Spectral Deconvolution and Feature Extraction With Robust Adaptive Tikhonov Regularization

Hai Liu, *Student Member, IEEE*, Luxin Yan, *Member, IEEE*, Yi Chang,
Houzhang Fang, *Student Member, IEEE*, and Tianxu Zhang

Abstract—Raman spectral interpretation often suffers common problems of band overlapping and random noise. Spectral deconvolution and feature-parameter extraction are both classical problems, which are known to be difficult and have attracted major research efforts. This paper shows that the two problems are tightly coupled and can be successfully solved together. Mutual support of Raman spectral deconvolution and feature-extraction processes within a joint variational framework are theoretically motivated and validated by successful experimental results. The main idea is to recover latent spectrum and extract spectral feature parameters from slit-distorted Raman spectrum simultaneously. Moreover, a robust adaptive Tikhonov regularization function is suggested to distinguish the flat, noise, and points, which can suppress noise effectively as well as preserve details. To evaluate the performance of the proposed method, quantitative and qualitative analyses were carried out by visual inspection and quality indexes of the simulated and real Raman spectra.

Index Terms—Deconvolution, feature extraction, peak detection, Raman spectroscopy, spectral analysis, Tikhonov regularization (TR), variational method.

I. Introduction

THE observed Raman spectra often suffer strong band overlapping and random noise. Raman spectra are often acquired aiming to characterize protein structure and structural changes of proteins upon ligand binding, as well as for many other chemical and biochemical studies. The determination of band parameters is often necessary for the interpretation of spectra. The analysis of Raman spectra can characterize the unknown chemical mixture, such as red wine [1], coal [2], amorphous carbon [3], zircon [4], and so on. Generally, spectral data points can be classified into three different types. Flat point is the point with no absorption at the wavenumber, and it has the same intensity with surrounding points. Steep point has a strong absorbance band and has a steep variation with neighborhood points. Noise point is the point at the local region, whose variation is middle between the flat point and steep point. Three kinds of points are shown in Fig. 1.
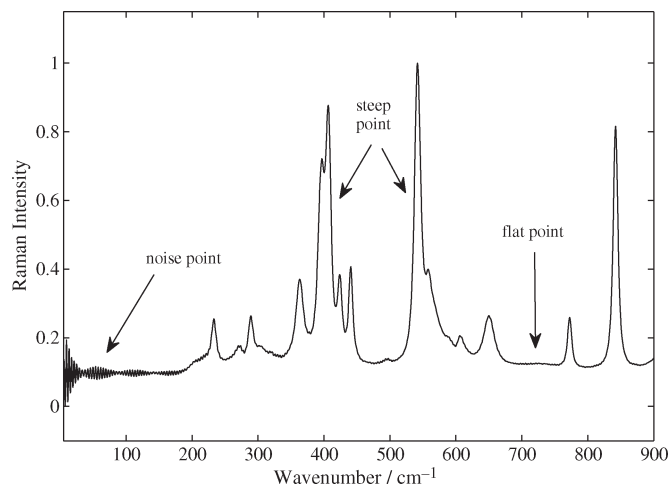
Fig. 1. Raman spectroscopic data series of (D+)-glucopyranose. Spectral points can be classified into three kinds, flat, noise, and steep points.

The first step of the analysis is challenged by the habitually strong band overlapping in Raman spectra. The causes of the band overlapping include natural effects, Doppler, collision broadening, and so on. However, Raman spectral resolving power can be significantly improved by mathematically removing the effect of the instrument response function and suppressing the noise. Curve fitting and spectral deconvolution are both classical methods, which are known to be effective and have attracted major research efforts; see, e.g., [3]–[14]. The former allows the estimation of their band parameters, such as positions, widths, and area, while the latter potentially provides more resolved spectra by narrowing the component bands.

Curve fitting is a special case of data modeling and parameter estimation. The experimental spectral data are modeled by a sum of bands to obtain a result as an estimation of band parameter. The simplest curve-fitting technique is to estimate overlapping bands by the least squares optimization procedure [5]. In the following, many other effective methods are also proposed, such as Voigt profile fitting [6], polynomial fitting [7], frequencies identified by the second-derivative spectrum [8], and parametric band fitting (Gaussian, Lorentzian, or Voigt band shape) [9]. Lórenz-Fonfría and Padrós [10] analyzed the robustness of the curve-fitting method and built a more accurate model to process noise errors. The curve-fitting method is only effective in the case of slight degradation, but for the highly overlapped spectra, it cannot fit the peaks enough, particularly at a low signal-to-noise ratio (SNR).

Deconvolution has become one of the most useful methods for resolving band-overlapping problems. The simplest

deconvolution technique is to process spectral data with the Fourier self-deconvolution (FSD) method [11], [12]. Crilly [15] discussed several iterative deconvolution algorithms and concluded that Jansson's method [16] generally gave the best results. To overcome the limitation of the conventional filter-based methods, blind deconvolution methods are proposed without any prior on filter, such as homomorphic filtering [17], homomorphic deconvolution, Van Cittert iterative [18], and high-order statistics [19], [20]. However, the results seemed rather noisy. Recently, Presser has deconvoluted the Raman spectrum of zircon with a weighted least squares method [4]. Lórenz-Fonfría and Padrós proposed an effective maximum entropy deconvolution method which allowed negative values in the entropy expression [13]. Then, they successfully estimated the composed bands of the highly overlapped bands, utilizing only Fourier transform module and first-order derivative of the spectrum [14].

