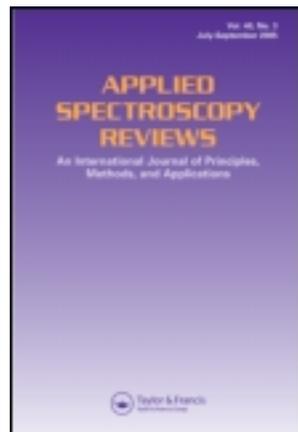


This article was downloaded by: [Duke University Libraries]

On: 24 July 2013, At: 05:56

Publisher: Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Applied Spectroscopy Reviews

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/laps20>

A WAVELET TRANSFORM AND ITS APPLICATION TO SPECTROSCOPIC ANALYSIS

Limin Shao^a, Xiangqin Lin^a & Xueguang Shao^b

^a Department of Chemistry, University of Science and Technology of China, Hefei, Anhui, 230026, P.R. China

^b Department of Chemistry, University of Science and Technology of China, Hefei, Anhui, 230026, P.R. China

Published online: 02 Feb 2007.

To cite this article: Limin Shao, Xiangqin Lin & Xueguang Shao (2002) A WAVELET TRANSFORM AND ITS APPLICATION TO SPECTROSCOPIC ANALYSIS, Applied Spectroscopy Reviews, 37:4, 429-450, DOI: [10.1081/ASR-120016391](https://doi.org/10.1081/ASR-120016391)

To link to this article: <http://dx.doi.org/10.1081/ASR-120016391>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>



APPLIED SPECTROSCOPY REVIEWS

Vol. 37, No. 4, pp. 429–450, 2002

**A WAVELET TRANSFORM AND
ITS APPLICATION TO
SPECTROSCOPIC ANALYSIS**

Limin Shao, Xiangqin Lin, and Xueguang Shao*

Department of Chemistry, University of Science and
Technology of China, Hefei, Anhui, 230026, P.R. China

CONTENTS

| | |
|---|-----|
| ABSTRACT | 430 |
| I. INTRODUCTION | 430 |
| II. THEORY AND ALGORITHMS | 431 |
| A. Theory of Wavelet Transform | 431 |
| B. Multiresolution Signal Decomposition (MRSD) | 432 |
| C. Improved Algorithm for Processing the Analytical Signal | 434 |
| III. APPLICATIONS | 434 |
| A. Signal-to-Noise Ratio (SNR) Improvement | 434 |
| B. Information Extraction | 437 |
| C. Resolution of Overlapping Signals..... | 441 |
| D. Data Compression | 444 |
| REFERENCES | 447 |

*Corresponding author. E-mail: xshao@ustc.edu.cn



ABSTRACT

An introduction to the wavelet transform and its applications in spectroscopy analysis are presented. Owing to the property of the dual localization both in time and in frequency domains, the wavelet transform exhibits several useful characteristics. With the wavelet transform a signal can be decomposed into series of contributions according to the frequency difference, these contributions are respectively called *discrete approximations* and *discrete details*. By applying certain treatment to the *discrete approximations* (the low frequency part) or the *discrete details* (the high frequency part), a part of the original signal within a certain frequency range can be obtained. The wavelet transform has been applied in various fields of the analytical chemistry, including removal of high frequency noise, information extraction, resolution of overlapping signals, and data compression etc. Research concerning application in analytical chemistry are introduced in respect of photoacoustic spectroscopy, EXAFS spectrum, NMR analysis, and Raman spectrum.

Key Words: Wavelet transform; Spectral analysis; De-noising; Resolution; Derivative; Data compression

I. INTRODUCTION

The wavelet transform originated in 1980 with Morlet^[1] working on seismic data analysis. Later, Grossmann and Morlet^[2] developed the geometrical formalism of the continuous wavelet transform based on invariance under the affine group, namely translation and dilation. In 1986, Daubechies et al.^[3] began research of the discrete wavelet transform by selecting a discrete subset called a *wavelet frame* from the continuous wavelet space. In 1988, Mallat^[4] introduced the concept of multiresolution analysis (MRA). This approach gives a general method for building orthogonal wavelet bases and leads to the implementation of fast wavelet algorithms.^[5,6]

The wavelet transform resembles the Fourier transform in some respects. The Fourier transform is localized in frequency but not in time domain. But the wavelet transform keeps the localization both in scale (frequency) and in position (time) domains. This dual localization makes the wavelet transform a powerful tool to study the local property of a signal and to eliminate the perturbation of the measuring error.



WAVELET TRANSFORM

431

By virtue of the dual localization, the wavelet transform soon got applied in various fields such as acoustic^[7] and fluid mechanics.^[8] The first application in analytical chemistry was to evaluate peak intensities in flow injection analysis (FIA) by Bos and Hoogendam.^[9] In 1994, they used the wavelet transform to preprocess IR spectra in the identification of substituted benzenes.^[10] In recent years, much research was performed in the de-noising,^[11–20] data compression,^[19–22] and analysis of complex signals including resolution of chromatographic peaks,^[23–25] NMR analysis,^[26,27] near infrared spectroscopy,^[28] electroanalytical chemistry,^[29] EXAFS analysis,^[30] chromatography analysis,^[31–33] and biology relevant analysis.^[34] The combination of the wavelet transform with other chemometric methods was also studied.^[35–39]

Our research concerning the wavelet transform in recent years has been focused on chromatography and spectroscopic analysis. In this paper, the implementation of the discrete wavelet transform to denoise and resolve spectroscopic data is discussed. A new algorithm based on the Mallat pyramid is described. Work performed in our laboratory in spectroscopic analysis is reported here as well as research by other workers.

II. THEORY AND ALGORITHMS

A. Theory of Wavelet Transform

If a function $\psi(x) \in L^2(R)$ ($L^2(R)$ denotes the Hilbert space of measurable, square integrable functions) satisfies the condition

$$\int_{-\infty}^{+\infty} |\hat{\psi}(x)| \cdot |x|^{-1} dx < \infty, \quad (1)$$

where $\hat{\psi}(x)$ denotes the Fourier transform of $\psi(x)$, it is called a ‘mother wavelet’.

A series of ‘self-similar’ wavelets can be generated from a mother wavelet by two kinds of processes, dilation and translation. If variable a and b are used to control the dilation and the translation respectively, the wavelet is defined as

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \quad a, b \in R, \quad a \neq 0, \quad (2)$$

where the normalization term $1/\sqrt{|a|}$ keeps the energy of the wavelet the same as that of the mother wavelet.



