Chemometrics and Intelligent Laboratory Systems 98 (2009) 24–30

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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

A simple background elimination method for Raman spectra

Seong-Joon Baek [a,*], Aaron Park [a], Jinyoung Kim [a], Aigou Shen [b], Jiming Hu [b]

[a] The Department of Electronics Engineering, Chonnam National University, Gwangju, Republic of Korea
[b] College of Chemistry and Molecular Sciences, Wuhan University, Wuhan, China

ARTICLE INFO

Article history:
Received 10 April 2008
Received in revised form 18 March 2009
Accepted 26 April 2009
Available online 6 May 2009

Keywords:
Background elimination
Raman spectroscopy
Interpolation

ABSTRACT

In this paper, we consider a new background elimination method for Raman spectra. The proposed method is based on peak detection, smoothing, and interpolation. Since the background is usually slowly varying with respect to wavelength, we could estimate the background by eliminating significant peaks. For this purpose, we seek the peaks by inspecting the smoothed derivative of a given spectrum. After clipping out the corresponding peak regions, we estimate the background by applying a modified linear interpolation. Then the background is eliminated from the measured Raman spectrum by simple subtraction. The experimental results showed that the proposed method gave satisfactory results for real Raman spectra as well as synthetic data. As the proposed method requires no prior knowledge of spectrum, we expect that the method could be applied to other spectral data as well.

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 and noise, lead to problems with instrument calibration and quantization of the spectral information [1,2]. According to previous works, one of the most significant sources of spectral variation is a variable curved background and the interference of background and noise leads to the worse precision.

Background elimination for spectral data has been paid much attention, therefore some methods have already been presented. But for the experimental Raman spectral data, the diverse sources of background and noise make it hard to eliminate them. In addition to it, as the background is usually varying from sample to sample, it is difficult to remove and only a few methods have been presented [1–5].

Recently, wavelet transform was used to eliminate backgrounds in Raman spectra [4]. The method relies on the filtering capabilities of the wavelet transform. However the method gave artificially looking results due to its peak correction step especially for experimental Raman spectra. Also the method is rather complex to implement. To overcome such issues, we propose a simple but effective background elimination method as an alternative, which relies on peak detection, smoothing, and interpolation.

For the experimental Raman data, the background consists of lower frequency components than the analyte signal. Hence, we could estimate background signal by eliminating peaks with high-frequency components which is supposed to originate from analytes. Once background signal is estimated, it can be eliminated from the measured Raman spectrum by simple subtraction. In this procedure, it is critical to identify which parts of a spectrum contain significant peaks. For the purpose, we propose a peak detection method based on the differentials of a measured spectrum. After clipping out the detected peak regions, we should fill the data in the clipped regions to estimate the background. We propose a modified linear interpolation method since simple linear interpolation gives unsatisfactory results according to our experiments. Accordingly, the proposed algorithm consists of three main parts: noise removal, peak detection, and interpolation, which will be discussed in detail.

Two simulated spectral data were used to validate the performance of the proposed method for elimination of both varying background and noise. Two kinds of experimental Raman spectral data were also used to verify the proposed method. The results showed that the proposed methods could handle various kinds of background signal. As the method requires no prior knowledge about the sample composition, no selection of suitable background correction points, and no mathematical assumption of the background distribution, we expect that the method could be applied to other spectral data as well.

2. Theory

2.1. Noise removal

All signals obtained as instrumental response of analytical apparatus are affected by high-frequency noise. The noise degrades the accuracy and precision of the analysis, and it also reduces the detection limit of an instrumental technique. So noise removal is highly desirable for analytical response optimization.

* Corresponding author.
E-mail address: tozero@jnu.ac.kr (S.-J. Baek).

0169-7439/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.chemolab.2009.04.007

There are some possible approaches to noise removal. In this work, we adopted a Savitzky–Golay filter. The filter is often used with frequency data or with spectroscopic data. For frequency data, the method is effective at preserving the high-frequency components of the signal. For spectroscopic data, the method is effective at preserving higher moments of the peak such as the line width. By comparison, the moving average filter tends to filter out a significant portion of the signal's high-frequency content, and it can only preserve the lower moments of a peak such as the centroid [6]. However if the height of peak is not quite important, the moving average filter may be used for a practical reason.