The results of both methods mentioned earlier are used for interpreting the overlapped Raman spectrum. In this paper, we developed a technique to recover latent spectra and extract feature parameters from degraded Raman spectra, simultaneously. Mutual support of Raman spectral deconvolution and feature-extraction processes within a joint variational framework is theoretically motivated. On one hand, had the correct feature curve of the Raman spectrum been extracted, spectral deconvolution would have been facilitated. The reason is that coupled correct feature curve with the constraint term can distinguish the three kinds of points and impose corresponding smooth constraints on them. On the other hand, had adequate spectral deconvolution been accomplished, successful feature extraction from feature curve would have been much easier to achieve, and accurate feature parameters could have been extracted from the feature curve.

Simultaneous feature-parameter extraction and restoration of an overlapped and noisy spectrum have been presented in [21] and [22]. In those works, the authors fitted a parameterized model based on least squares criterion and achieved impressive results. However, only the peak positions and magnitudes are estimated.

The remainder of this paper is organized as follows. In Section II, the Raman spectrum observation model and the regularization-based deconvolution model are formulated. The optimization method and parameter determination are presented in Section III. Experiment results are provided in Section IV, and Section V concludes this paper.

II. PROBLEM FORMULATION

A. Observed Model

Raman spectrum analysis systems usually operate on over-lapped and noisy spectra. The standard model

$$y(v) = f(v) \otimes g(v) + n(v) \qquad (1)$$

is applicable to a large variety of spectrum degradation processes encountered in practice. Thereinto, $\otimes$ denotes the convolution operation $f(v) \otimes g(v) = \int_L f(v-z)g(z)dz$, where v is the wavenumber. $y(v)$ is the observed spectroscopic data, $f(v)$ stands for actual spectrum, and $g(v)$ stands for the instrument response function (also called blur kernel), which collects the

line-shape function and the instrumental broadening. It is usually assumed that $n(v)$ is distributed as additive Gaussian white noise with zero mean and standard deviation σ.

B. Spectral Deconvolution

When the Raman data are slightly degraded (and not noisy), we often expect least squares solutions of (1) to yield adequate results. Let F denote the set of all solutions, which are consistent with the prior information; a least squares solution $\hat{f}$ will satisfy

$$\hat{f} = \arg \min_{f \in F} \left\{ \|f \otimes g - y\|^2 \right\}. \qquad (2)$$

When the Raman data are overlapping and noisy, finding the minimum $\hat{f}$ is an ill-posed inverse problem: Small perturbations in the data may produce unbounded variations in the solution. Regularization method incorporated *a priori* knowledge about the ideal spectrum into smoothing functional to constrain the solution space. To settle the problem, smoothness regularization is often formed by first- [23] or second-order [24] differences. In this paper, we use the first-order difference to describe the spectral smoothness. Using Tikhonov regularization (TR) [23], [25], [26], the smoothness of the solution is imposed, and our regularized solution f will minimize

$$Q(f) = \frac{1}{2} \int (f \otimes g - y)^2 dv + \lambda \int |f'|^2 dv \qquad (3)$$

where $f' = (f_i - f_{i-1})/2$ and measured by the L_2 norm. The coefficient λ is the positive regularization constant, to weight the constraint term in the cost functional.

C. Spectral-Feature-Parameter Extraction

Raman spectral feature usually includes band numbers, positions, widths, area, and so on. These parameters are very important to interpret Raman spectra, but these feature parameters usually need to be measured manually. In this paper, we proposed a simple way to extract the spectral feature parameters automatically. The proposed extraction method is considered under two assumptions. First, the spectra are noiseless (or noise free). Second, bands of Raman spectra are Gaussian or Gaussian-like shape. The synthetic spectrum with its component subbands (three Gaussian bands), which are shown in Fig. 2(a). We construct a nonnegative monotonically decreasing function

$$w(f') = \exp\left(-\left(\frac{f'}{k}\right)^2\right) \qquad (4)$$

to show spectral feature curve. Fig. 2(b) shows the feature curve of the Raman spectrum [in Fig. 2(a)], computed by (4).

Feature parameters are extracted based on spectral feature curve. We define a new concept called steepest distance (SD), which is the horizontal distance between the steepest points. The steepest point has a maximum first-order derivative at two sides of the subband, which is shown in Fig. 2(a). The peak
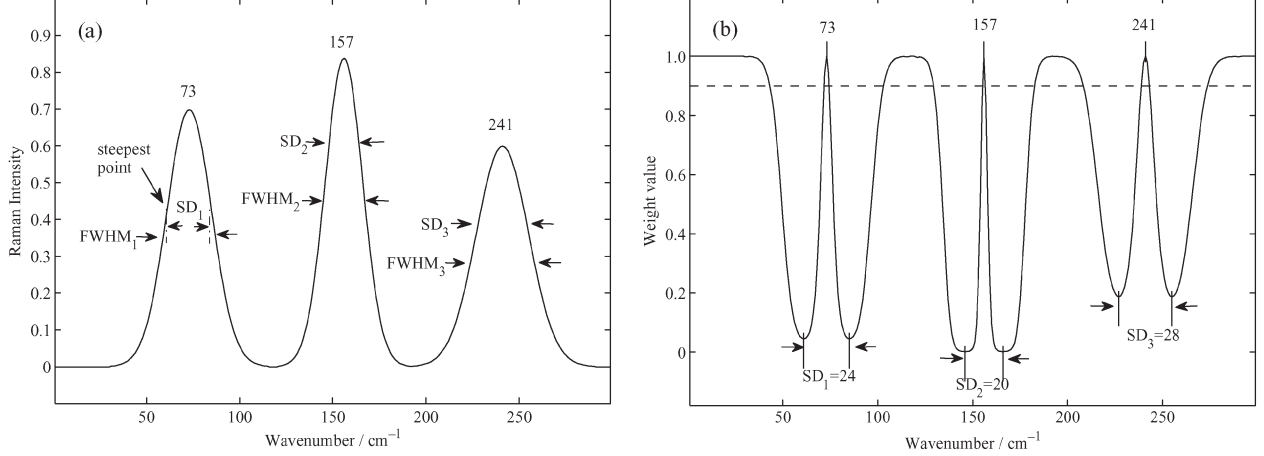
Fig. 2. Bandwidth computation from the simulated bands. (a) Synthetic Raman spectrum with three Gaussian-shape bands. (b) Band-parameter extraction according to the feature curve, using (4). SD (steepest distance).

