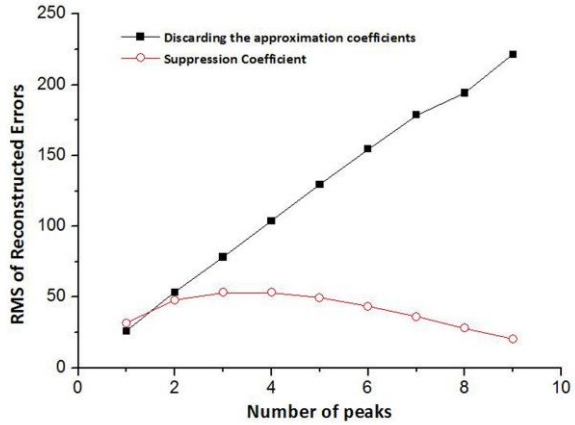


Figure 4. Background correction of spectra with different kinds of background, (a–c) Flat, linear, and sigmoidal, SNR=150 dB. Different SNR using the improved algorithm, (d–f) Flat, linear, and sigmoidal, SNR=30 dB.

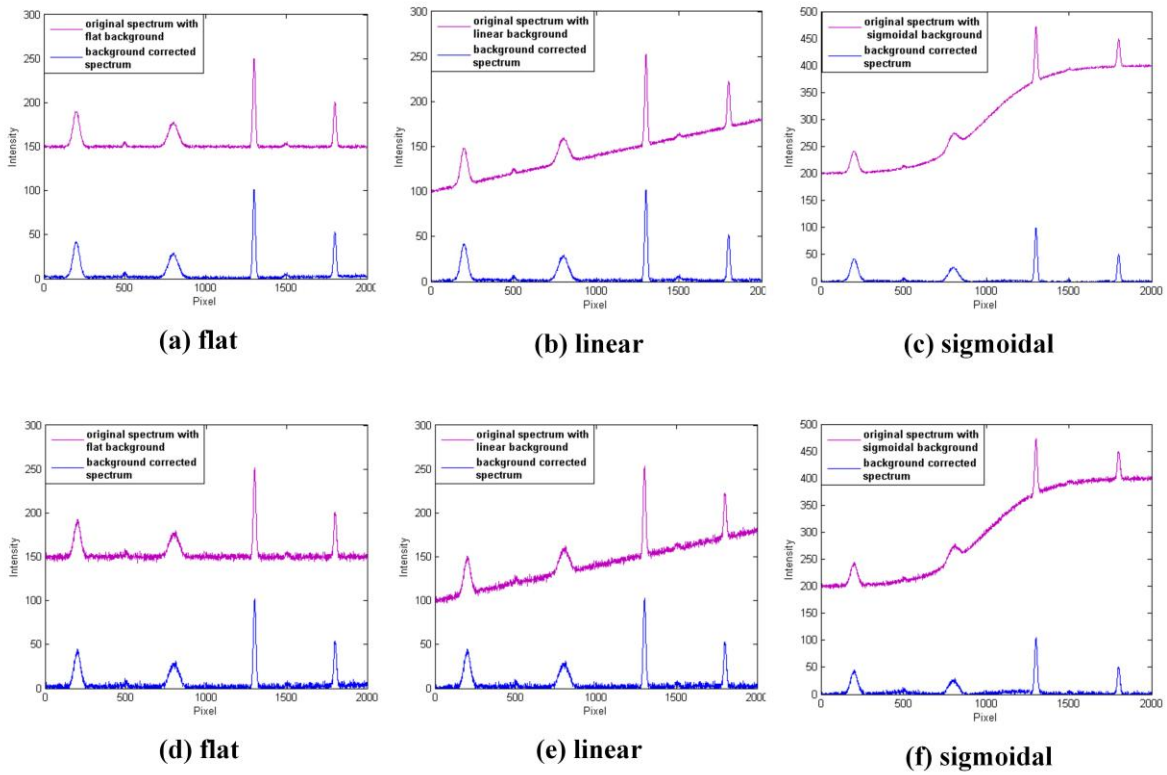


Figure 5. The performance comparison between the improved WT algorithm and the iterative polynomial algorithm. (a) The RMS values of the reconstructed errors when the two algorithms are implemented on simulated spectra with different SNR. (b) The running time of

the two algorithms when the SNR of the simulated Raman spectrum is 190 dB.