The Savitzky–Golay filter can be thought of as a generalized moving average filter. The filter coefficients are derived by unweighted linear least squares fit using a polynomial of a given degree. Once the filter coefficients w are obtained, we can compute the smoothed value according to the following equation.

$$\begin{aligned} y[i] &= \sum_{k=-N}^{N} w_k x[i-k] \\ &= w_{-N}x[i-N] + w_{-N+1}x[i-N+1] + \cdots + w_N x[i+N], \end{aligned}$$

where $y[i]$ is the smoothed value for $x[i]$, N is the number of neighboring data points on either side of $x[i]$. In this case, the span which is the length of a filter is $2N+1$. An efficient implementation of finding Savitzky–Golay filter coefficients is presented in [7] and some precomputed values for the coefficients are found in [8]. If the coefficients are precomputed and stored for some frequently used span, then the Savitzky–Golay filtering reduces to simple additions and multiplications.

2.2. Peak detection and elimination

Baseline, i.e., background in a Raman spectrum shows a very broad spectral feature which is not originated from a certain molecular structure. Compared with the analyte signal in a Raman spectrum, the variation of baseline is moderate, i.e. it has lower frequency components. Therefore, if we identify which parts of a spectrum contain peaks of analyte which represents higher frequency components, we could successfully estimate the background by peak elimination and interpolation.

A typical Gaussian peak and its first derivative are shown in Fig. 1. Inspecting the figure, we can say that the peak maximum can be found with the sign change of the derivative from positive to negative. The boundary of the peak can be determined with adjacent zero positions of the derivative.

If a peak is buried in a background signal, the derivative of a spectrum is the sum of the derivative of analyte and background. Let's see Fig. 2, which shows three Gaussian peaks with and without curved background. Both of derivatives are depicted in Fig. 3 along with the curved background's derivative. Let $g(\lambda)$, $b(\lambda)$ be Gaussian peaks and background. Gaussian peaks buried in the background, $s(\lambda)$, is $g(\lambda)+b(\lambda)$. Then its derivative is expressed as:

$$\frac{\mathrm{d}s(\lambda)}{\mathrm{d}\lambda} = \frac{\mathrm{d}g(\lambda)}{\mathrm{d}\lambda} + \frac{\mathrm{d}b(\lambda)}{\mathrm{d}\lambda}.$$

If we know the derivative of the background, it is easy to get the derivative of Gaussian peaks from the observed signal s.

$$\frac{\mathrm{d}g(\lambda)}{\mathrm{d}\lambda} = \frac{\mathrm{d}s(\lambda)}{\mathrm{d}\lambda} - \frac{\mathrm{d}b(\lambda)}{\mathrm{d}\lambda}.$$

For the discrete data, derivative is approximated with difference:

$$\frac{\mathrm{d}s(\lambda)}{\mathrm{d}\lambda} \approx s[k] - s[k-1],$$

where k is a sample index corresponding to λ. We denote this operation as $\mathrm{d}s = \mathit{diff}(s)$ in the following.

Now the problem is how to estimate the derivative of the background. As we have mentioned, the background shows very broad spectral features and its derivative is also very slowly varying as you see in Fig. 3. Therefore we can say that the elimination of some typical peaks from ds would make ds similar to db.

Here we used a very simple approach to do it. As is known, smoothing with large span preserves the slowly varying components while suppressing the fast varying components. As a smoothing

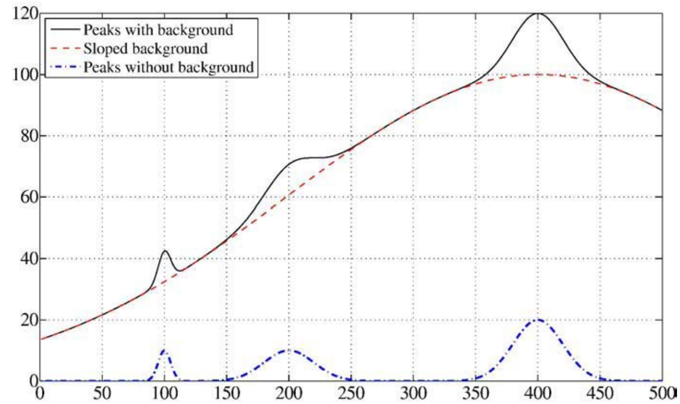

Fig. 2. Three Gaussian peaks with and without curved background.

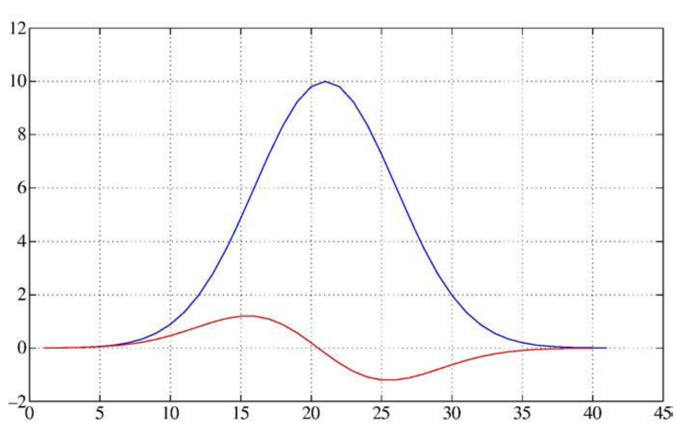

Fig. 1. A Gaussian peak and its derivative.