TABLE I
BAND PARAMETERS OBTAINED FROM THE
DECONVOLUTED SPECTRUM IN FIG. 2

Band position[a]	73	157	241
FWHM	28.3	23.5	33
SD[b]	24	20	28
FWHM$_G$	28.25	23.54	32.97

[a] In cm^{-1}, obtained from the band maximum;
[b] Steepest Distance (SD);

position of the subband equals the wavenumber point, which is the local maximum value point between the two steepest points.

For any-height-and-width Gaussian-shape band, we get the relationship between full-width at half-maximum (FWHM) and SD: FWHM $= \sqrt{2\ln 2} \times$ SD. Its derivation detail is presented in Appendix I. The component bands are Gaussian bands with parameters presented in Table I. We can easily extract the feature parameters (band number, positions, and bandwidth) according to the feature curve.

D. Coupling Spectral Deconvolution With Feature Extraction

Coupling with the feature-extraction function, we proposed a new model

$$Q(f) = \frac{1}{2}\int (f \otimes g - y)^2 dv + \lambda \int \varphi(f') dv \qquad (5)$$

where $\varphi(f') = w(f')|f'|^2$ and $w(f') = \exp(-(f'/k)^2)$; we call $\varphi(f')$ adaptive function, which is shown in Fig. 3(b). The symbol k determines the decreasing speed of the weighting function. The smaller the k is, the faster the weighting function decreases.

In the flat region, f' takes on the small values, leading to $w(f')$ achieving the large values (approximate to one), which will impose a small penalty $\varphi(f')$ on these flat regions [Fig. 3(b)]. In the steep region, f' takes on large values; the weighting function will take on small values (approximate to zero), which thus leads to a small smoothing penalty $\varphi(f')$

on these detail points also, which is shown in the right hand of Fig. 3(b). Thus, $\varphi(f')$ preserves the original information in the flat and steep regions, whereas in the noise region, f' takes on the middle values (between the flat point and the detail point), getting the middle weighting function values, but results in large penalty values on these noise regions. Thus, strong smoothing constraint is imposed on the noise points. It is shown in the middle of Fig. 3(b). Compared with the TR function [shown in Fig. 3(a)], the adaptive TR (ATR) function distinguishes three kinds of points and imposes stronger constraint on the noise point on than other points. Therefore, the weight curve $w(f')$ has two effects in model (5). First, it can compute the feature parameters, such as positions and SD [shown in Fig. 2(b)], and second, it is coupled with the TR to impose different constraints $\varphi(f')$ on the three types of points [shown in Fig. 3(b)].

However, the first-order derivative of the spectrum is very sensitive to the noise. The coefficient value $w(f')$ will become incorrect if the Raman spectrum suffers heavy noise. To settle this problem, we prefiltered the data with a Gaussian function with the whole spectrum to smooth random noise. The cost function can be constructed as

$$Q(f, w) = \frac{1}{2}\int (f \otimes g - y)^2 dv + \lambda \int w(f'_\sigma)|f'|^2 dv. \qquad (6)$$

Thereinto, $f'_\sigma = f' \otimes G_\sigma$, and G_σ denotes the Gaussian kernel with a small standard deviation $\sigma = 2$. It is noteworthy that we only impose the smooth constraint on the weight factor $w(f'_\sigma)$ but not on the spectral signal $|f'|^2$. According to the selections of smoothness constraint function, we call model (5) as ATR and model (6) as robust ATR (RATR).

III. OPTIMIZATION AND PARAMETER DETERMINATION

A. Optimization Method

First, we get the accuracy coefficient from the noisy Raman spectra by

$$w_\sigma = \exp\left(-\left(\frac{f' \otimes G_\sigma}{k}\right)^2\right). \qquad (7)$$
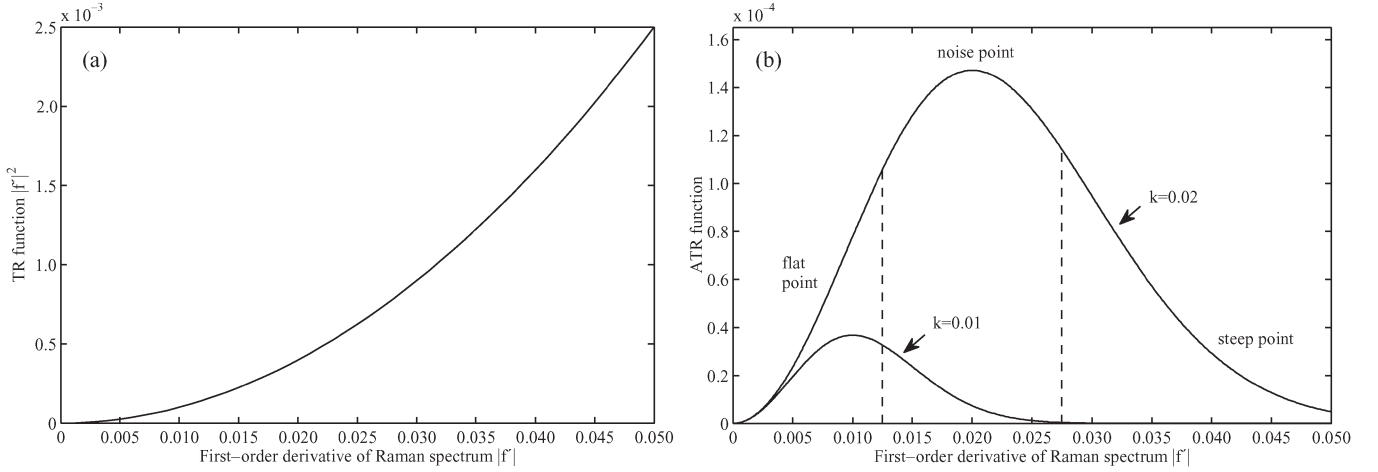
Fig. 3.   (a) TR function in (3). (b) Different constraint $\varphi(f')$ on different points in ATR. TR gives the strong smoothness constraint on the steep points, but the ATR method can distinguish three points.

