Available online at www.sciencedirect.com
ScienceDirect
Chemometrics and Intelligent Laboratory Systems 85 (2007) 94–101
Chemometrics and intelligent laboratory systems
www.elsevier.com/locate/chemolab

A background elimination method based on wavelet transform for Raman spectra

Yaogai Hu [a,b], Tao Jiang [b,c], Aiguo Shen [b], Wei Li [b], Xianpei Wang [a], Jiming Hu [b,*]

[a] College of Electronic Information, Wuhan University, Wuhan, 430079, China
[b] College of Chemistry and Molecular Sciences, Wuhan University, Wuhan, 430072, China
[c] School of Stomatology, Wuhan University, Wuhan, 430079, China

Received 15 March 2005; received in revised form 21 November 2005; accepted 9 May 2006
Available online 3 July 2006

Abstract

A new hybrid algorithm is proposed to eliminate the varying background of spectral signals. The method is based on the use of multi-resolution, which is one of the main advantages provided by wavelet transform. Compared with the analyte signal, the background has a low frequency. The new method firstly split the signals into different frequency components, and then removes the varying low-frequency background. The method is successfully applied to simulated spectral data set and experimental Raman spectral data. The results showed that the wavelet transform technique could handle all kinds of background and low signal-to-background ratio spectra, and required no prior knowledge about the sample composition, no selection of suitable background correction points, and no mathematical assumption of the background distribution. The proposed procedure was illustrated, by processing real spectra, to be an effective and practical tool for background elimination in Raman spectra. In addition, the proposed strategy can be applied to other spectral signals as well.

Keywords: Wavelet transform; Background elimination; Raman spectroscopy

1. Introduction

Infrared spectroscopy and Raman spectroscopy are being increasingly used to measure, both directly and indirectly, a large number of chemical and physical properties of materials. Spectral interferences, including varying backgrounds/baselines and noise, lead to problems with instrument calibration and quantitation of the spectral information [1–3]. As in Ref. [1], the results indicated that one of the most significant sources of spectral variation was a variable curved background. The interference of background and noise will lead to the worse precision. Recently, background elimination for spectral data has been paid much attention, and some methods have already been presented [2–5]. For the experimental Raman spectral data, the baseline is usually non-constant and varying from sample to sample, it is difficult to remove this kind of baseline and only a few methods have been presented [2,3,6,7]. The diverse sources of background and noise make it as a hard task to eliminate them.

Wavelet transform is a new signal processing method [8,9] and has recently been applied to signal processing in analytical chemistry, such as in noise filtering, data compression and overlapping peaks separation. By means of the wavelet transform, an original signal can be decomposed into localized contributions characterized by a scale parameter. Each contribution represents a portion of the signal with a different frequency. Furthermore, the wavelet transform is a linear operation, which is important for quantitative analysis. In recent years more works concerning the background removal by using the wavelet transform were published [10–19].

In the present paper, wavelet transform was used to eliminate backgrounds in spectra. For a Raman measurement, the background has a lower frequency than the analyte signal. Hence, by the wavelet decomposition of the measured data set and by removal of the background contribution in the wavelet transformed vector, the pure analyte signal can be created through wavelet reconstruction. The new hybrid algorithm consists of three parts: noise removal (denoising), low-frequency (approximation) removal based on

* Corresponding author. Tel.: +86 27 876824398701; fax: +86 27 87647617.
E-mail address: jmhu@whu.edu.cn (J. Hu).

0169-7439/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.chemolab.2006.05.004

wavelet decomposition algorithm, and peaks correction based on first derivative methods.

Two simulated spectral data were used to validate the performance of the proposed method for elimination of both varying background and noise. The method is also successfully applied to an experimental Raman spectral data. The results showed that the wavelet transform technique could handle all kinds of background and low signal-to-background ratio spectra, and required no prior knowledge about the sample composition, no selection of suitable background correction points, and no mathematical assumption of the background distribution. The proposed procedure was illustrated, by processing real spectra, to be an effective and practical tool for background elimination in Raman spectra. In addition, the proposed strategy can be applied to other spectral signals as well.

2. Theory

2.1. Wavelet decomposition

Wavelet transform decomposes a signal into localized contributions (details and approximations) labeled by a scale and a position parameter. And each of the contributions at different scale (or resolution level) represents the information of different frequency contained in the original signal. Similarly to the Fourier transform, the discrete wavelet transform (DWT) decomposes the time record $x(t)$ (t=1,2, …, N) into dyadic wavelet functions $\psi_{j,k}(t)$ and scaling functions $\phi_{j,k}(t)$. The basis for this decomposition is formed from mother wavelet $\psi(t)$ and father wavelet $\phi(t)$, by translating in time and dilating in scale [20].

$$\begin{aligned}\psi_{j,k}(t) &= 2^{-j/2}\psi(2^{-j}t-k),\\ \phi_{j,k}(t) &= 2^{-j/2}\phi(2^{-j}t-k), \quad j,k\in Z\end{aligned} \tag{1}$$

where k=1, 2, …, $N/2$, N is the number of data record. j=1, 2, …, J, J is often a small natural number, Z is the set of integers.