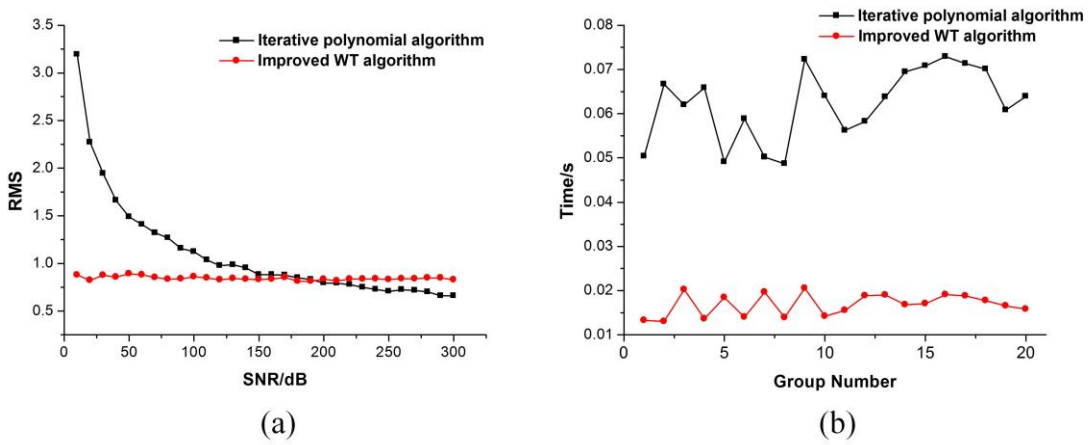


Figure 6. The simulated experiments on the choice of proper decomposition level and wavelet type for different kinds of background. (a) The relationship between the RMS and the decomposition level when using the coif4 wavelet. (b) The relationship between the RMS and different wavelet when the decomposition level is 9.