For simplicity of computation, we introduce a new iteration scheme. The weighting factor w_σ is gotten from the last iteration and considered as a fixed constant. The differentiation of $Q(f)$ with respect to f is given by

$$\frac{\delta Q}{\delta f} = (f \otimes g - y) \otimes g(-v) - 2\lambda w_\sigma f''. \qquad (8)$$

The Euler–Lagrange (E–L) equation is a linear partial differential equation; its derivation is presented in detail in Appendix II. Studying the objective functional (6), it can be seen that it is strictly convex. Thus, the minimization of $Q(f)$ and solution f are both obtained iteratively by a gradient descent method. At the nth iteration, the next solution f^{n+1} is obtained from the current solution f^n as

$$\begin{cases} f^0 = y \\ f^{n+1} = f^n + \Delta t^n \left(-\frac{\delta Q}{\delta f}(f^n) \right) \end{cases} \qquad (9)$$

where Δt^n is the time step. If the Δt^n is too small, the convergence will be very slow. On the other hand, if it is too large, the algorithm will be unstable or divergent. By making a second-order Taylor series approximation to the cost functional at current state f^n, a quadratic step size approximation becomes [27]

$$\Delta t^n = \frac{(\nabla Q(f^n))^T \nabla Q(f^n)}{(\nabla Q(f^n))^T (\nabla^2 Q(f^n)) \nabla Q(f^n)} \qquad (10)$$

where $\nabla Q(f^n)$ is the Hessian matrix of the cost function $Q(f^n)$.

B. Discretization

Upwind finite-difference schemes use a backward or forward finite-difference approximation to the derivative operator [28]. In this paper, the discretization scheme is upwind finite difference. The algorithm calls for the solution of (6). Following

discretization, we obtain a linear system of (7) and (8), which are discrete as

$$w_\sigma = \exp\left(-\left(\frac{\left(\frac{(f_{i+1}-f_i)}{2} \right) \otimes G_\sigma}{k} \right)^2 \right) \qquad (11)$$

$$\sum_{i \in L} ((f \otimes g - y) \otimes g)_i - 2\lambda w_\sigma \sum_{i \in L} (f_{i-1} - 2f_i + f_{i+1}) = 0. \qquad (12)$$

The discrete support of the Gaussian is limited to about $6\sigma + 1$, which, in our experiments, is much smaller than the spectrum length. The convolution operator can be computed by the Matlab function "*imfilter*."

C. Slit Width Selection and Choice of the Parameters

The ideal spectrum usually has the narrowest and sharpest absorption bands. It is important to determine accurate line widths and peak positions from slit-distorted spectra. It is often considered that instrument response function can be approximated as a known function, such as Lorentzian, Gaussian [3], or a combine of both. Reference [29] concludes that a Voigt profile representation for the observed Raman spectra is the most accurate. Voigt shape is defined as

$$g(v) = \frac{1}{\text{HWHM}_D} \sqrt{\frac{\ln 2}{\pi}} \frac{\mu}{\pi} \int_{-\infty}^{+\infty} \frac{\exp(-\theta^2)}{\mu^2 + (\xi - \theta)^2} d\theta \qquad (13)$$

which is the convolution of the intrinsic line-shape function with the instrumental broadening, where

$$\xi = \frac{v}{\text{HWHM}_D} \sqrt{\ln 2} \qquad \mu = \frac{\text{HWHM}_P}{\text{HWHM}_D} \sqrt{\ln 2}.$$