Unlike conventional techniques, wavelet decomposition produces a family of hierarchically organized decompositions. At each level j, the j-level approximation $A_j(t)$, and a deviation signal called the j-level detail $D_j(t)$ can be calculated according to the following equations.

The detail $D_j(t)$ is defined as

$$D_j(t) = \sum_{k\in Z} C(j,k)\psi_{j,k}(t), \quad j,k\in Z \tag{2}$$

where, $C(j, k)$ is the wavelet coefficients, and

$$C(j,k) = \int_{-\infty}^{+\infty} x(t)\psi_{j,k}(t)\mathrm{d}t \tag{3}$$

The signal $x(t)$ is the sum of all the details:

$$x(t) = \sum_{j\in Z} D_j(t) \tag{4}$$

Then, take a reference level called J; there are two sorts of details. Those associated with indices $j \le J$ correspond to the fine details, the others, which correspond to $j>J$, are the coarser details, we group these latter details into

$$A_J(t) = \sum_{j>J} D_j(t), \tag{5}$$

which defines what is called an approximation of the signal $x(t)$. Apparently, with the increase of the level J, the resolution defined as 2^{-J} decreases, and $A_J(t)$ will only contain the "lower frequency" components of $x(t)$.

The detail $D_j(t)$ and the approximation $A_J(t)$ are connected by Eq. (6)

$$x(t) = A_J(t) + \sum_{j\le J} D_j(t). \tag{6}$$

From the preceding formulas, it is obvious that the approximations are related to one another by

$$A_{J-1}(t) = A_J(t) + D_j(t). \tag{7}$$

2.2. Denoising

All signals obtained as instrumental response of analytical apparatus are affected by noise. The noise degrades the accuracy and precision of an analysis, and it also reduces the detection limit of the instrumental technique. Signal denoising is therefore highly desirable in analytical response optimization. There are different possible approaches to signal denoising and one of the most recent methods is based on the wavelet transform.

Donoho and Johnstone [21] and Donoho [22] have developed a method known as the wavelet shrinkage to estimate an unknown smoothed signal from data with noise. The methodology can be applied to the DWT, as well as to the wavelet packet transform (WPT). In the present article, we only focus on the DWT.

The general denoising procedure involves three steps. The basic version of the procedure follows the steps described below.

1) Decompose
   Choose a wavelet and level L. Compute the wavelet decomposition of the signal s at level L.
2) Threshold detail coefficients
   For each level from 1 to L, select a threshold and apply soft or hard thresholding to the detail coefficients.
3) Reconstruct
   Compute wavelet reconstruction using the original approximation coefficients of level L and the modified detail coefficients of levels from 1 to L.

Two different approaches are usually applied to denoise: hard thresholding or soft thresholding. The hard thresholding method consists in setting all the wavelet coefficients below a given threshold value equal to zero, while in soft thresholding the wavelet coefficients are reduced by a quantity equal to the threshold value [22]. The denoising procedure requires the estimation of the noise level. There are many possible approaches to do so, and a systematic investigation about their performance can by

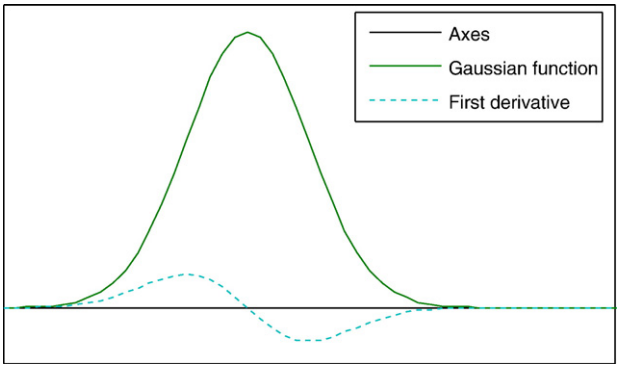

Fig. 1. Gaussian function and its first derivative curve.

found in Ref. [23]. The most popular one, known as 'universal thresholding' (ThU), is based on statistical properties of white Gaussian noise, and defined as:

$$\mathrm{th} = \sigma\sqrt{2\log N} \tag{8}$$

where N=*length(s)* is the signal length and σ is the standard deviation of the noise. The threshold level is calculated from the standard deviation of the detail coefficients of the first level of decomposition.

2.3. Baseline removal based on wavelet decomposition

We know that very broad spectral features were not to be expected to originate from the molecular structure and could therefore safety be ascribed to unwanted baseline and background fluorescence effects. Compared with the analyte signal in a Raman spectrum, the background variation is moderate, i.e. it has a lower frequency. By means of the wavelet decomposition, the measured data are changed into a discrete wavelet transform vector. After discarding the elements attributed to the background in the vector, the pure analyte signal is obtained. Therefore, the wavelet transform can be used to separate analyte signal from background in Raman spectra and fulfill the background elimination. This approach [14,15,19] requires no prior assumption of the background distribution and selection of suitable background correction points according to the appearance of the background, and can handle all kinds of background including a curved background such as line wing background due to the interferents in the sample. But it has some shortcomings and deficiency. For example, there are some meaningless negative values in the reconstructed signal. Chen [16] substituted all negative values with zeroes in the reconstructed signal. The result of this simple treatment is still unsatisfactory to the peak shape, peak height and peak width of spectral signal, except that the peak position can be extracted. One of the focal points of this work is to do the further peaks correction to reconstructed spectrum.