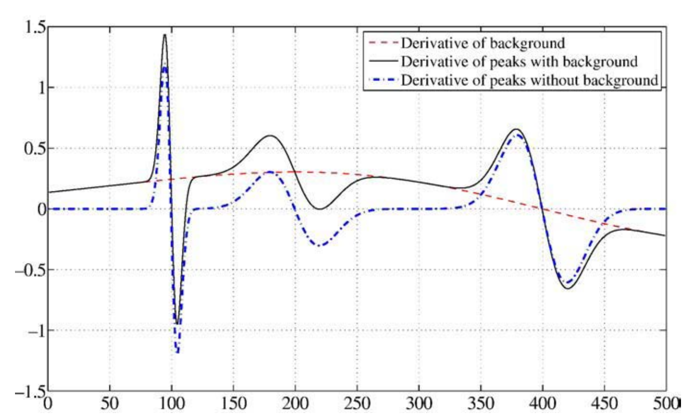

Fig. 3. The derivative of three Gaussian peaks with and without curved background.

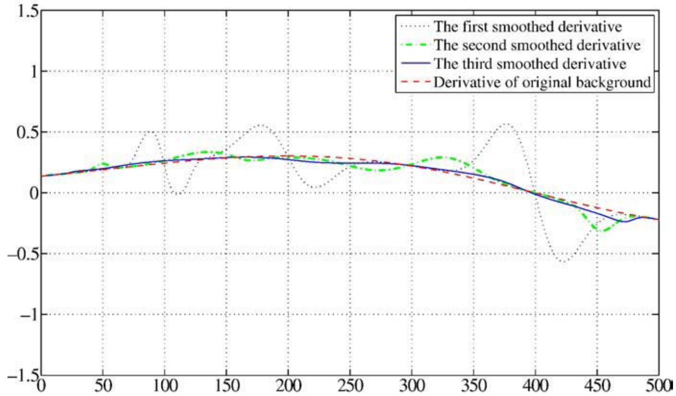

Fig. 4. The smoothed derivatives of three Gaussian peaks with curved background.

method, we used a simple moving average filter for a practical reason. It smoothes data by replacing each data point with the average of the neighboring data points defined within the span. This process is equivalent to lowpass filtering with the response of the smoothing given by the difference equation:

$$y[i] = \frac{1}{2N+1}\sum_{k=-N}^{N} x[i-k],$$

where $y[i]$ is the smoothed value for $x[i]$, N is the number of neighboring data points on either side of $x[i]$, and $2N+1$ is the span. In the following, we denote a moving average filter as $y = smoothing$ $(x, span)$.

In Fig. 4, we showed three successively smoothed derivatives along with the background's derivative. The span was set to about 20% of the length of total data. While the first smoothed derivative suppresses high peaks in the derivative ds, there remain some evident peaks. Since repeated use of a moving average filter does not affect low frequency components, we apply the smoothing filter again to moderate the peaks. As you see in the figure, the second and the third smoothed derivatives are more similar to the background's derivative than the first one. In this work, we used the third smoothed derivative, which is enough according to our experiments.

Let s3ds be the third smoothed derivative of ds, which can be considered as the estimate of the background's derivative, db. Then we can easily obtain the estimate of dg by subtracting s3ds from ds. The results are shown in Fig. 5. According to the figure, the estimated dg is quite similar to the original dg as expected.

Now we can determine peaks and their boundaries by inspecting the sign of the estimated dg. The detailed procedure is as follows. First find positions where the sign of dg changes from + to −, then mark them as peaks. From the peak position, find the local maximum on the left. Then successively find a local minimum position or the last position where dg is positive. The first position found is a left boundary.

Similarly, from the same peak position, find the local minimum on the right then find local maximum position or the last position where dg is negative. The position found is a right boundary.

As dg never gets to zero on the boundary sometimes, the condition that dg is positive or negative should be relaxed with the condition that dg is above some small positive value or below some small negative value in the implementation of an algorithm. Now we can obtain three peak regions, 80–120, 130–250, and 340–460 from Fig. 5.

After finding out all peak boundaries, we perform two adjustments for practical reason. If the peak width between the left and right boundary is smaller than a given threshold, we drop it since it should not be an important peak.

In addition to it, there are some overly estimated boundaries due to imprecise dg. To handle the problem, we draw a straight line between every left and right boundary. If there are some data below the line, we find the local minimum on the region and mark the position of it as a new boundary. The detailed procedures of adjustment are given in the following section.