$\text{HWHM}_P = \text{FWHM}/2 = \sqrt{2 \ln 2}\,\sigma$ is the half width of the combination of natural and collisional broadening (Gaussian function), and $HWHM_D$ is the half width of Doppler
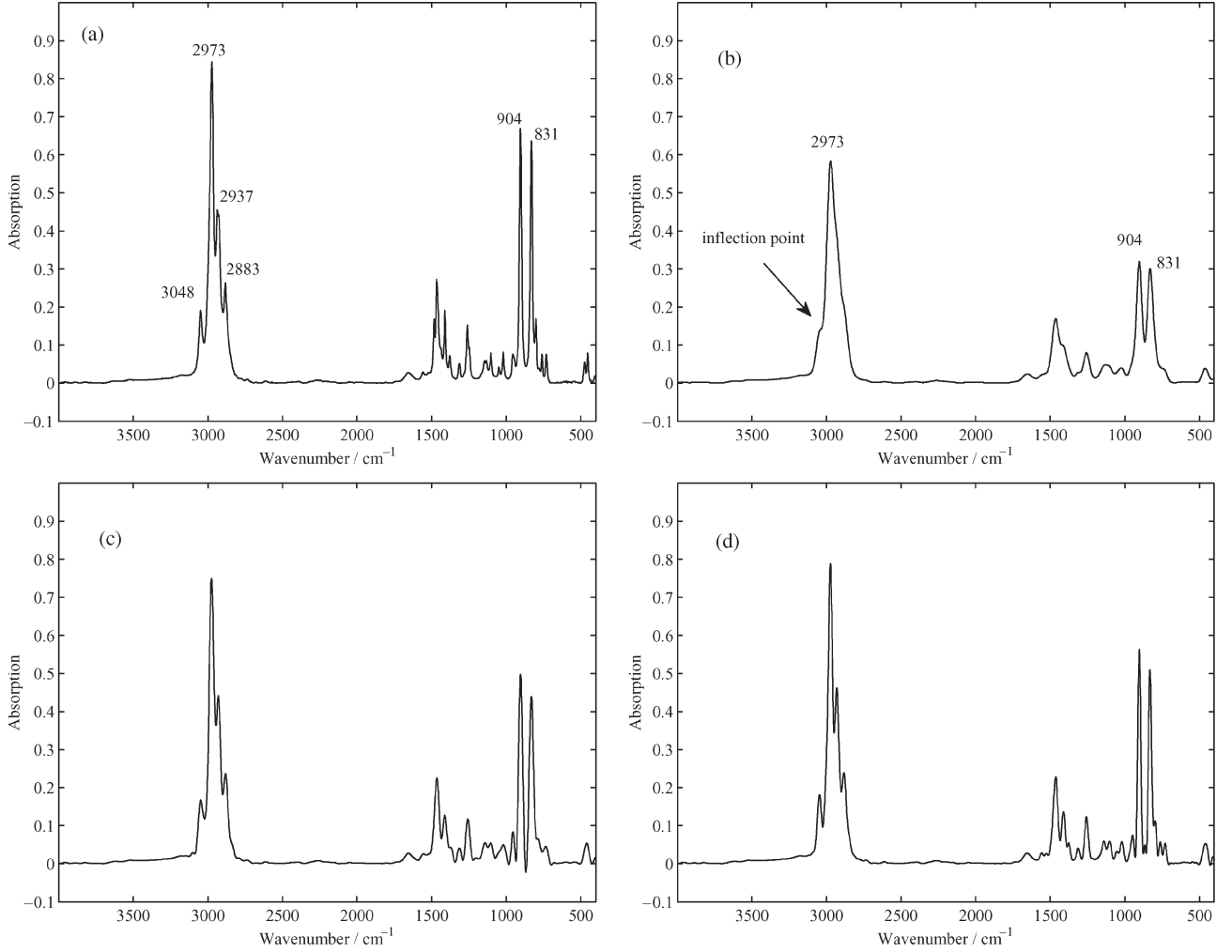
Fig. 4. (a) Original spectrum. (b) Degraded spectrum, convolved with Voigt kernel $\sigma = 16$ cm^{-1}. (c) Deconvolution by the FSD [28] method with a kernel width of 16 cm^{-1} (narrowing factor of 2.0). (d) Deconvolution by the proposed method with the true kernel.

broadening (Lorentzian function). The Lorentzian character is chosen as 0.8 [29].

The choice of an optimal coefficient λ plays a key role in the regularization. On one hand, for a better noise suppressing, a larger λ should be chosen to favor a smoother solution. However, a too large value of λ will induce oversmoothing. On the other hand, if λ is too small, the deconvoluted peaks are resolved enough but contain much residual noise. In other words, the minimization of (6) will equal the least square method (2) when $\lambda \to 0$. Therefore, for the proposed RATR method, the regularization parameter λ should be chosen as a balance between the narrowing capability and noise suppression, particularly at low SNRs.

For the regularization coefficient, we set $\lambda = \lambda/t$; then, after each iteration of optimization, the value of λ is divided by t. We usually set $t = 1.01$ to reduce the influence of the smooth term and increase that of the least square term. The basic idea behind such techniques is not to evaluate the optimal λ itself but to converge to the solution provided by its value. In this paper, we adopt a direct adaptation of the methodology described in [30]

$$\delta = \frac{1.4826}{\sqrt{2}} \text{median} \{|y_i - y_{i-1}|, \qquad i = 2, \ldots, N\}. \quad (14)$$

The same parameters are used in the following deconvoluted experiments $\lambda = 20\delta$. Crilly [31] analyzes the computing error in spectral deconvolution algorithms and concluded that small residual may not imply an accurate solution. We declare convergence when, for more than two consecutive iterations, both the cost function and solution changed less than threshold values: $|Q^{n+1} - Q^n| < d_1$ and $\|f^{n+1} - f^n\|/\|f^n\| < d_2$, where d_1 and d_2 are predetermined coefficients. The deconvolution procedure described in this section can be summarized with the following:

Algorithm 1 Raman spectral deconvolution algorithm

Input Observed Raman spectrum
Initialize $f^0 = y$, $\lambda = 20\delta$, $w_\sigma^0 = 1$
% Iterative minimization
While $\{|Q^{n+1} - Q^n| > d_1$ and $\|f^{n+1} - f^n\|/\|f^n\| > d_2\}$
    Update $\lambda^n = \lambda/2$
    % Compute feature curve to get parameters
    1) Get feature curve $w_\sigma^n = \exp[-((f_\sigma')^n/k)^2]$, where
        $f_\sigma' = f' \otimes G_\sigma$;
    % Raman spectrum recovery

2) Fix $w_\sigma^{n+1} = w_\sigma^n$, Solve $\hat{f}^{n+1} = \arg\min_f Q(f^{n+1}, w_\sigma^n)$ using gradient descent flow method.

end **While**

Output deconvoluted Raman spectrum $\hat{f}$ and feature parameters.

Here, d_1 and d_2 are small positive constants between 10^{-9} and 10^{-7} in this paper.