2.4. Peaks correction based on the first derivative method

Generally, the individual peak can be described by a Gaussian distribution function, which contains three parameters indicating the peak height, width and position, and is usually expressed as follows:

$$g(\lambda) = \frac{k}{\sqrt{2\pi}\sigma} e^{-\frac{(\lambda-\mu)^2}{2\sigma^2}} = Ae^{-\frac{(\lambda-\mu)^2}{2\sigma^2}} \tag{9}$$

where μ and σ are the peak position and the standard deviation (width), and A is the peak height.

One important area of signal analysis is the finding of zero crossings of higher derivatives. This would be a trivial matter if the signals were noise-free [19]. According to Eq. (9), we can deduce that

$$\frac{\mathrm{d}(g(\lambda)+C)}{\mathrm{d}\lambda} = \frac{\mathrm{d}g(\lambda)}{\mathrm{d}\lambda} = -\frac{\lambda-\mu}{\sigma^2} Ae^{-\frac{(\lambda-\mu)^2}{2\sigma^2}} \tag{10}$$

where C is an arbitrary constant. A Gaussian function and its first derivative curve are shown in Fig. 1.

For the practical Raman spectra, the intensity of individual peak is often discrete, finite and nonnegative sampling data. Hence, the first derivative of one peak will be discrete and finite data. From Fig. 1, we can know that the first derivative of Gaussian is similar to the sinusoidal curve. The second zero crossing corresponds to the peak position ($\lambda=\mu$); the left maximum point and the right minimum point correspond to two inflexions of Gaussian peak ($\lambda=\mu\pm\sigma$); the first zero crossing and the third zero crossing correspond to the left and right boundary of Gaussian peak. The range of Gaussian peak can be determined through these three zero crossings. The derivative of arbitrary constant equals to zero, and therefore the first derivative do not change for one signal with a flat background added. This is the mathematics foundation of determining the range of the peak through its first derivative curve.

3. Algorithms

The proposed method includes following four main aspects: denoising, multiple-level decomposition, reconstruction and peak correction.

1) Denoising based on DWT
   The names of the Daubechies family wavelets are written as db$N$, where N is the order, and db is the "surname" of the wavelet. In this work, denoising using soft heuristic SURE thresholding and scaled noise option, on detail coefficients obtained from the decomposition of the original signal s, at level 3 by db4 wavelet.
2) Multiple-level decomposition
   Perform decomposition at level 7 of s using db6 wavelet.
3) Signal reconstruction
   The approximation coefficients were set equal to zeroes. Compute wavelet reconstruction using the modified to zero approximation coefficients of level L and the original detail coefficients of levels from 1 to L. The reconstructed signal s* was obtained.
4) Peak correction

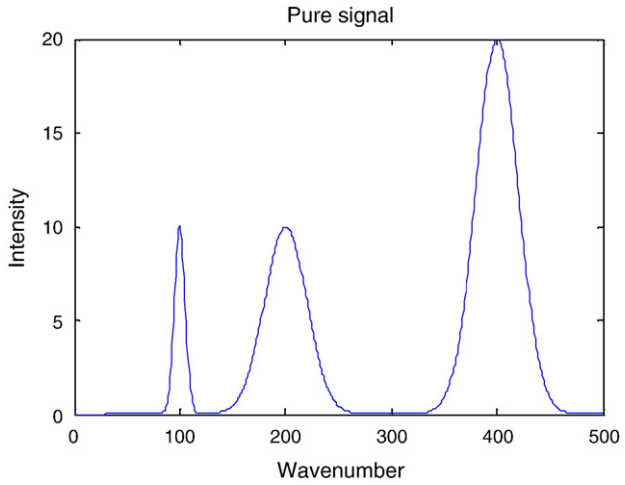

Fig. 2. The simulated pure signal which is composed of three Gaussian peaks.

In practical computation, the signal to be analyzed is often discrete sampling data, which can be denoted as $s(i)$, where $i=1, 2, \ldots, n$, n is the number of s. In this work, the zero crossing of s was defined as: if $s(i)=0$ or $s(i-1)\times s(i)<0$, then i is one of zero crossings of s. It should be stressed that there may exist some false zero crossings which must be discarded due to the noise.

(1) Firstly, obtain the first derivative Ds of the reconstructed signal s^*;

(2) Obtain the zero crossings of Ds: $O(i)=1, 2, \ldots, M$, M is the number of zero crossings.