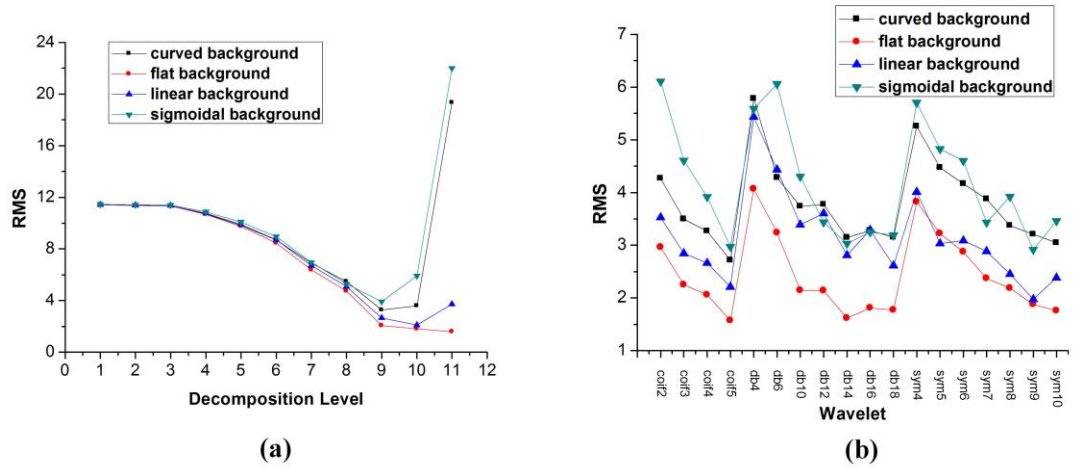
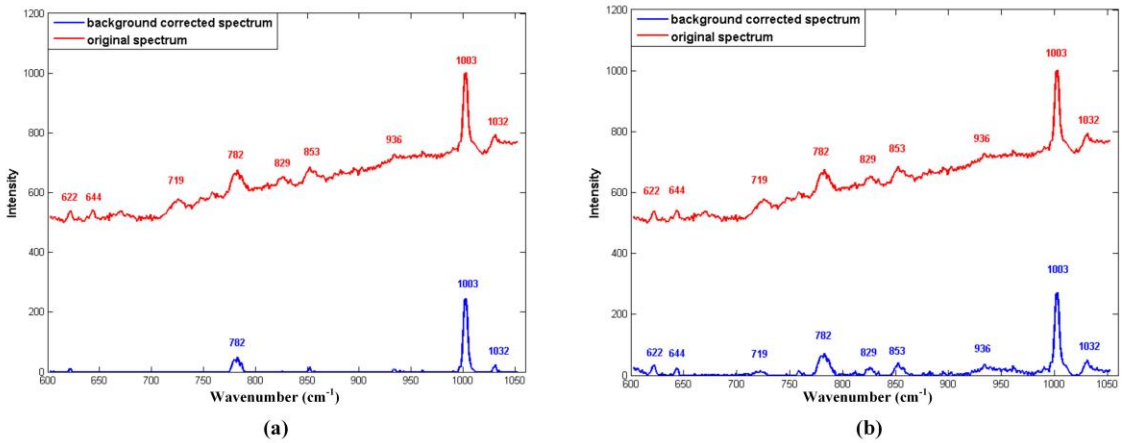
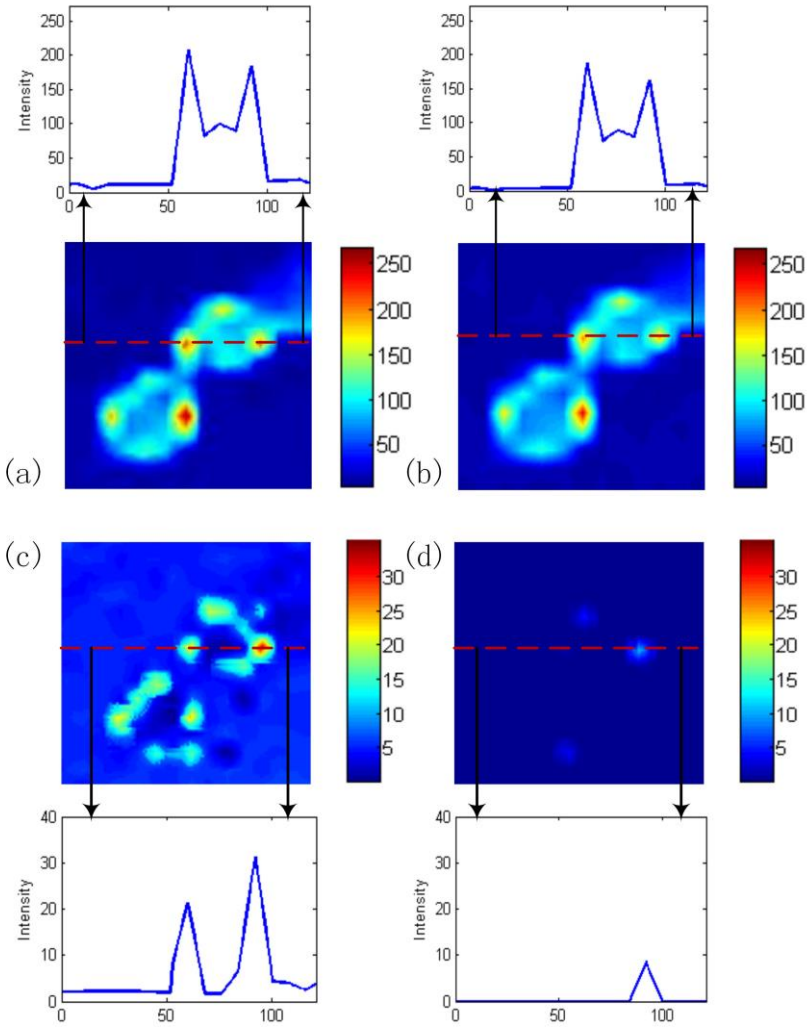


Figure 7. The performance comparison between the traditional algorithm which just discards the approximation coefficients and the improved algorithm which utilize the suppression coefficient in the background correction of experimental Raman spectrum. (a) The original Raman spectrum and the reconstructed spectrum obtained with the traditional algorithm. (b) The original Raman spectrum and the reconstructed spectrum obtained with the improved



algorithm.

Figure 8. The performance comparison of the traditional and improved background correction algorithms for Raman spectra with high SNR (1003 cm^{-1}) and low SNR (644 cm^{-1}). (a) The Raman image at 1003 cm^{-1} constructed by the spectral data processed with the improved algorithm using the suppression coefficient. (b) The Raman image at 1003 cm^{-1} constructed by the spectral data processed with the traditional algorithm which just discards the approximation coefficient. (c) The Raman image at 644 cm^{-1} constructed by the spectral data processed with the improved algorithm. (d) The Raman image at 644 cm^{-1} constructed using the spectral data processed with the traditional algorithm.


Intensity
(a)
(b)
(c)
(d)