The wavelet transform is defined as the projection of a function $f(t)$ ($f(t) \in L^2(R)$) onto the wavelet

$$\begin{aligned} Wf(a, b) &= \int_{-\infty}^{+\infty} f(t)\psi_{a,b}^*(t) dt \\ &= \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} f(t)\psi^*\left(\frac{t-b}{a}\right) dt, \end{aligned} \quad (3)$$

where the superscript * represent the complex conjugate. If we denote the dilation of $\psi(x)$ with the scale parameter a by $\psi_a(x) = \sqrt{|a|}\psi(ax)$ and denote $\psi(-x)$ by $\tilde{\psi}_a(x)$, the transform at a point u and a scale a can be written as a convolution product with $\tilde{\psi}_a(x)$

$$W_a f(u) = f * \tilde{\psi}_a(u). \quad (4)$$

Therefore, the wavelet transform can be viewed as a filtering of $f(x)$ with a band-pass filter whose impulse response is $\tilde{\psi}_a(x)$.

The inverse transform of $Wf(a, b)$ can be described as

$$f(t) = \frac{1}{C} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} Wf(a, b)\psi_{a,b}\left(\frac{t-b}{a}\right) \frac{da db}{a^2}, \quad (5)$$

where

$$C = \int_{-\infty}^{+\infty} \frac{\hat{\psi}^*(\omega)\hat{\psi}(\omega)}{\omega} d\omega.$$

In practice applications, the discrete wavelet transform is generally used, i.e., the values for variables a and b are restricted as follows

$$\begin{cases} a = a_0^m \\ b = b_0^n \end{cases}, \quad (m, n \in \mathbb{Z}, a_0 \neq 0). \quad (6)$$

Consequently, the discrete form of Eqs. (2) and (3) can be expressed as

$$\begin{aligned} \psi_{m,n}(t) &= a_0^{-m/2}\psi(a_0^{-m}t - nb_0) \\ Wf(m, n) &= a_0^{-m/2} \int_{-\infty}^{+\infty} f(t)\psi(a_0^{-m}t - nb_0) dt. \end{aligned} \quad (7)$$

B. Multiresolution Signal Decomposition (MRSD)

In order to implement the computation of the discrete wavelet transform, several algorithms have been developed.^[2,5,6] The multiresolution signal decomposition (MRSD) by Mallat^[5,6] is commonly used.



C. Improved Algorithm for Processing the Analytical Signal

The MRSD algorithm requires 2^N where N is an integer. This number is reduced to half after each decomposition, which makes the decomposed results e.g., spectra, or chromatograms, discontinuous. Therefore, Eqs. (10) and (11) are reformulated as

$$C^{(j)}(n) = \sum_{l=0}^{L^{(j)}-1} \tilde{h}^{(j)}(l)C^{(j-1)}(n-l), \quad (16)$$

$$D^{(j)}(n) = \sum_{l=0}^{L^{(j)}-1} \tilde{g}^{(j)}(l)C^{(j-1)}(n-l), \quad (17)$$

where $L^{(j)}$ is the data number of $\tilde{H}^{(j)} = \{\tilde{h}^{(j)}(l), l \in Z\}$ or $\tilde{G}^{(j)} = \{\tilde{g}^{(j)}(l), l \in Z\}$, which are generated by inserting $2^{j-1} - 1$ zeros into every adjacent item of \tilde{H} or \tilde{G} , and the value of $L^{(j)}$ will be doubled when j increases. These two equations can process signals of any length, and smooth decomposed results can be obtained. In most cases, they are more suitable than Eqs. (10) and (11).

III. APPLICATIONS

A. Signal-to-Noise Ratio (SNR) Improvement

Noise and signal are simultaneously collected by an analytical instrument. The irreproducibility of the noise makes it difficult to analyze the original signal. Therefore, the technique of de-noising is very important in many fields. A number of filtering methods have been developed, such as Fourier filtering, Savitzky-Golay smoothing, and Kalman filtering etc. In spite of the existence of the methods, the problems of complicated computation and poor de-noising performance still exist.

The wavelet transform has been proven to be an effective method for de-noising chemical data.^[19] The general procedure of the de-noising with the wavelet transform can be done in three steps. (1) Perform the wavelet transform of a noisy signal f_{noise} , and obtain the wavelet coefficients W ; (2) suppress those elements in W that are determined to be from noise by certain thresholding method; (3) apply the inverse wavelet transform to the suppressed W to obtain the de-noised signal $f_{\text{de-noised}}$. There are several methods^[11,14,19] to estimate the threshold and to perform the thresholding.

In our studies,^[33] the de-noising based on the improved algorithm in Sec. II.C was preferred. Because both the *discrete approximations* and the



WAVELET TRANSFORM

435

discrete details have the same data length as the original signal, it is unnecessary to go through the whole three steps, we can obtain de-noised result simply by selecting one of the *discrete approximations*. This improved algorithm was applied to the de-noising of the photoacoustic spectrum of $\text{NdCl}_3 \cdot 6\text{H}_2\text{O}$.^[40]

Figure 1A and 1B are the decomposed results from the normalized PA spectrum of $\text{NdCl}_3 \cdot 6\text{H}_2\text{O}$. The Symmlet ($N=8$) wavelet was used. Curves C and D in Fig. 1 respectively denote the *discrete approximations* and the *discrete details*, and $C^{(0)}$ denotes the original spectrum. As $C^{(0)}$ shows, the noise made it difficult to determine the positions of the peaks, especially in the region of 300–450 nm. Therefore, direct analysis from the raw spectrum is surely inaccurate. From Fig. 1, it is apparently that with the wavelet transform the noise is gradually separated into the *discrete details*, $D^{(j)}$ s, and that the amplitude of the noise in the *discrete approximations*, $C^{(j)}$ s, is gradually low. It can also be found that $D^{(1)}-D^{(5)}$ are the high frequency noise, and that the photoacoustic information begins to emerge in $D^{(6)}$. Meanwhile, the photoacoustic information begins to be lost from $C^{(7)}$, which can be found from the comparison between $C^{(6)}$ and $C^{(7)}$ in the region of 300–500 nm. Therefore, $C^{(6)}$ is chosen as the de-noised spectrum. It can be seen that the SNR of the de-noised result was greatly improved. Consequently, the positions of the peaks in the region of 300–450 nm can be easily determined. Furthermore, the correlation between intramolecular energy transition bands and spectral range was compared in Ref. [40], it was found that all the peaks' positions are consistent with the theoretical energy transition bands.

We also applied the wavelet transform to de-noise Raman spectra.^[41] Ethanolic solutions of CCl_4 (liquid in a capillary with concentration of 30, 45, 60, 75, and 90%, V/V) were prepared. Figure 2 shows the Raman spectra of the solutions.

It can be seen that all the spectra suffer from a high noise level, especially that the peak almost disappears when the concentration is 30%. These spectra are decomposed by our improved algorithm with Haar wavelet. According to the same criteria introduced above, the *discrete approximations*, $C^{(6)}$ s, are chosen as the de-noised results shown in Fig. 3.