For all $O(i)$ except for which being equal to 1 or n, judge whether $O(i)$ is the false zero crossing, if so, modify $Ds(O(i))$:

If $Ds(O(i)-1)\geq 0$ and $Ds(O(i)+1)\geq 0$, then $Ds'(O(i))=\min(Ds(O(i)-1), Ds(O(i)+1))$; if $Ds(O(i)-1)\leq 0$ and $Ds(O(i)+1)\leq 0$, then $Ds'(O(i))=\max(Ds(O(i)-1), Ds(O(i)+1))$; otherwise $Ds'(O(i))=Ds(O(i))$. $Ds'(j)=Ds(j)$, $\forall j\in\{1, 2, \ldots, n\}$ and $j\neq O(i)$, where $i=1, 2, \ldots, M$.

Obtain the zero crossings of the corrected first derivative Ds': $O'(i)$, $i=1, 2, \ldots, M'$, M' is the number of zero crossings.

(3) Let $j=0$,

For $i=0$ to M'

If $\min s^*(\lambda)>0$, $\lambda\in[O'(i), O'(i+1))$ and $\max s^*(\lambda)<0$, $\lambda\in[O'(i+1), O'(i+2))$, then $O'(i)$, $O'(i+1)$ and $O'(i+2)$ are three zero crossings for determining one peak. Let $j=j+1$ and $\lambda_1^j=O'(i)$, $\lambda_2^j=O'(i+1)$, $\lambda_3^j=O'(i+2)$.

Next i.

(4) Based on all three zero crossings $(\lambda_1^i,\lambda_2^i,\lambda_3^i$, where $i=1, 2, \ldots, P$, P is the number of peaks.), we can determined the positions and the ranges of all Gaussian peaks, and then corrected them: $s'(\lambda)=s^*(\lambda)-\max(s^*(\lambda_1^i), s^*(\lambda_3^i))$, $\lambda\in(\lambda_1^i,\lambda_3^i)$, and $s'(\lambda)=0, \lambda\notin\cup_{i=1}^{P}(\lambda_1^i, \lambda_3^i)$.

(5) Discarding all the Gaussian peaks which peak height are below a given threshold value TH: $\forall\lambda\in(\lambda_1^i, \lambda_3^i)$, $i=1, 2, \ldots, P$, if $\max s'(\lambda)<\mathrm{TH}$, then $s'(\lambda)=0$.

4. Experimental

Two simulated spectral data sets and an experimental Raman spectrum were used to validate the performance of the proposed method for elimination of varying background and noise. Calculations were carried out with the Matlab 7.0 software package (Math-Works, Natick, MA, USA) equipped with the Wavelet Toolbox 3.0.

4.1. Simulated data

The simulation carried out was intended to imitate real spectral data sets that contain varying background, analytical signals, and random noise. A broader Gaussian peak is treated as curved background, a linear function is treated as sloping background, and a narrower Gaussian peak is treated as the spectra of interest.

The simulated signal $s(\lambda)$ is composed of the pure signal $g(\lambda)$, background $b(\lambda)$ and random noises $Noise(\lambda)$:

$$s(\lambda)=g(\lambda)+b(\lambda)+\mathrm{Noise}(\lambda) \tag{11}$$

As shown in Fig. 2, the simulated pure signal $g(\lambda)$ is composed of three Gaussian peaks:

$$g(\lambda)=10e^{-\frac{(\lambda-100)^2}{2\times 5^2}}+10e^{-\frac{(\lambda-200)^2}{2\times 20^2}}+20e^{-\frac{(\lambda-400)^2}{2\times 20^2}} \tag{12}$$

The sloping and curved backgrounds are appointed as $b(\lambda)=1+0.9*\lambda$ and $b(\lambda)=100e^{-\frac{(\lambda-400)^2}{2\times 200^2}}$, respectively.

$Noise(\lambda)$ is the random real number between 0 and 1, $\lambda=1, 2, \ldots, 500$.

Two simulated data sets (one with the sloping background and another one with the curved background) are shown in Fig. 3.

4.2. Experimental Raman spectrum

4.2.1. Preparation of gastric mucosa tissue samples

Excision specimens from one histologically proven adenocarcinoma of stomach were obtained from an individual from Department of Oncology of Zhongnan Hospital, Wuhan University, after informed consent. A biopsy was taken from each sample, snap-frozen in liquid nitrogen and stored at −80 °C until use. Cryosections (25 μm thickness) were obtained from the biopsy specimens and placed on gold sheet for Raman spectroscopy. During the measurement, the gastric mucosa tissue section was thawed to reach room temperature in air. Histopathological analysis has been performed on haematoxylin and eosin stained sections by a single GI-registry pathologist.

4.2.2. Spectroscopic instrumentation

A Renishaw Raman microspectrometer (Renishaw Raman system RM1000) has been optimised for maximum throughput, detection sensitivity and fluorescence suppression. The argon ion laser provided a 20 mW incident light at 514.5 nm. After attenuation through prisms and filters, the power of laser exposed on samples was only about 4 mW, which almost impossible lead to degradation of the samples. Spectra were measured from tissue and cell specimens with a ×20 short-working-distance objective