IV. EXPERIMENT AND DISCUSSION

The proposed algorithm is applied to Raman and IR spectra and compared with Fourier self-deconvolution [11], [32] (Kauppinen *et al.* proposed and improved by Lórenz-Fonfría and Padrós), as implemented in Matlab. The performance of the deconvolution methods on the simulated data is evaluated on the basis of the root-mean-square error (RMSE). RMSE is defined as

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(\hat{f}_i - f_i)^2}{N}} \qquad (15)$$

where N is the spectrum length and subscript i denotes the ith data point. RMSE represents the average difference between the two spectra, and a small RMSE corresponds to a good match.

A. Simulated Experiments

To demonstrate the feasibility of this method, we construct degraded spectrum data on the basis of the standard infra spectra. Fig. 4(a) shows the IR spectrum of 1,2-EPOXYBUTANE ($C_4H_8O_1$) from 4000 to 400 cm^{-1} at a 1-cm^{-1} resolution. We regarded it as a true spectrum and convolved it with a Voigt kernel with width $\sigma = 16$ cm^{-1} (not shown here), to generate the degraded spectrum [Fig. 4(b)]. The degraded spectrum became highly overlapped and less resolved, in which bands became wider and lower, such as 904 and 831 cm^{-1}. In Fig. 4(b), it is difficult to distinguish the peaks at 3048, 2973, 2937, and 2883 cm^{-1}.

For the RATR method, we tested the performance with the adaptive regularization. Using (14), we estimated the $\delta = 0.00081$ and set $\lambda = 0.02$. Because both FSD and the proposed method need a known point spread function, we applied a predefined kernel.

Fig. 4(c) and (d) shows the deconvolution results for the noise-free degraded spectrum in Fig. 4(b). For the FSD and RATR methods, we used the accurate Voigt kernel with the same width as the convolved kernel. The recovery spectrum converged after 500 iterations. In Fig. 4(d), the peaks at 3048, 2973, 2937, and 2883 are split from each other distinctly, while in Fig. 3(c) (FSD method), the overlapped bands are separated slightly. Consequently, for the noise-free case, it seems that the RATR method produces a much narrower spectrum with more details than conventional FSD.

How many subbands can be split from one overlapped band? Why can the peak at 2973 cm^{-1} be split to several peaks but the
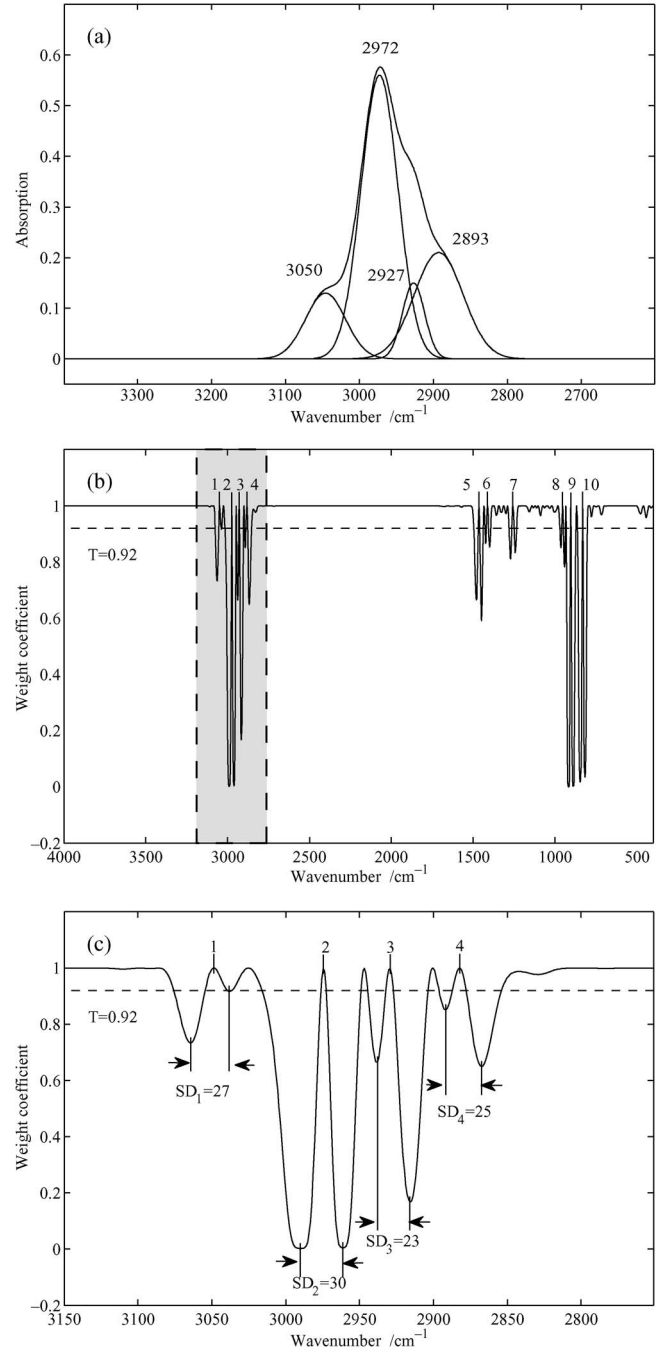

Fig. 5. Spectral feature extraction. (a) Curve-fitting method [9]. (b) Weight coefficient of the whole spectrum. (c) Zooming of local spectrum from 3150 to 2750 cm^{-1}.

peak at 904 cm^{-1} can only recover to one peak? In fact, not all of the peaks can be split into two or more subpeaks. Although the Raman spectrum is degraded, the observation spectrum has much similarity with the original spectrum in "data tendency." The number of peaks is determined by the "data tendency." In Fig. 4(b), there exist the inflection points, which are corresponding to the original Raman spectral valley. The band 2973 cm^{-1} has inflection points, but the band 904 cm^{-1} does not. Therefore, the number of inflection points determines the data tendency and determines how many peaks can be split. With the iterations, some latent inflection points become clear.