From Fig. 3 it can be seen that, compared with that of the original spectra, the SNR of de-noised signals is greatly improved, especially in the cases of concentrations being 30 and 45% where the signal peaks are almost invisible in the original spectra. But the wavelet transform yields meaningful de-noised results. Owing to the high level and the non-uniform distribution of the noise, the baselines of the five de-noised spectra in Fig. 3 are not restrictedly zero, and side bands appear. However, the intensity of the side bands is very weak compared with that of the signal peak, and the

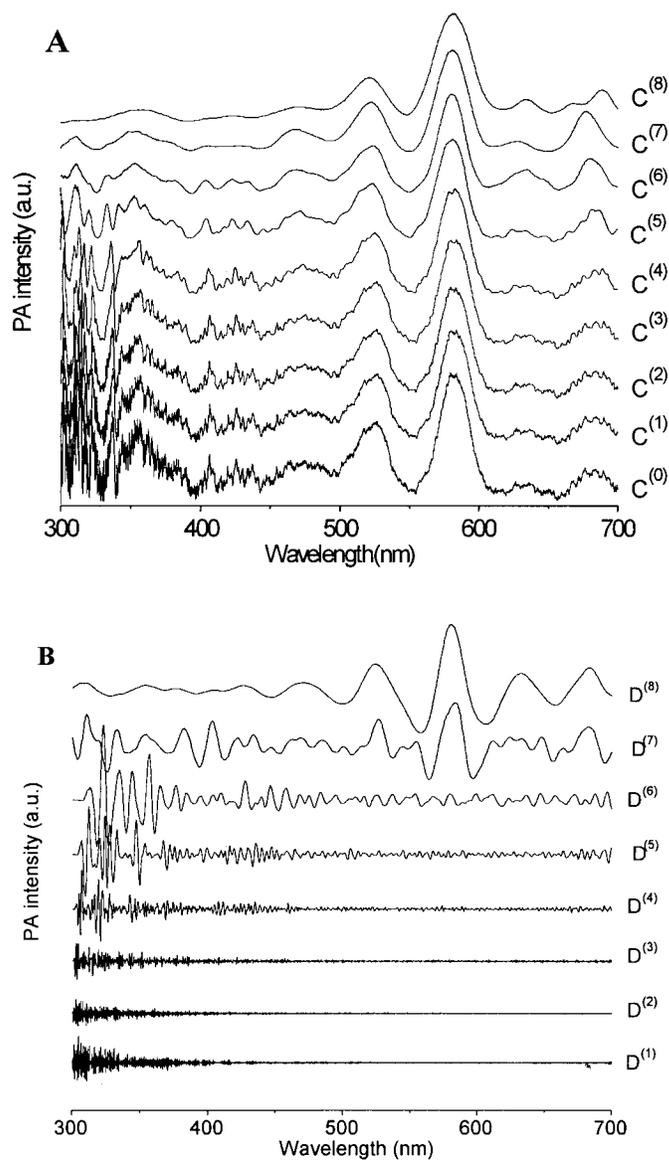


Figure 1. The normalized PA spectrum of $\text{NdCl}_3 \cdot 6\text{H}_2\text{O}$, and the decomposed results by wavelet transform. (A) The *discrete approximations*; (B) the *discrete details*.



WAVELET TRANSFORM

437

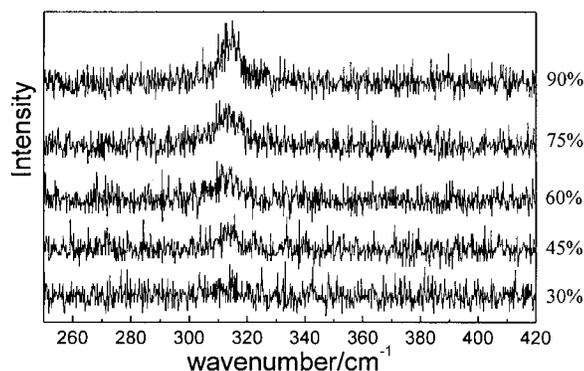


Figure 2. Raman spectra of ethanolic solutions of CCl_4 of different concentrations.

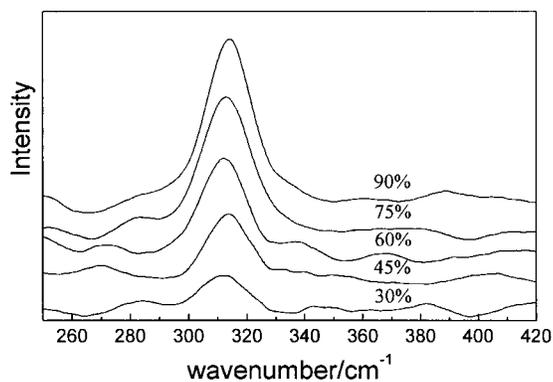


Figure 3. Filtered results from the Raman spectra in Fig. 2, which correspond to the *discrete approximations*, $C^{(6)}$ s, in the wavelet decompositions.

comparative intensity of the side bands decreases with the increase of the SNR of the original spectra.

B. Information Extraction

From the theory of MRSD in Sec. II.B, it is concluded that the wavelet transform can decompose a given signal into two kinds of contributions. One is the high frequency (between 2^j and 2^{j+1}) parts, i.e., the *discrete details*, $D^{(j)}(n)$ s, and the other is the low frequency (lower than 2^j) parts,



i.e., the *discrete approximations*, $C^{(j)}(n)$ s. If all the $D^{(j)}(n)$ s are set zero, and the reconstruction is done by Eq. (15), the reconstructed signal is the low frequency part of the original signal. Similarly, if all the $C^{(j)}(n)$ s are set zero, the reconstructed signal is the high frequency part of the original signal. Therefore, wavelet transform can be used to extract the information of different frequency from the original signal.

Wavelet transform was applied to extract the extended X-ray absorption fine structure (EXAFS) information from the measured spectrum.^[42] The conventional procedure is to use cubic spline interpolation to obtain the background absorption, then subtract it from the measured spectrum. However, in the conventional process, several points (usually 3–5) must be selected from the original spectrum for the cubic spline interpolation. The manual selection is tedious and time-consuming, and the result by trial and inspection surely results in poor reproducibility. There exists the fact that the measured EXAFS spectrum is composed of the high frequency EXAFS oscillation information and the smooth low frequency background absorption. Based on the fact, with the wavelet transform the EXAFS oscillation information can be extracted because its frequency is higher than the background's.