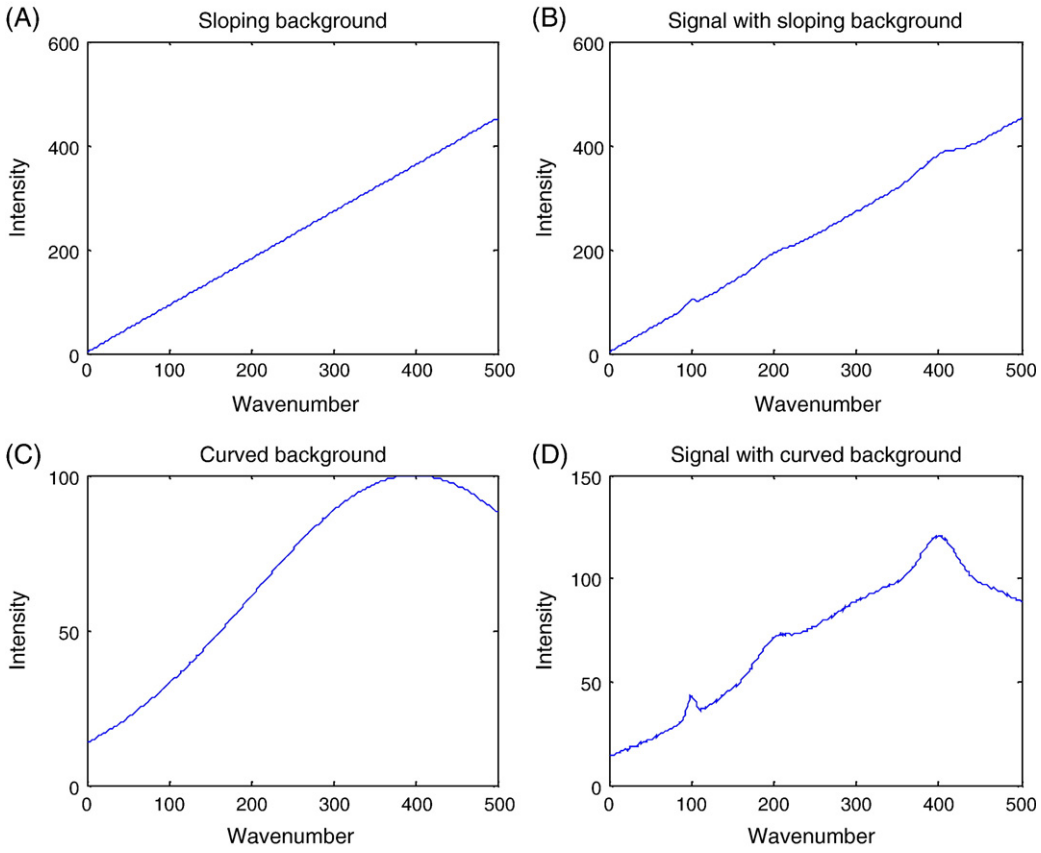

Fig. 3. Two simulated signals with noise and different backgrounds. (A) The simulated sloping background. (B) Original signal with noise and the sloping background. (C) The simulated curved background. (D) Original signal with noise and the curved background.

(NA 1.20), and the signals were integrated for 30–120 s and measured over a spectral range of 600–1800 cm^{-1} with respect to the excitation frequency. The system includes a stigmatic spectrometer with two motorized gratings, of which the 1800 grooves/mm grating was used to provide a spectral resolution of the Raman scatter of around 4 cm^{-1}. Raman scattering was detected by using an air cooled 578×385 pixels CCD camera. Peak frequencies and rapid checking of instrumental performance were calibrated with the silicon phonon line at 520 cm^{-1}. Spectrum data were visualized on a computer and exported TXT data by the Grams/32 spectroscopic software.

5. Results and discussion

The treatments of sloping and curved backgrounds have been done by using the wavelet transform technique, especially for low signal-to-background ratio spectra. As a special case of sloping and curved backgrounds, the flat background was not concerned.

5.1. Determination of the wavelet function

There exist several different wavelet families and orders of wavelets. An important aspect of the choice of a wavelet is to select a wavelet with more vanishing moments than the polynomial order of the interesting signal being analyzed. In this paper, only the Daubechies wavelets were tested for the ability of noise removal and background elimination. It should be stressed that there exist a very large number of possible wavelet bases. The Daubechies wavelets are smooth, orthogonal, and compactly supported [9]. For db$N$ wavelet, the support length of ψ and ϕ is $2N-1$, and the number of vanishing moments of ψ is N. The bigger the order N is, the better the regularity is.

The popular db4 wavelet, therefore, is a good choice for low-frequency signals such as Raman spectral signal. For our calculations, we used the db4 mother wavelet at a three-level decomposition. The db4 wavelet is often used for analyzing and denoising analytical signals and three- or four-level decomposition is recommended for denoising [24].

Fig. 4 shows the reconstructed signals of the simulated signal with sloping background obtained with different Daubechies wavelet (db4–db8) under the same decomposition number 7. For db4 and db8, there are two false peaks in the range of 50–80 and 250–350. For db5, there is one false peak in the range of 250–350. For db7, there is one false peak in the range of 50–80. It was found that the Daubechies wavelets db6 performed well. After processing of a spectrum using a transform with db6, the peak in the processed spectrum became symmetrical with a flat baseline and the intensity at the top of the line approached the true value.