2.3. Interpolation

After peak detection and adjustment, we have to replace the data in peak regions with some interpolated values which are the estimates of the background signal. We first tried a simple linear interpolation method [9], but it gave poor estimates especially when background signal is convex or concave like in the third peak region. So we used derivative information in s3ds to remedy the matter. Considering that s3ds is the background's derivative, we can obtain interpolated values by integrating s3ds in the peak region. However direct integration of s3ds gives some mismatched values on the peak boundaries since s3ds is only an estimate of dg. To fix the problem, we combined linear interpolation with the integration of s3ds.

For the discrete data, integration is approximated with the following cumulative sum, which is denoted as $y = cumsum(x)$.

$$y[k] = \sum_{i=1}^{k} x[i].$$

In order to avoid interpolation mismatch on the boundaries, we force the cumulative sum to start from zero and end in zero, which is

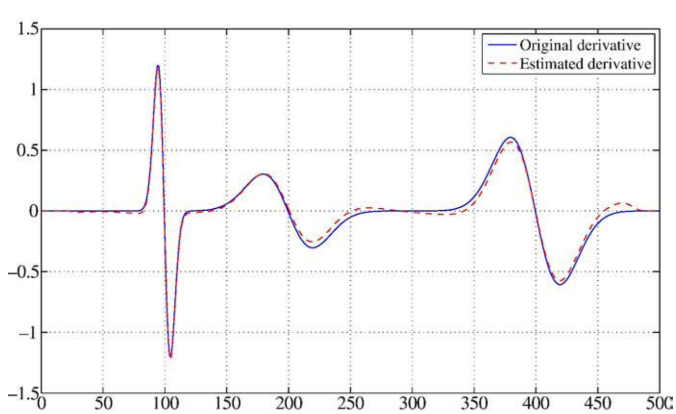

Fig. 5. The estimated derivative of three Gaussian peaks and the original derivative.

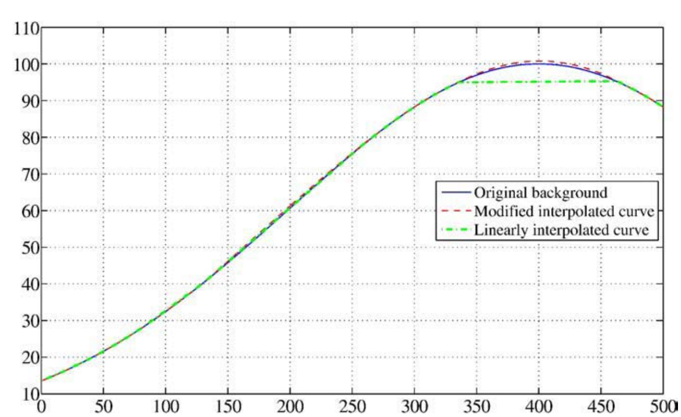

Fig. 6. The estimated background and the original.

accomplished by simply subtracting the mean from the derivative. Then we add the results to linearly interpolated values. The modified linear interpolation, mintline, on a given region is expressed as

$$\begin{aligned}&\mathit{diffline} = s3ds - \mathit{mean}(s3ds),\\ &\mathit{mintline} = \mathit{lintline} + \mathit{cumsum}(\mathit{diffline}),\end{aligned}$$

where *lintline* is the linearly interpolated curve.

The results of the modified linear interpolation are shown in Fig. 6. While the linearly interpolated curve differs much from the original background especially in the region from 340 to 460, the modified interpolated curve is quite similar to the original one.

As an estimated background curve is usually not so smooth if there are much more peaks in a spectrum, we fix it by smoothing the estimated background once again. Finally, background removed signal is easily obtained by subtracting the estimated background from the original spectrum. The background removed result of the given example is presented in Fig. 7. The figure clearly shows that the result is quite similar to the original Gaussian peaks.

3. Algorithms

The proposed method includes high-frequency noise removal, background's derivative finding, peak detection, interpolation, and background elimination. As smoothing function is repeatedly used in our method, we have to preset the span, which is the length of smoothing filter. Two kinds of span are used for the smoothing. One is span L_n for noise removal and the other is span L_b for background estimation. Large L_n gives more smoothed spectrum but it often eliminates small peaks. So it should be set to small so as not to affect important peaks. According to our experiments, the value below half length of the smallest width of important peaks would be enough. On the other hand, L_b should be set large enough to cancel the effect of the peaks of analyte. The value above the largest width of important peaks would be enough according to our experiments. But both of them can be varied according to the specific application.