Figure 4 is a spectrum of Cu in k -space obtained (with $k = [0.263 \times (E - E^0)]^{1/2}$, and $E^0 = 8393.5$ eV) from the measured spectrum in E -space. The decomposed results by the wavelet transform are shown in Fig. 5 using

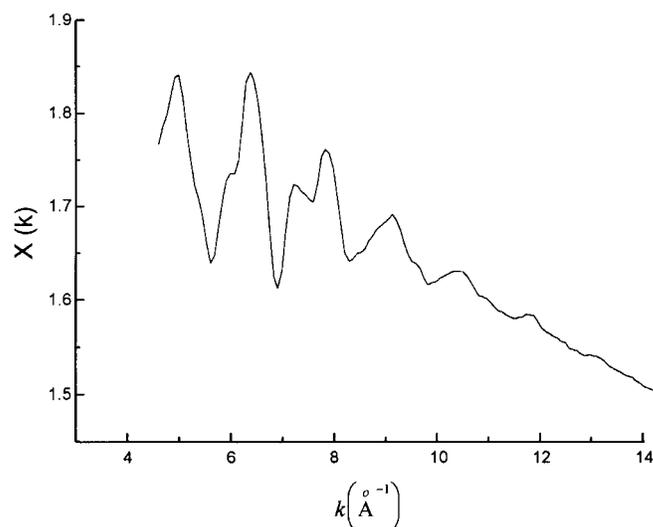


Figure 4. The experimental EXAFS spectrum of Cu in k -space ($E^0 = 8393.5$ eV).



WAVELET TRANSFORM

439

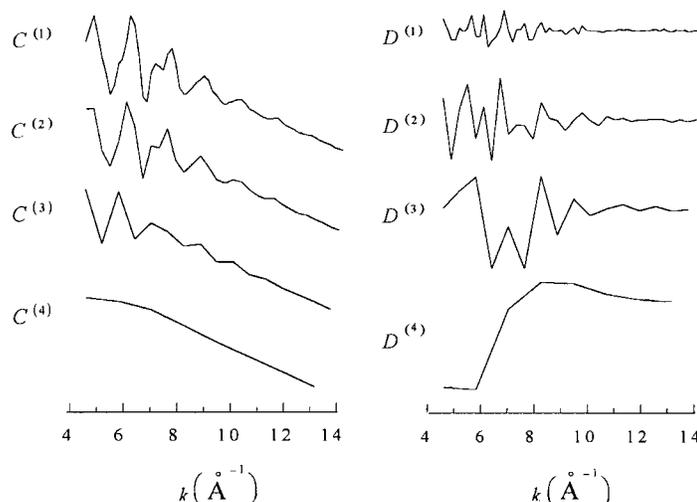


Figure 5. Decomposed results by the wavelet transform from the spectrum in Fig. 4.

Daubechies ($N=4$) wavelet. It is clear that the EXAFS oscillation information is gradually separated from the *discrete approximations*, $C^{(j)}$ s, into the *discrete details*, $D^{(j)}$ s. The characteristic of the background absorption gradually emerges in the $C^{(j)}$ s, and is totally exhibited by $C^{(4)}$. Therefore, the extraction can be performed by setting all the $C^{(j)}$ s zero in the reconstruction.

The extracted spectrum is shown in Fig. 6 in solid curve. As a comparison, the result of the conventional cubic spline interpolation is also given as dotted curve. It can be found that the main characteristics of the two curves are similar. But in the high k region, the amplitude of the signal by the wavelet transform is greater than that by the conventional method, the higher SNR means that more accurate results could be obtained. It should be mentioned here that although the dotted curve in Fig. 6 is not the best result of the conventional cubic spline interpolation, it was obtained from many times of trials and inspections.

The structural parameters of the first coordination shell of Cu obtained from the extracted EXAFS oscillation (the solid and the dotted curves in Fig. 6) are listed in Table 1. They are the coordination number, N , the Debye-Waller factor, σ , the coordination distance, r , and the electron mean free path, λ . The errors of curve fitting are also given in Table 1.

The curve fitting errors in Table 1 show that the data analysis based on the wavelet transform is more accurate than that based on the cubic spline interpolation. Consequently, the results by the wavelet transform are more

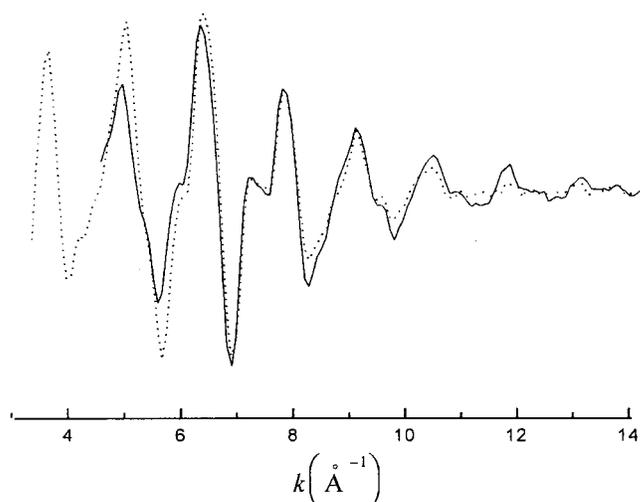


Figure 6. The extracted EXAFS oscillation by the wavelet transform (the solid curve), and the result obtained by the conventional cubic spline interpolation (the dotted curve).

Table 1. Comparison of the Structural Parameters Obtained by Least Square Curve Fitting from the EXAFS Oscillations Extracted by the Wavelet Transform and the Cubic Spline Interpolation

| Spectrum | Method | N | r | σ | λ | Error |
|----------|--------------|-----|------|----------|-----------|--------|
| I | Wavelet | 12 | 2.50 | 0.112 | 6.5 | 0.0026 |
| | Cubic spline | 8 | 2.52 | 0.077 | 4.4 | 0.0074 |
| II | Wavelet | 12 | 2.49 | 0.109 | 6.3 | 0.0027 |
| | Cubic spline | 9 | 2.53 | 0.082 | 4.5 | 0.0029 |
| III | Wavelet | 12 | 2.51 | 0.113 | 6.3 | 0.0011 |
| | Cubic spline | 9 | 2.52 | 0.084 | 4.5 | 0.0056 |

credible. Overall, the values obtained by the wavelet transform are more in agreement with those values by other methods.^[43] On the other hand, the values of the structural parameters show that the reproducibility of the wavelet transform method is obviously superior to that of the cubic spline interpolation. The reason of the superiority is that there was no user's interference in the information extraction by the wavelet transform, while the interpolating points have to be specified by the user in the background absorption removal by the cubic spline interpolation.



C. Resolution of Overlapping Signals

Nuclear magnetic resonance (NMR) is a powerful tool in analytical chemistry, organic chemistry, and bioorganic chemistry. But it is always difficult to obtain the structural information of a sample when the peaks of its NMR spectrum are overlapping. Though several techniques^[44-47] had been developed to enhance the resolution of NMR spectra, the problems of low resolution in the spectra of most biological macro-molecules still exist.