5.2. Effect of decomposition number

Once the wavelet function for removal of background and noise are decided, the optimal parameters of scales should be determined to obtain the best quantitative results.

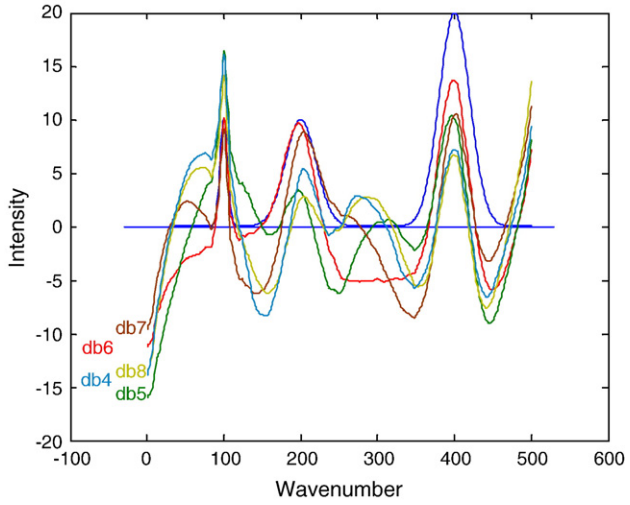

Fig. 4. The reconstructed signals of the simulated signal with sloping background obtained with different Daubechies wavelet (db4–db8) under the same decomposition number 7. The blue curve is the simulated pure signal. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The reconstructed signals of the simulated signal with sloping background corresponding to different decomposition number 5–9 are shown in Fig. 5. With increase of decomposition number, the analyte contribution was gradually excluded from the spectrum. If the number was too small, the fitted background curve would involve much information of the analyte signal. Consequently, over subtraction of the background would be observed and the concentration of the analyte would be underestimated. From Fig. 5 we can see that the reconstructed signal is an oscillating curve which includes six false peaks when decomposition number is 5. There is one false peak in the range of 250–350 when decomposition number is 6. On the contrary, too much decomposition would result in an excessively low estimation of the background under the analyte line

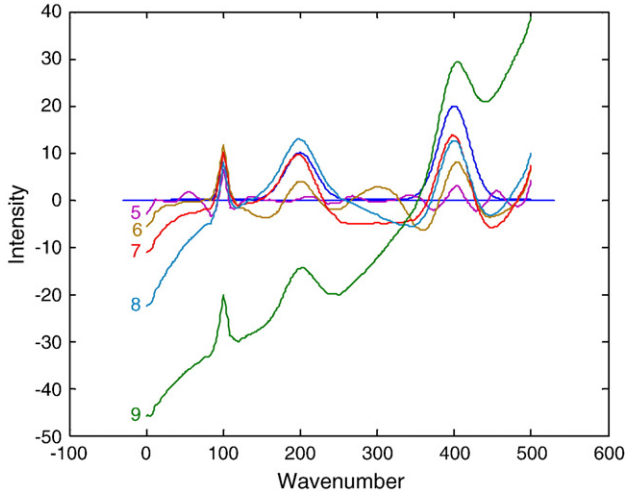

Fig. 5. The reconstructed signals of the simulated signal with sloping background corresponding to different decomposition number 5–9. The blue curve is the simulated pure signal. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

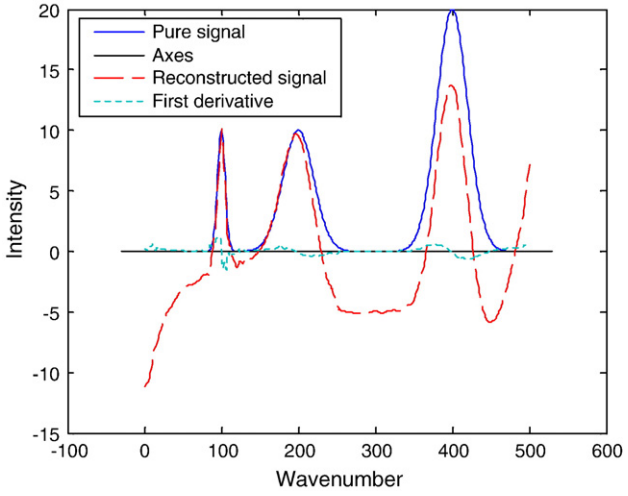

Fig. 6. The reconstructed signal and its first derivative curve for the simulated signal with sloping background.

and a great error in the result. For example, there are obvious background trends in the reconstructed signal when the decomposition number is 9.

The proper decomposition number could be determined by visual inspection of the appearance of the processed spectrum. In these experiments, the appropriate decomposition number was found to be 7 and 8.

5.3. Elimination of sloping and curved backgrounds

The reconstructed signal of the simulated signal with sloping background and its first derivative curve are shown in Fig. 6. From this figure, it is obviously that all three Gaussian peak positions and the range of peaks have been effectively extracted by the first derivative curve.