3.1. Noise removal

Apply Savitzky–Golay filter with span L_n to the given spectrum. The degree of a polynomial is set to 2. Alternatively, simple averaging filter can be used for practical reason. In the following, s denotes the noise removed data.

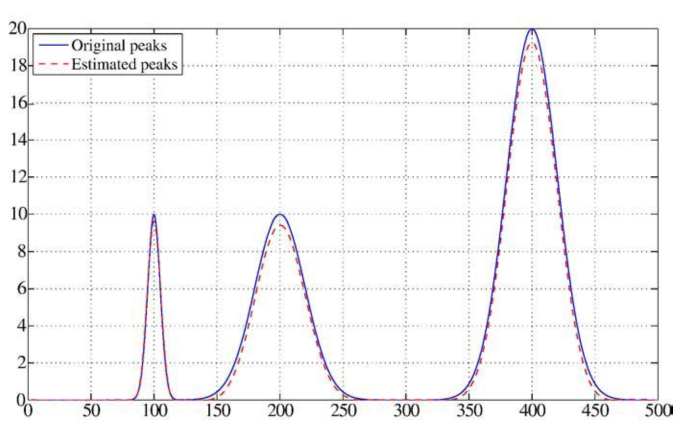

Fig. 7. The estimated Gaussian peaks and the original.

3.2. Finding derivatives

As the noise removed data still suffer from the remaining noise, we apply smoothing to s with span L_n before and after differentiation to make the derivative smoother.

$$ds = \mathit{smooth}(\mathit{diff}(\mathit{smooth}(s, L_n)), L_n).$$

Now the derivative of the background signal can be obtained by smoothing ds three times with span L_b.

$$s3ds = \mathit{smooth}(\mathit{smooth}(\mathit{smooth}(ds, L_b), L_b), L_b)$$

Then the derivative of a given spectrum without background is simply estimated as $dg = ds - s3ds$.

3.3. Peak detection

Find peak positions first where the sign of dg changes from + to −. Let the peak position be p_k, $k = 1,2,\cdots,m$, where m is the number of peaks found. For a given p_k, left and right boundary can be found in the following way.

Let i decrease from p_k. Find the position where $dg(i-1) < dg(i)$ for local maximum. From there, proceed to find the position where $dg(i-1) > dg(i)$ or $dg(i-1) < \mathit{threshold}$. The position $i-1$ is left boundary, pl_k. Similarly, let i increase from p_k. Find the position where $dg(i+1) > dg(i)$, then find the position where $dg(i+1) < dg(i)$ or $dg(i+1) > -\mathit{threshold}$. The position $i+1$ is right boundary, pr_k. The threshold is set to 0.02 in our experiments.

Once all pl_k and pr_k are found, adjust them in the following way. First check if $pr_k - pl_k < L_n$, then drop it. And draw a straight line from every $s(pr_k)$ to $s(pl_k)$. Let the line be $pline_k$ and the segment of s on the region be s_k. If there is a region where $s_k < pline_k$, then find the minimum position of $s_k - pline_k$ and mark the position as new boundaries.

3.4. Interpolation and background elimination

Interpolate every peak regions of s with a straight line. Let $lints_k$ be the linearly interpolated line and $s3ds_k$ be the parts of $s3ds$ in one peak region. Interpolated samples on that region, mb_k, are obtained in the following way.

$$\begin{aligned}&\mathit{diffline} = s3ds_k - \mathit{mean}(s3ds_k)\\ &mb_k = lints_k + \mathit{cumsum}(\mathit{diffline})\end{aligned}$$

Now background estimate, mb, can be obtained by aggregating interpolated samples of all peak regions together with untouched parts of s. Then smoothed background, b, is obtained as $b = \mathit{smooth}(mb, L_b)$.

The spectrum without background can be estimated with simple subtraction as $x = s - b$. In case x contains some negative values, fix it by adding a constant or forcing them to zero. In our experiments, a constant adding method is adopted.

4. Experimental results

We used two simulated spectral data sets and two experimental Raman spectra to validate the performance of the proposed method. Experiments were carried out with the Matlab software package (MathWorks, MA, USA) equipped with the Curve Fitting Toolbox.