As mentioned above, the wavelet transform can decompose a signal into contributions of different frequency. Therefore, if we amplify one or some of the relatively high frequency contributions, i.e., *the discrete details*, which represent the useful information, then the frequency of the reconstructed signal will be increased. In another word, with the wavelet transform, an overlapping spectrum can be further resolved.^[48]

Figure 7 shows the experimental NMR spectra of a biomolecule (gramicidin-S) on a Bruker DMX-500 NMR spectrometer (Switzerland) at room temperature. In Fig. 7, spectra (a) and (b) were measured with low resolution for the investigation, and spectrum (c) with the highest resolution for comparison. In order to improve the resolution of the spectra, the original spectra were decomposed using Daubechies ($N=4$) wavelet, since it is the most suitable for resolution.^[48,49] The *discrete details* $D^{(1)}$ and $D^{(2)}$ are chosen to be amplified 10 times, then the reconstruction was performed.

For comparison, parts of the spectra between 4.0 and 5.0 ppm are shown in detail in Fig. 8. It can be seen clearly that the resolution is greatly improved. Three groups located at about 4.75, 4.57, and 4.30 ppm are well resolved.

The derivative spectrum is another commonly used technique to analyze the overlapping spectrum. Generally, the calculation is done with several methods, such as numerical differentiation, Fourier transform, and Savitzky-Golay polynomial filtering, etc. However, these methods cannot yield smooth curves when treating the noise involved spectrum, which affects the accuracy. Leung et al.^[50] attempted to use the discrete wavelet transform to perform the differentiation, and obtained high SNR derivative spectra. But the high order derivative with their method is seriously distorted when the original data number is small, and for a low SNR spectrum, their method suffers from complexity in computation and decreasing in accuracy.

According to our research,^[51] a novel method to perform the differentiation was proposed. The method is based on the continuous wavelet transform using the Haar wavelet. The new method was tested with the simulated Gaussian peak with 10% white noise added. The original simulated spectrum

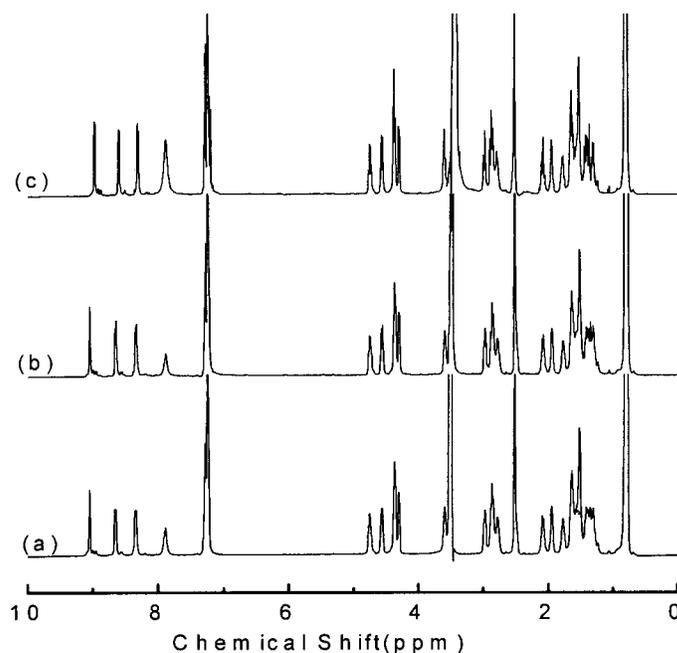


Figure 7. The NMR spectra of gramicidin-S at different resolution. (a) The NMR of the lowest resolution, (b) the NMR of the middle resolution, (c) the NMR of the highest resolution.

is shown in Fig. 9 as curve (a). Curves (b)–(f) are the results from the above-mentioned methods. It is obvious that the derivative spectrum obtained with the continuous wavelet method is smooth and clean, the SNR is greatly improved compared with the results of the other methods. Moreover, the SNR is even higher than that of the original spectrum. It is because the Haar wavelet is a filter, in the derivative calculation, the smoothing is performed simultaneously. Therefore, the method based on the continuous wavelet transform is superior to the mentioned methods in treating noisy data.

The effect of the dilation factor a was studied with simulated data. The result is that the larger the factor a is, the more smooth the result is, but large factor a may eliminate some small peaks. Therefore, in practice uses, the value of factor a should be determined properly.

The performance of the novel method was further investigated with the photoacoustic spectrum of $\text{Pr}(\text{Gly})_3\text{Cl}_3 \cdot 3\text{H}_2\text{O}$. From the original spectrum, curve (a) in Fig. 10, it is clear that no accurate determination of any



WAVELET TRANSFORM

443

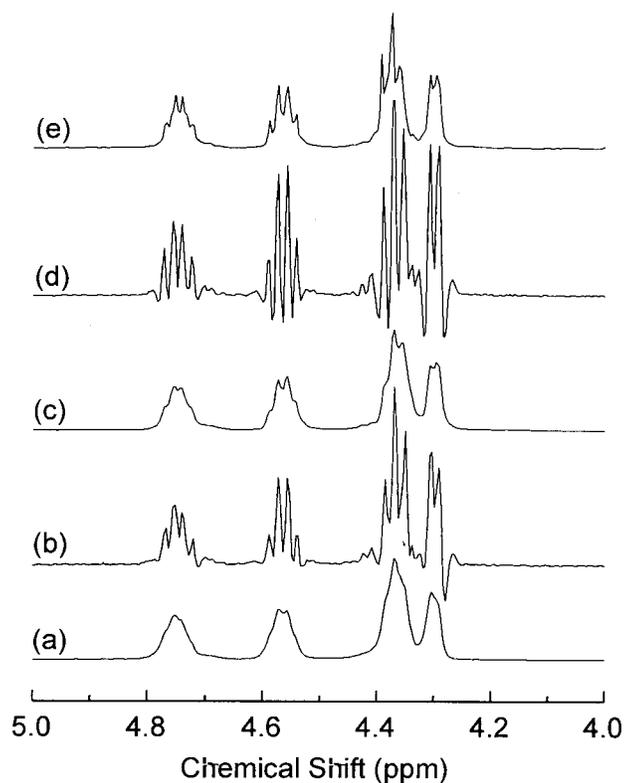


Figure 8. The comparison between the resolved and the experimental NMR spectra. (a) The experimental spectrum of the low resolution; (b) The resolved spectrum from (a); (c) The experimental spectrum of the middle resolution; (d) The resolved spectrum from (c); (e) The experimental spectrum of the highest resolution.

peak positions can be performed due to the noise. According to our research, neither the numerical differentiation nor the other methods mentioned above could yield high SNR derivative spectra. However, the continuous wavelet method yields satisfactory results. In Fig. 10 it can be found that the derivative spectra are all smooth and clean. Although there is small fluctuation in the baselines, it does not affect the positions of the peaks. The peaks' positions (for the second derivative spectrum) and the zeros' positions (for the first derivative spectrum) are all clear enough to indicate the peaks' positions of the original spectrum. The correlation between intramolecular energy transition bands and spectral range is listed in Table 2. It can be seen that the consistence is very good.