After treatment of the reconstructed signal using above-mentioned peak correction method in Section 3, the background eliminated signal of the simulated signal with sloping background are obtained and shown in Fig. 7. Similarly, the background

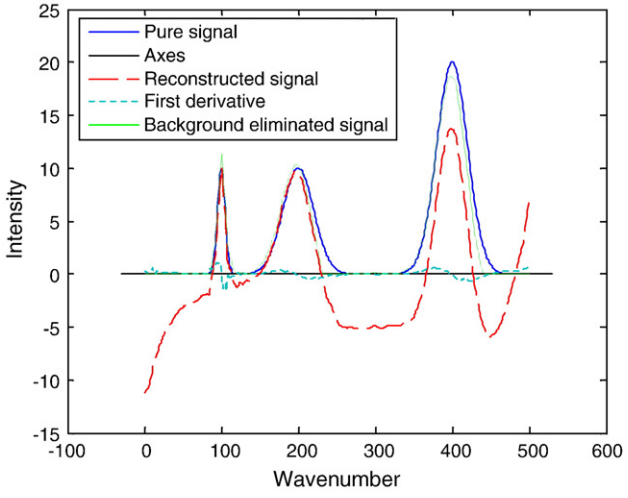

Fig. 7. The background eliminated signal of the simulated signal with sloping background.

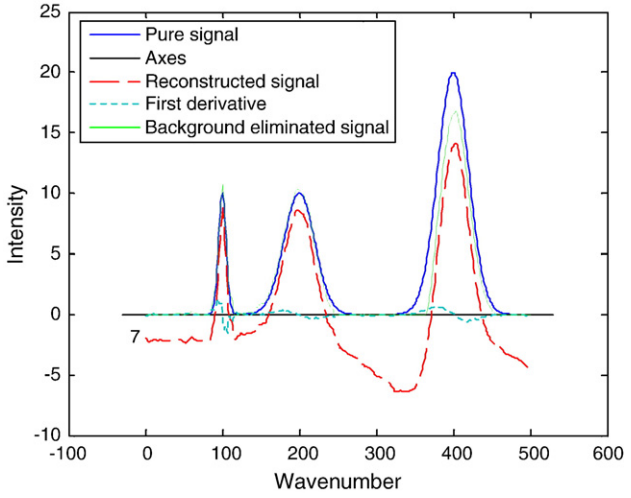

Fig. 8. The background eliminated signal of the simulated signal with curved background.

eliminated signal of the simulated signal with curved background are obtained and shown in Fig. 8. Table 1 compares the background eliminated results with the expected values. The positions and heights of three Gaussian peaks approached the true value. It is clear that the proposed approach had good performances in subtracting both simple sloping backgrounds and complex curved backgrounds.

5.4. Experimental Raman spectrum

The above-mentioned approach was applied to experimental Raman spectrum analysis. Fig. 9(A) was the original Raman spectrum of malignant gastric mucosa tissues in the 600–1800 cm^{-1} region. It can be seen that the spectrum is affected by noise and the backgrounds. The wavelet denoised Raman spectrum was shown in Fig. 9(B). From visual inspection of the result it is obvious that the wavelet denoising method has performed well. The reconstructed signal and its first derivative curve were shown in Fig. 9(C). The last background eliminated spectrum obtained by the proposed method was shown in Fig. 9(D).

In the next work, based on applying the above-mentioned background elimination algorithms to the Raman spectra of the normal and malignant gastric mucosa tissue samples, we will focus on extracting the effectual features of the pattern and classifying them into two categories: normal and malignant

Table 1
Compare the background eliminated results and the expected value

Peak position, μ			Peak height, A		
Expected value	Processed results 1	Processed results 2	Expected value	Processed results 1	Processed results 2
100	100	100	10	9.2595	9.235
200	199	198	10	10.884	10.489
400	398	401	20	18.485	17.418

Processed value 1 is the background eliminated results for the simulated signal with the sloping background.
Processed value 2 is the background eliminated results for the simulated signal with the curved background.