TABLE II
PEAK DISTORTIONS IN DECONVOLUTED SPECTRA BY CURVE-FITTING, FSD, AND RATR METHODS [a]CANNOT BE CALCULATED MANUALLY FOR OVERLAPPED BANDS. [b]"+" OR "−" INDICATES VALUE LARGER OR SMALLER THAN THE ORIGINAL, RESPECTIVELY.

Band position		3048	2973	2937	2883	904	831	RMSE
Position	FSD	-1	+1	-6	-1	-1	+1	2.61
	Curve-Fitting	+2	-1	-10	+10	-4	-1	6.08
	RATR	+1	+2	-4	-2	0	0	**2.04**
FWHM	FSD	--a	--	--	--	25.20	25.10	2.42
	Curve-Fitting	27	25	15	33	21	22	4.67
	RATR	31.79	35.32	27.08	29.44	27.00	30.01	**1.55**
Height	FSD	-0.013	-0.036	+0.014	-0.004	-0.115	-0.148	0.0036
	Curve-Fitting	-0.060	-0.210	-0.310	-0.040	-0.350	-0.340	0.2528
	RATR	-0.010	-0.054	+0.017	-0.026	-0.106	-0.120	**0.0025**

B. Feature-Parameter Extraction

The feature parameters are extracted according to the weight coefficient curve. Spectrum characteristics include band numbers, positions, widths, area, and so on. With the iteration, the weight factor becomes more and more accurate, and the deconvoluted spectrum more and more approximates to the true spectrum.

Fig. 5(a) shows the result of the curve-fitting method [9]. Four main peaks are fitted from the heavy overlapped bands [Fig. 4(b)], but the peaks 2973 and 2883 are distorted from the original position 10 cm^{-1}. The result shows that the curve-fitting method can fit the number of the peaks but not peak positions and heights. Fig. 5(b) shows the weight coefficient of the whole spectrum [Fig. 4(d)]. The larger the first-order derivative of Raman spectrum is, the smaller the weighting coefficient achieves. The value of the adaptive weight term represents the importance of the band. We set threshold $T = 0.92$ to distinguish the main features from the residual noise. Ten peaks are extracted, which are labeled with the numbers 1–10. Fig. 5(c) shows the zoomed part from 3150 to 2750 cm^{-1}. The local minimums respect the steepest points of the peaks. We can easily extract the steepest points, whose second-order derivative equals zero.

For the single peaks, such as 904 and 831 cm^{-1}, the parameters automatically extracted by the proposed method can be as accurate as those extracted by the manually marked method, but for the heavy overlapped peaks, such as 3048, 2973, and 2937 cm^{-1}, the conventional method hardly extracts the spectral feature. We mark it with "–" in Table II. However, the proposed method does not suffer such a problem. We may estimate the FWHM according to the SD.

Peak distortions of deconvoluted spectra are also investigated between the FSD and proposed method. We concentrate on how distorted the positions and FWHM of recovered bands compared with those of the original ones are. In the actual spectrum [see Fig. 4(a)], six bands are taken as reference, which are labeled with 3048, 2973, 2937, 2883, 904, and 831 cm^{-1}. Table II lists those band distortions between the deconvoluted [Fig. 4(c) and (d)] and actual spectra [Fig. 4(a)] in their peak position and FWHM. Compared with the FSD method, the proposed method has smaller RMSEs at the indices of position, FWHM,

and height, shown in Table II. Therefore, we may conclude that the RATR method provides much narrower bands and accurate positions than the FSD method for the noise-free spectra.

C. Effect of the SNR

Afterward, to investigate the robustness of deconvolution methods to noise, additive zero-mean random noise $n(v)$ is added to the degraded spectrum [Fig. 4(b)] with four different SNRs of 500, 200, 100, and 50, as shown in Fig. 6(a) (from top to bottom). To improve the visual analysis of deconvolution performances, the results illustrated are extracted from a limited part (1000-cm^{-1} length) of the entire spectrum.

Fig. 6(b) and (c) shows the deconvoluted results from the noisy spectra by FSD and RATR, respectively. In all conditions, the four main overlapped peaks are all deconvoluted. For both methods, with SNR decreasing (from top to down), it can be seen that the recovered spectrum seems to be less deconvoluted and with more residual noise [see Fig. 6(b) and (c)], and at the same SNR level, the residual noise in Fig. 6(c) is less than that in Fig. 6(b), and also, the peaks are separated more distinctly than those in Fig. 6(b). In other words, Fig. 6(c) seems to have less noise left, while the peaks are separated distinctly at the same time. It is helpful to extract the feature parameters. We may conclude that the proposed method performs much better than FSD, which can suppress noise point effectively as well as preserve details at the steep point.

According to (14), we estimated the noise levels and set the regularization parameters. For SNR $= 500$, 200, 100, and 50, we computed the $\delta = 0.0021$, 0.0041, 0.0078, and 0.0132 and set $\lambda_0 = 0.03$, 0.05, 0.06, and 0.08, respectively.

D. Effect of Adaptive Term

To demonstrate the effectiveness of the robust adaptive regularization, we compared the performance of the TR, ATR, and RATR methods. Three models are set with the same parameter settings ($\delta = 0.0041$ and $\lambda_0 = 0.05$) and kernel width $\sigma = 16$ cm^{-1} and thus have the same narrowing ability.