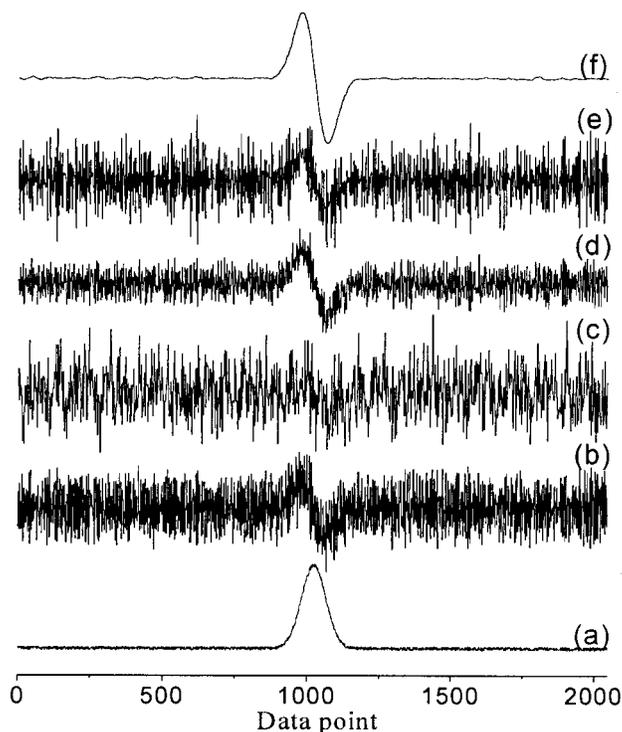


Figure 9. The simulated Gaussian signal with noise added (a), and its derivative spectra by numerical differentiation (b), Fourier transform (c), Savitzky-Golay method (d), Leung's method (e), and our proposed method with $a = 60$ (f).

D. Data Compression

One important technique of modern instrumental analysis is the high-speed data acquisition, which produces more abundant information. However, this technique brings side effects too. More space is needed for the storage, and in some cases the time consumed by the analysis is too long because of the redundant data. Therefore, pretreatment to remove the redundant information from the raw analytical signal is necessary, i.e., the employment of compressing tools.

The wavelet transform has been proven to be an efficient tool for data compression. Chau et al.^[21] applied fast wavelet transform (FWT) to compress Ultraviolet-Visible spectra. They tested different Daubechies wavelets, thresholding values, and maximum decomposition levels, and found that



WAVELET TRANSFORM

445

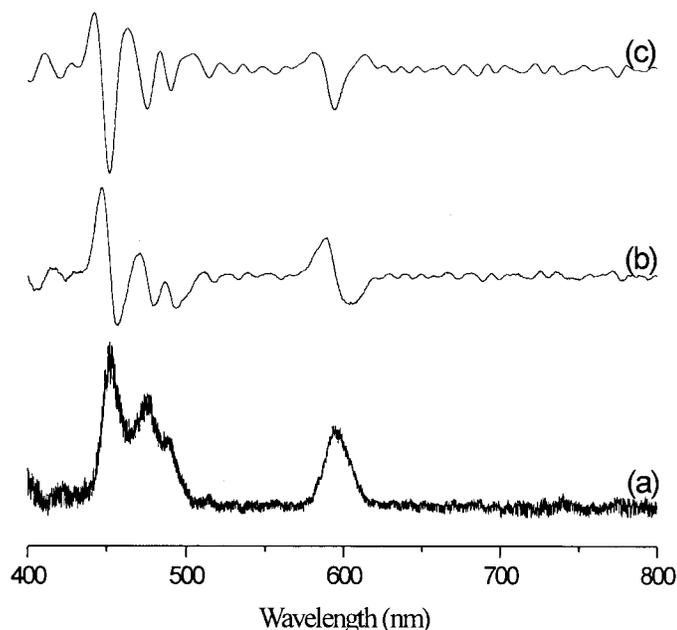


Figure 10. The photoacoustic spectrum of $\text{Pr}(\text{Gly})_3\text{Cl}_3 \cdot 3\text{H}_2\text{O}$ (a) and its first (b) and second derivative (c) spectra with $a = 100$.

Daubechies wavelet ($N = 16$), thresholding 0.003, and decomposition level 6 are good and efficient coefficients in compressing the UV-vis spectra with significant storage reduction. Chau's group^[52] later employed biorthogonal filter bank (BFB) technique to solve the side-lobe problem. They found that the linear phase property of filters with the adoption of the symmetric extension method could effectively solve that problem. They also used other method, such as differential pulse code modulation (DPCM), optimal bit allocation (OBA), and variable length coding (VLC) etc., to further improve the compressing efficiency. Wavelet packet transform (WPT) was also applied to the compression^[20] and the library search^[53] of infrared spectra.

In our study, a NMR spectrum of biological molecules composed of 32,768 points was investigated. Spectrum (a) in Fig. 11 is the measured spectrum, and spectra (b), (c) and (d) are the reconstructed spectra from 2048, 1024, and 512 wavelet coefficients respectively. In Fig. 11, it can be seen that there is almost no difference between the spectra. The effect of the wavelet basis on the reconstructed result, the degree of



Table 2. The Correlation Between the Positions of the Peaks Determined from the Derivative Spectrum and the Corresponding Transition Bands of $\text{Pr}(\text{Gly})_3\text{Cl}_3 \cdot 3\text{H}_2\text{O}$

| Positions of the Peaks (nm) | Spectral Range (nm) | Transition Bands | Group State |
|-----------------------------|---------------------|------------------|-------------|
| 451.6 | 400–460 | 3P_2 | 3H_4 |
| 475.4 | 460–500 | 3P_1 | |
| 490.4 | | 3P_0 | |
| 594.3 | 500–636 | 1D_2 | |

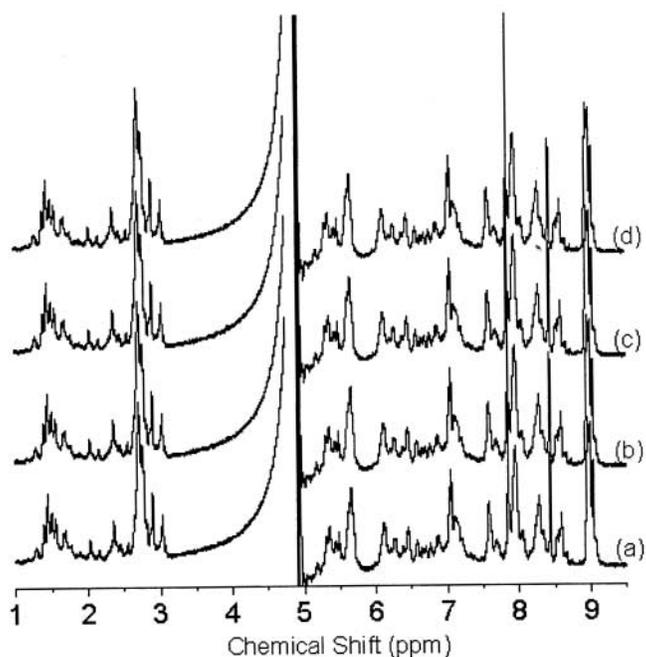


Figure 11. Comparison of the original spectrum (a) and the reconstructed spectra from 2048 (b), 1024 (c) and 512 (d) wavelet coefficients.

the decomposition, etc., were also discussed. Results are that, among the wavelets of Haar, Daubechies, and Symmlet, the Symmlet ($N=0$) is the most suitable wavelet for the compression, and the number of the decomposition should be greater than 8.