4.1. Simulated data

The same data set as in [4] were used for simulated data. The data were intended to imitate real spectral data sets that contain varying

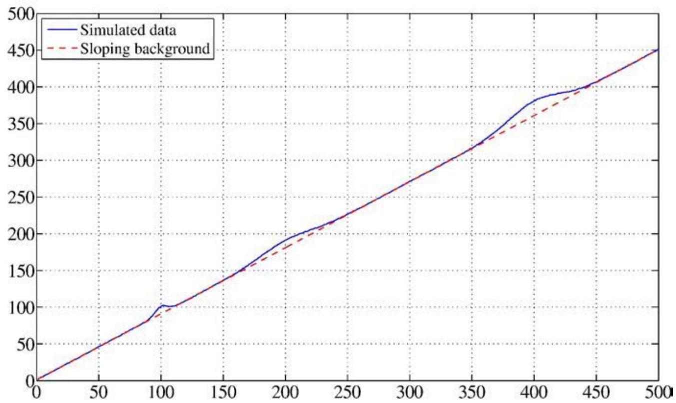

Fig. 8. Simulated data with the sloping background.

background, analytical signals, and random noise. A broader Gaussian peak is treated as curved background while a linear line is treated as sloping background. Narrow Gaussian peaks are treated as the spectra of interest. The simulated signal s is composed of pure signal g, background b, and random noise n. The simulated pure signal g contains three Gaussian peaks.

$$g[k] = 10\exp\left(-\frac{(k-100)^2}{2\cdot 5^2}\right) + 10\exp\left(-\frac{(k-200)^2}{2\cdot 20^2}\right) + 20\exp\left(-\frac{(k-400)^2}{2\cdot 20^2}\right).$$

The sloping and curved backgrounds are represented as $b[k]=1+0.9\ k$ and $b[k]=100\exp\left(-\frac{(k-400)^2}{2\cdot 200^2}\right)$, respectively. Noise n is the random real number between 0 and 1, $k=1,2,\cdots,500$. Two simulated data sets (one with the sloping background and the other with the curved background) are shown in Figs. 8 and 9 together with the respective background.

In the experiments with simulated data, L_n and L_b are set to 21 and 137, respectively. The experimental result with the sloping background is shown in Fig. 10. Except for somewhat underestimated peak at 200, it is clear that the proposed method gives a nearly perfect result. This underestimation is mainly due to inexact peak boundary determination.

The result with the curved background is shown in Fig. 11. In this case, underestimated peak is shown at 400. Without it, the result is very similar to that of the sloping background. Compared to the results in [4], both results are more accurate without any post processing.

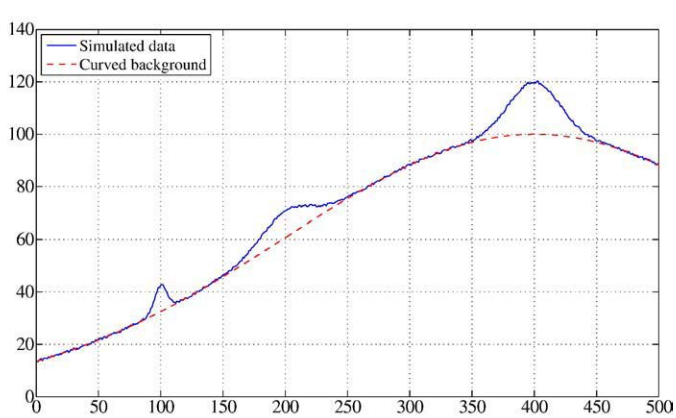

Fig. 9. Simulated data with the curved background.

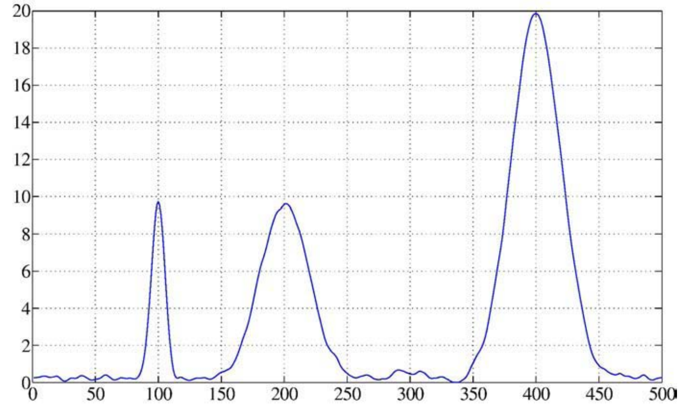

Fig. 10. Baseline eliminated signal for the data with the sloping background.

4.2. Experimental Raman spectrum

4.2.1. Measurement of experimental Raman Data

Excision specimens from one histologically proven adenocarcinoma of the stomach were obtained from an individual from the 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 before use. Cryosections (25 μm thickness) were obtained from the biopsy specimens and placed on a gold sheet for Raman spectroscopy. During the measurement, all the samples were thawed to reach room temperature in air.

A Renishaw Raman microspectrometer (Renishaw Raman system RM1000) was used for measurement. 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 do not cause any degradation to the samples. Spectra were measured from tissue and cell specimens with a ×20 short-working-distance objective, and the signals were integrated for a given time and measured over a spectral range of 600–1800 cm^{-1} with respect to the excitation frequency.