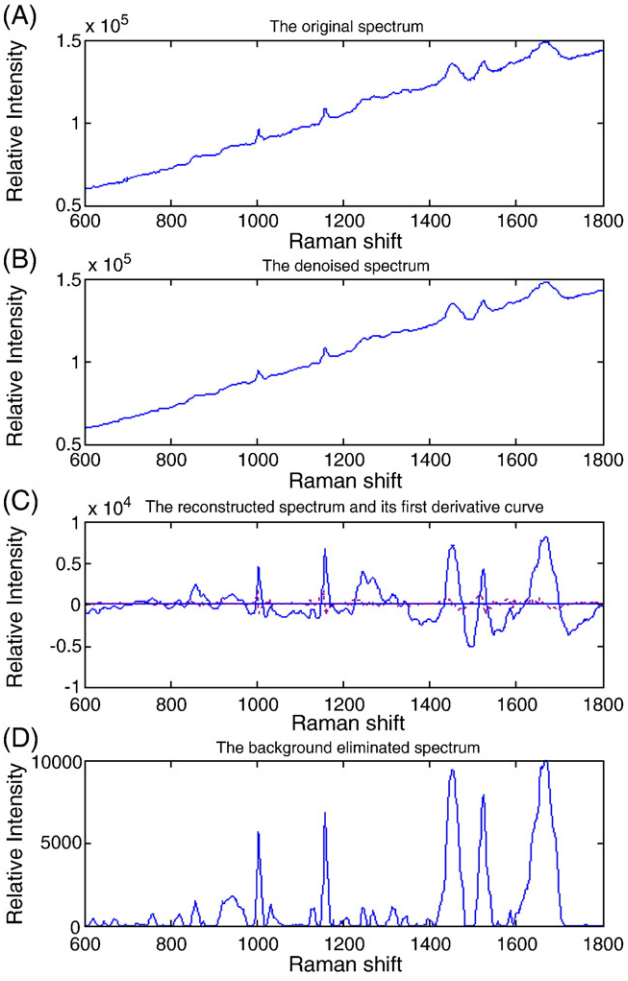

Fig. 9. (A) The original Raman spectrum of malignant gastric mucosa tissues in the range of 600–1800 cm^{-1}. (B) The wavelet denoised Raman spectrum. (C) The reconstructed Raman spectrum and its first derivative curve. (D) The last background eliminated Raman spectrum.

mucosa, and finally develop a new diagnostic method which would assist the early detection of stomach cancer.

6. Conclusions

The objective of this study was to apply the wavelet transform technique to removal of noise and backgrounds in spectra. For the Raman spectral signals, the background has a lower frequency than the analyte signal. Hence, by the wavelet decomposition of the measured spectral data and by removal of the background contribution in the wavelet transformed vector, the pure analyte signal can be created through wavelet reconstruction. The proposed method was successfully applied to simulated data sets and experimental Raman data for removal of non-constant background and noise. The results showed that the wavelet transform technique could handle all kinds of background and low signal-to-background ratio spectra, and required no prior knowledge about the sample composition, no selection of suitable background correction points, and no mathematical assumption of the background distribution. The proposed procedure was illustrated, by processing real spectra, to be an effective

and practical tool for background elimination in Raman spectra. In addition, the proposed strategy can be applied to other spectral signals as well.

Acknowledgements

The work reported in this paper was financed by National Nature and Science Foundation of China Grant no. 20375029 and no. 30400507. The authors wish to thank the anonymous reviewers for their comments and suggestions.

References

[1] S.C. Rutan, O.E. de Noord, R.R. Andréa, Anal. Chem. 70 (1998) 3198.
[2] C.R. Mittermayr, H.W. Tan, S.D. Brown, Appl. Spectrosc. 55 (2001) 827.
[3] H.W. Tan, S.D. Brown, J. Chemom. 16 (2002) 228.
[4] T.V. Karstang, O.M. Kvalheim, Chemom. Intell. Lab. Syst., Lab. Inf. Manag. 12 (1991) 147.
[5] P.J. Gemperline, J.H. Cho, B. Archer, J. Chemom. 13 (1999) 153.
[6] R.P. Raradkar, R.P. Williams, Appl. Spectrosc. 51 (1997) 92.
[7] J.G. Sun, Appl. Spectrosc. 11 (1997) 525.
[8] C. Chui, An Introduction to Wavelets, Academic Press, New York, 1992.
[9] Y. Meyer, Wavelets: Algorithms and Applications, SIAM, Philadelphia, PA, 1992.
[10] P. Ramos, I. Ruisanchez, J. Raman Spectrosc. 36 (2005) 848.
[11] D. Chen, X.G. Shao, B. Hu, Q.D. Su, Anal. Chim. Acta 511 (2004) 37.
[12] C.X. Ma, X.G. Shao, J. Chem. Inf. Comput. Sci. 44 (2004) 907.
[13] B.F. Liu, Y. Sera, N. Matsubara, K. Otsuka, S. Terabe, Electrophoresis 24 (2003) 3260.
[14] X. Qin, L.S. Shen, Comput. Appl. Chem. 20 (2003) 23.
[15] X.G. Ma, Z.X. Zhang, Anal. Chim. Acta 485 (2003) 233.
[16] G. Chen, Z.Y. Wen, G.R. Yang, S.L. Huang, Opt. Prec. Eng. 10 (2002) 552.
[17] C. Perrin, B. Walczak, D.L. Massart, Anal. Chem. 73 (2001) 4903.
[18] T.T. Cai, D.M. Zhang, D. Ben-Amotz, Appl. Spectrosc. 55 (2001) 1124.
[19] B.K. Alsberg, A.M. Woodward, D.B. Kell, Chemom. Intell. Lab. Syst., Lab. Inf. Manag. 37 (1997) 215.
[20] M.B. Priestley, J. Time Ser. Anal. 17 (1999) 85.
[21] D.L. Donoho, I.M. Johnstone, Biometrika 81 (1994) 425.
[22] D.L. Donoho, IEEE Trans. Inf. Theory 41 (1995) 613.
[23] B.K. Alsberg, A.M. Woodward, M.K. Winsor, J. Rowland, D.B. Kell, Analyst 122 (1997) 645.
[24] L. Pasti, B. Walczak, D.L. Massart, P. Reschiglian, Chemom. Intell. Lab. Syst., Lab. Inf. Manag. 48 (1999) 21.