REFERENCES

1. Morlet, J. Sampling Theory and Wave Propagation. *Proc. 51st Annu. Meet. Soc. Explor. Geophys.*; Los Angeles, USA, 1981.
2. Grossmann, A.; Morlet, J. Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape. *SIAM J. Math. Anal.* **1984**, *15*(4), 723–736.
3. Daubechies, I.; Grossmann, A.; Meyer, Y. Painless Nonorthogonal Expansions. *J. Math. Phys.* **1986**, *27*, 1271–1283.
4. Mallat, S. Review of Multifrequency Channel Decompositions of Images and Wavelet Models. In *Acoustic, Speech and Signal Processing, Multidimensional Signal Processing*; IEEE Special Issue, 1988.
5. Mallat, S. A Theory for Multiresolution Signal Decomposition: the Wavelet Representation. *IEEE, Trans. on Pattern Anal. Machine Intell.* **1989**, *11*, 674–693.
6. Mallat, S. Multiresolution Approximation and Wavelet Orthonormal Bases of $L^2(\mathbb{R})$. *Trans. Am. Math. Soc.* **1989**, *315*, 69–87.
7. Kronland-Martinet, R.; Morlet, J.; Grossmann, A. Analysis of Sound Patterns Through Wavelet Transforms. *Int. J. Pattern Recognit. Artif. Intell.* **1987**, *1*(2), 273–302.
8. Liandrat, J.; Morlet-Bailly, F. The Wavelet Transform: Some Applications to Fluid Dynamics and Turbulence. *Eur. J. Mechanics, B: Fluids* **1990**, *9*(1), 1–19.
9. Bos, M.; Hoogendam, E. Wavelet Transform for the Evaluation of Peak Intensities in Flow-injection Analysis. *Anal. Chim. Acta* **1992**, *267*, 73–80.
10. Bos, M.; Vrieling, J.A.M. The Wavelet Transform for Preprocessing IR Spectra in the Identification of Mono- and Di-substituted Benzenes. *Chemom. Intell. Lab. Syst.* **1994**, *23*, 115–122.
11. Alsberg, B.K.; Woodward, A.M.; Winson, M.K.; Rowland, J.; Kell, D.B. Wavelet Denoising of Infrared Spectra. *Analyst* **1997**, *122*, 645–652.
12. Zou, X.L.; Mo, J.Y. Spline Wavelet Multifiltering Analysis. *Chin. Sci. Bullet.* **1997**, *42*(8), 640–644.
13. Wolkenstein, M.; Hutter, H.; Grasserbauer, M. Wavelet Filtering for Analytical Data. *Fresenius J. Anal. Chem.* **1997**, *358*(1/2), 165–169.
14. Mittermayr, C.R.; Rosenberg, E.; Grasserbauer, M. Detection and Estimation of Heteroscedastic Noise by Means of Wavelet Transform. *Anal. Commun.* **1997**, *34*, 73–75.
15. Lu, X.Q.; Mo, J.Y. Spline Wavelet Multifrequency Channel Filters for High Noise Digital Signal Processing in Voltammetry. *Acta Sci. Nat. Univ. Sunyatseni* **1997**, *36*, 129–130.



16. Wolkenstein, M.; Hutter, H.; Nikolov, S.G.; Grasserbauer, M. Improvement of SIMS Image Classification by Means of Wavelet De-noising. *Fresenius J. Anal. Chem.* **1997**, *357*(7), 783–788.
17. Mittermayr, C.R.; Nikolov, S.G.; Hutter, H.; Grasserbauer, M. Wavelet Denoising of Gaussian Peaks: A Comparative Study. *Chemom. Intell. Lab. Syst.* **1996**, *34*, 187–202.
18. Mittermayr, C.R.; Lendl, B.; Rosenberg, E.; Grasserbauer, M. The Application of The Wavelet Power Spectrum to Detect and Estimate 1/f Noise in the Presence of Analytical Signals. *Anal. Chim. Acta* **1999**, *388*(3), 303–313.
19. Barcly, V.J.; Bonner, R.F.; Hamilton, I.P. Application of Wavelet Transforms to Experimental Spectra: Smoothing, Denoising and Data Set Compression. *Anal. Chem.* **1997**, *69*(1), 78–90.
20. Walczak, B.; Massart, D.L. Noise Suppression and Signal Compression Using the Wavelet Packet Transform. *Chemom. Intell. Lab. Syst.* **1997**, *36*(2), 81–94.
21. Chau, F.T.; Shih, T.M.; Gao, J.B.; Chan, C.K. Application of the Fast Wavelet Transform Method to Compress Ultraviolet-Visible Spectra. *Appl. Spectrosc.* **1996**, *50*(3), 339–348.
22. Chau, F.T.; Gao, J.B.; Shih, T.M.; Wang, J. Compression of Infrared Spectral Data Using the Fast Wavelet Transform Method. *Appl. Spectrosc.* **1997**, *51*(5), 649–659.
23. Shao, X.G.; Sun, P.Y.; Cai, W.S.; Zhang, M.S. Resolution of Overlapping Chromatograms by Wavelet Transform. *Chin. J. Anal. Chem.* **1997**, *25*(6), 671–674.
24. Shao, X.G.; Cai, W.S.; Sun, P.Y.; Zhang, M.S.; Zhao, G.W. Quantitative Determination of the Components in Overlapping Chromatographic Peaks Using Wavelet Transform. *Anal. Chem.* **1997**, *69*(9), 1722–1725.
25. Shao, X.G.; Sun, P.Y.; Cai, W.S.; Zhang, M.S. Wavelet Analysis and Its Application to Resolution of Overlapping Chromatograms. *Chemistry* **1997**, *8*, 59–62.
26. Serrai, H.; Senhadji, L.; Decertaines, J.D.; Coatrieux, J.L. Time-Domain Quantification of Amplitude, Chemical Shift, Apparent Relaxation Time T^* , and Phase by Wavelet-Transform Analysis. Application to Biomedical Magnetic Resonance Spectroscopy. *J. Magn. Reson.* **1997**, *124*(1), 20–34.
27. Barache, D.; Antonie, J.P.; Dereppe, J.M. The Continuous Wavelet Transform, an Analysis Tool for NMR Spectroscopy. *J. Magn. Reson.* **1997**, *128*(1), 1–11.
28. Walczak, B.; Bouveresse, E.; Massart, D.L. Standardization of Near-infrared Spectra in the Wavelet Domain. *Chemom. Intell. Lab. Syst.* **1997**, *36*, 41–51.