Rats were kept in the animal research facility in Renmin Hospital, Wuhan University. They were introperitoneally anesthetized by pentobarbital sodium (40 mg/kg) before operation. The incision was made on the back and the spinal cord was transected at T_{10-11}. Rats were killed by giving an excess amount of anesthetic. Bone tissues were obtained at 1, 2, 3, and 5 weeks after surgery. Femur tibia and humeri were excised from the rats and soft tissues were carefully removed. Then they were stored at −80 °C before use.

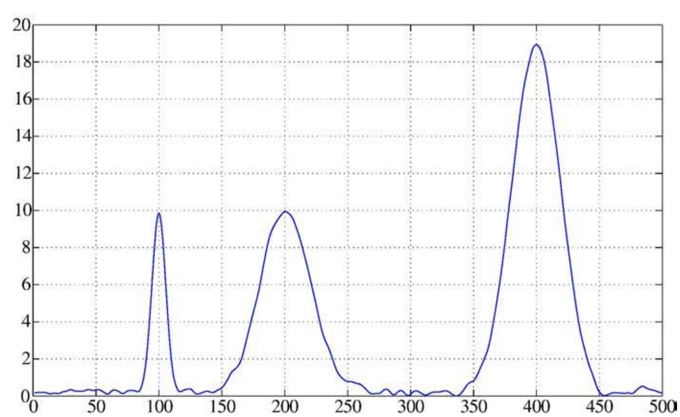

Fig. 11. Baseline eliminated signal for the data with the curved background.

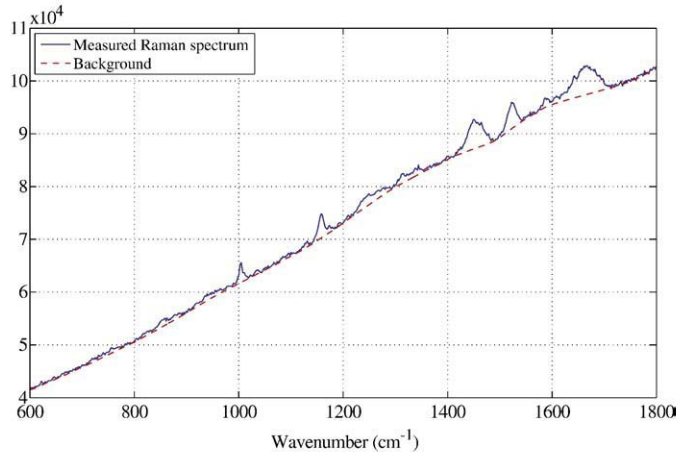

Fig. 12. Measured Raman spectrum of malignant gastric mucosa tissue.

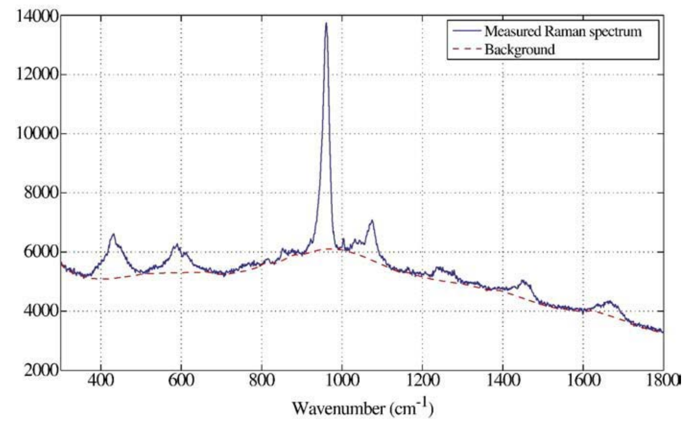

Fig. 14. Background eliminated Raman spectrum of malignant gastric mucosa tissue.

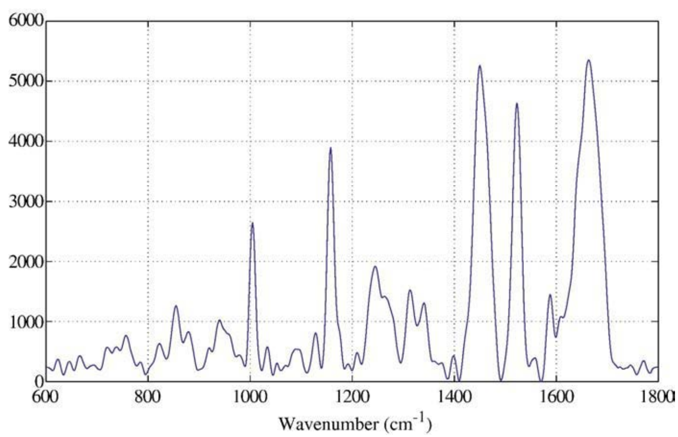

Fig. 13. Measured Raman spectrum of spinal cord injured rat's bone tissue.