WAVELET TRANSFORM

449

29. Fang, H.; Chen, H.Y. Wavelet Analyses of Electroanalytical Chemistry Responses and an Adaptive Wavelet Filter. *Anal. Chim. Acta* **1997**, *346*, 319–325.
30. Shao, X.G.; Zhong, H.B.; Zhang, M.S.; Zhao, G.W. Application of Wavelet Transform in the Analysis of EXAFS Data. *Spectrosc. Spectr. Anal.* **1997**, *17*(6), 85–88.
31. Pan, Z.X.; Shao, X.G.; Zhong, H.B.; Liu, W.; Wang, H.; Zhang, M.S. Correction of Baseline Drift in High-Performance Liquid Chromatography by Wavelet Transform. *Chin. J. Anal. Chem.* **1996**, *24*(2), 149–153.
32. Shao, X.G.; Hou, S.Q.; Fang, N.H.; He, Y.Z.; Zhao, G.W. Quantitative Determination of Plant Hormones by High Performance Liquid Chromatography with Wavelet Transform. *Chin. J. Anal. Chem.* **1998**, *26*(1), 107–110.
33. Shao, L.M.; Tang, B.; Shao, X.G.; Zhao, G.W.; Liu, S.T. Wavelet Transform Treatment of Noise in High Performance Liquid Chromatography. *Chin. J. Anal. Chem.* **1997**, *25*(1), 15–18.
34. Guler, I.; Hardalac, F.; Muldur, S. Determination of Aorta Failure with the Application of FFT, AR and Wavelet Methods to Doppler Technique. *Comput. Bio. Medic.* **2001**, *31*(4), 229–238.
35. Collantes, E.R.; Duta, R.; Welsh, W.J.; Zielinski, W.L.; Brower, J. Preprocessing of HPLC Trace Impurity Patterns by Wavelet Packets for Pharmaceutical Fingerprinting Using Artificial Neural Networks. *Anal. Chem.* **1997**, *69*(7), 1392–1397.
36. Shen, H.L.; Wang, J.H.; Liang, Y.Z.; Pettersson, K.; Josefson, M.; Gottfries, J.; Lee, F. Chemical Rank Estimation by Multiresolution Analysis for Two-way Data in the Presence of Background. *Chemom. Intell. Lab. Syst.* **1997**, *37*(2), 261–269.
37. Shao, X.G.; Shao, L.M.; Hou, S.Q.; Zhao, G.W. Application of Data Compression Based on Wavelet Transformation in Window Factor Analysis. *J. Chin. Univ. Sci. Tech.* **1997**, *27*(3), 460–465.
38. Shao, X.G.; Cai, W.S. Resolution of Multicomponent Chromatograms by Window Factor Analysis with Wavelet Transform Preprocessing. *J. Chemom.* **1998**, *12*(2), 85–93.
39. Jouan-Rimbaud, D.; Walczak, B.; Poppi, R.J.; Noord, O.E.; Massart, D.L. Application of Wavelet Transform To Extract the Relevant Component from Spectral Data for Multivariate Calibration. *Anal. Chem.* **1997**, *69*(21), 4317–4323.
40. Shao, X.G.; Li, W.; Chen, G.; Su, Q.D. On-line Analysis of the Photoacoustic Spectral Signal Using Wavelet Transform. *Fresenius J. Anal. Chem.* **1999**, *363*(3), 215–218.



41. Cai, W.S.; Wang, L.Y.; Pan, Z.X.; Zuo, J.; Xu, C.; Shao, X.G.J. Application of Wavelet Transform in Quantitative Analysis of Raman Spectrum. *Raman Spectrosc.* **2001**, *32*(3), 207–209.
42. Shao, X.G.; Shao, L.M.; Zhao, G.W. Extraction of Extended X-ray Absorption Fine Structure Information from the Experimental Data Using the Wavelet Transform. *Anal. Commun.* **1998**, *35*(4), 135–137.
43. Mustre, J.; Yacoby, Y.; Stern, E.A.; Rehr, J.J. Analysis Of Experimental Extended X-Ray-Absorption Fine-Structure (Exafs) Data Using Calculated Curved-Wave, Multiple-Scattering Exafs Spectra. *Phys. Rev. B* **1990**, *42*(17), 10843–10851.
44. Lawler, R.G.; Barbara, P.F. Practical Aspects of Quantitative CIDNP Using FT NMR. *J. Magn. Reson.* **1980**, *40*, 135–155.
45. Wittbold, W.M.; Fischman, A.J.; Ogle, C.; Cowburn, D. Optimal Filtering in Fourier Transform NMR. *J. Magn. Reson.* **1980**, *39*, 127–135.
46. Pearson, G.A. Optimization of Gaussian Resolution Enhancement. *J. Magn. Reson.* **1987**, *74*, 541–545.
47. Brereton, I.M.; Galloway, G.J.; Field, J.; Marshman, M.F.; Doddrell, D.M. Gradient-induced Water-suppression Techniques for High-resolution NMR Spectroscopy. *J. Magn. Reson.* **1989**, *81*, 411–417.
48. Shao, X.G.; Gu, H.; Wu, J.H.; Shi, Y.Y. Resolution of the NMR Spectrum Using Wavelet Transform. *Appl. Spectrosc.* **2000**, *54*(5), 731–738.
49. Shao, X.G.; Cai, W.S.; Sun, P.Y. Determination of Component Number in the Overlapping Multicomponent Chromatogram Using Wavelet Transform. *Chemom. Intell. Lab. Syst.* **1998**, *43*, 147–155.
50. Leung, A.K.; Chau, F.T.; Gao, J.B. Wavelet Transform: A Method for Derivative Calculation in Analytical Chemistry. *Anal. Chem.* **1998**, *70*, 5222–5229.
51. Shao, X.G.; Pang, C.Y.; Su, Q.D. A Novel Method to Calculate the Approximate Derivative Photoacoustic Spectrum Using Continuous Wavelet Transform. *Fresenius J. Anal. Chem.* **2000**, *367*(6), 525–569.
52. Ho, H.L.; Cham, W.K.; Chau, F.T.; Wu, J.Y. Application of Biorthogonal Wavelet Transform to the Compression of Ultraviolet-Visible Spectra. *Comput. Chem.* **1999**, *23*(1), 85–96.
53. Leung, A.K.M.; Chau, F.T.; Gao, J.B.; Shih, T.M. Application of Wavelet Transform in Infrared Spectrometry: Spectral Compression and Library Search. *Chemom. Intell. Lab. Syst.* **1998**, *43*, 69–88.