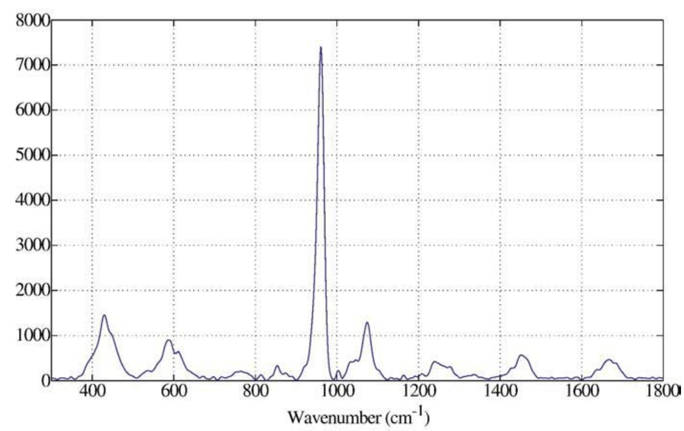

Fig. 15. Background eliminated Raman spectrum spinal cord injured rat's bone tissue.

Raman spectra for rat's bone were measured by HR800 Raman system (Jobin Yvon) with a microscope. Bones were cut vertically with low speed spindle at condylus femoris or caput humeri. 25 separate points were detected for each sample and the data were managed on an average. Measured spectral range is 300–1800 cm^{-1} with respect to the excitation frequency.

4.2.2. Background elimination of experimental Raman spectrum

Measured Raman spectrum of malignant gastric mucosa tissue and spinal cord injured rat's bone tissue are shown in Figs. 12 and 13. The estimated background is also shown in the respective figures. Roughly speaking, the first one resembles the sloping background while the second one resembles the curved background. Background eliminated spectra are shown in Figs. 14 and 15. From the visual inspection of the results, it is obvious that the background eliminated spectrum preserves significant peaks. Also it would make it easy to determine the height and position of the peaks.

Based on this work, we are currently focusing on resolving some shoulder peaks by post processing. In addition to it, we are studying the method extracting the effectual features of the pattern and classifying them into two or three categories to develop a new diagnostic method which would help the early detection of disease.

5. Conclusion

The objective of this paper is to develop a simple and computationally efficient method for the removal of noise and backgrounds in Raman spectra. For Raman measurement, the background has lower frequency components than the analyte signal. Hence, we could estimate the background by eliminating peaks with high-frequency components. Once the background is estimated, background elimination is achieved by simple subtraction. For this purpose, we presented a simple background estimation method, which consists of peak detection and interpolation based on the differentials of given data.

Experimental results show that the proposed method is successfully applied to experimental Raman data as well as simulated data. It means that the proposed method could handle various kinds of background. As the method requires no prior knowledge about the sample composition, no selection of suitable background correction points, and no mathematical assumption of the background distribution, we expect that the proposed method would be applied to other spectral data as well.

Acknowledgments

This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD, Basic Research Promotion Fund) (F01-2007-000-10133-0) and the NSFC-KOSEF Scientific Cooperation Program (fund number is 20711140390 supported by NSFC).

References

[1] H.-W. Tan, C.R. Mittermayr, S.D. Brown, Robust calibration with respect to background variation, Appl. Spectrosc. 55 (2001) 827–833.

[2] H.-W. Tan, S.D. Brown, Wavelets applied to removing non-constant, varying spectroscopic background in multivariate calibration, J. Chemom. 16 (2002) 228–240.
[3] P.J. Gemperline, J.H. Cho, B. Archer, Multivariate background correction for hyphenated chromatography detectors, J. Chemom. 13 (1999) 153–164.
[4] Y. Hu, T. Jiang, A. Shen, W. Li, X. Wang, J. Hu, A background elimination method based on wavelet transform for Raman spectra, Chemometr. Intell. Lab. Syst. 85 (2007) 94–101.
[5] A. Likar, T. Vidmar, A peak-search method based on spectrum convolution, J. Phys. D: Appl. Phys. 36 (2003) 1903–1909.
[6] The MathWorks, Curve Fitting Toolbox User's Guide, The MathWorks, Inc, USA, 2006.
[7] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes in C, Cambridge University Press, USA, 1992.
[8] A. Savitzky, M.J.E. Golay, Smoothing and differentiation of data by simplified least squares procedures, Anal. Chem. 36 (1964) 1627–1639.
[9] S.V. Vaseghi, Advanced Signal Processing and Digital Noise Reduction, Wiley, UK, 1996.