Their deconvolution results are shown in Fig. 7(a) (TR), Fig. 7(b) (ATR), and 7(c) (RATR), deconvoluted from the
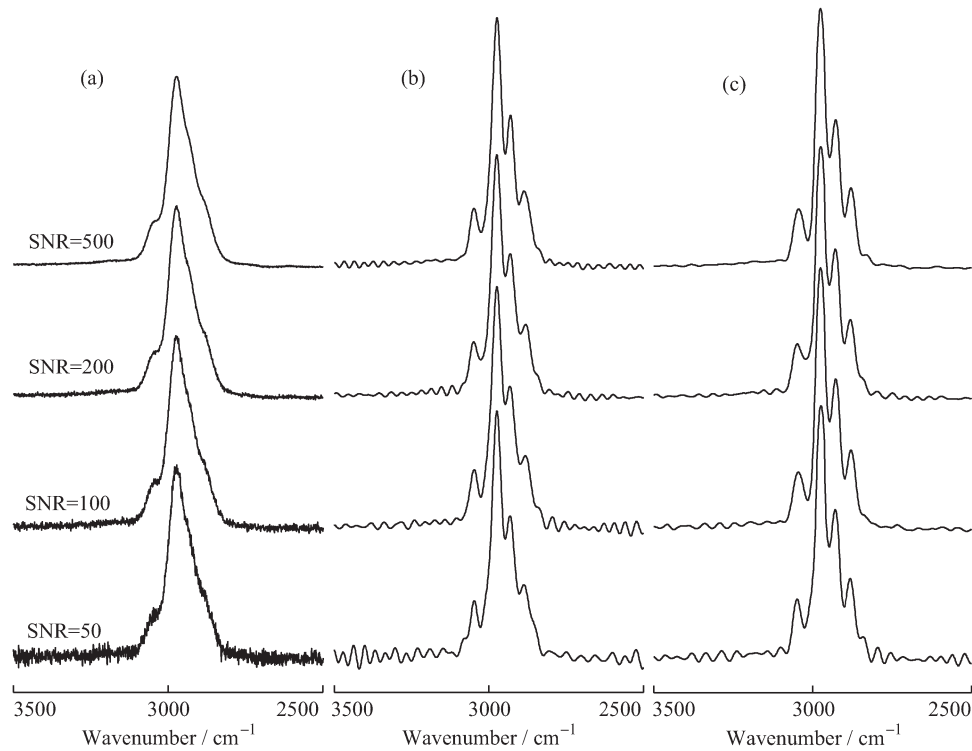
Fig. 6.   Deconvolution results between the proposed method and the FSD method with different noise levels. (a) Simulated noisy absorption spectra at a 1-cm^{-1} resolution (from top to bottom SNRs of 500, 200, 100, and 50). (b) FSD [32] (the width of Voigt kernel $\gamma_v = 16$ cm^{-1} and, from bottom to top, narrowing factors of 1.4, 1.45, 1.5, and 1.6). (c) Spatially adaptive spectral deconvolution with $\gamma_v = 16$ cm^{-1}.
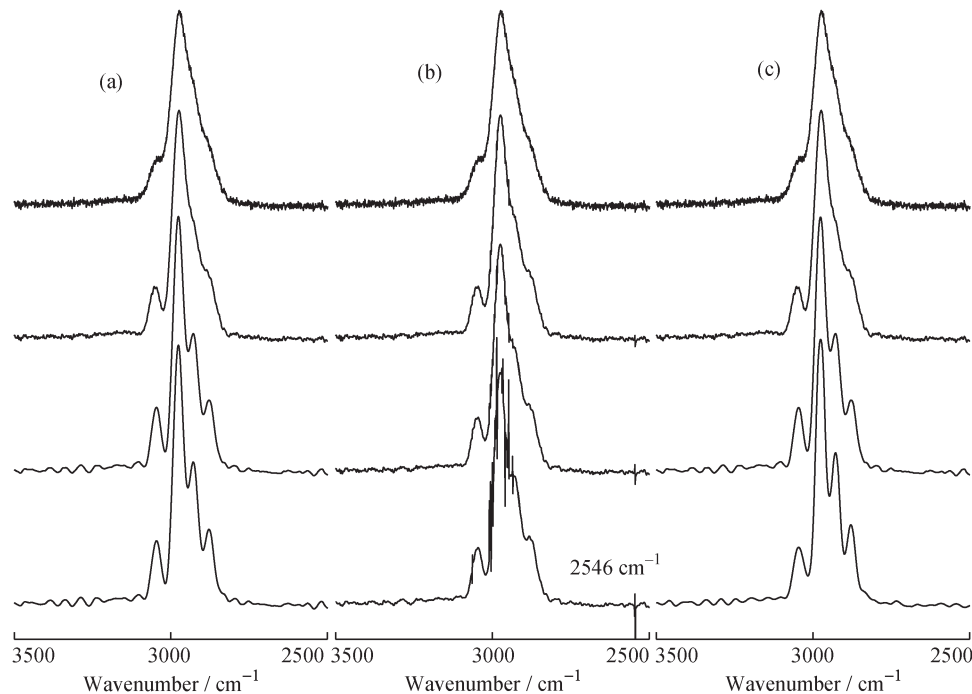

Fig. 7.   Deconvolution results with TR and ATR at SNR = 100. (a) TR (iteration numbers at 1, 10, 100, and 315). (b) ATR (1, 10, 15, and 30). (c) RATR (1, 10, 100, and 456).

overlapped spectrum in Fig. 6(a) (SNR = 100). From Fig. 7(a), we can see the noise suppressed well but the peaks not split enough, particularly at the peaks 2973 and 2937 cm^{-1}. Without the adaptive term, it will lead to oversmoothing solution and loss of sharp band information in the spectra. Fig. 7(b) shows the whole deconvoluted process at the iteration numbers of 1,

10, 15, and 30. The noise point at 2546 cm^{-1} is treated as the steep point to preserve. The coefficient value will become small if the Raman spectrum suffers heavy noise. As a result, significant noise is preserved and propagated through the iteration process, becoming amplified into large artifacts in the solution [bottom of Fig. 7(b)]. Fig. 7(b) shows that the ATR method is
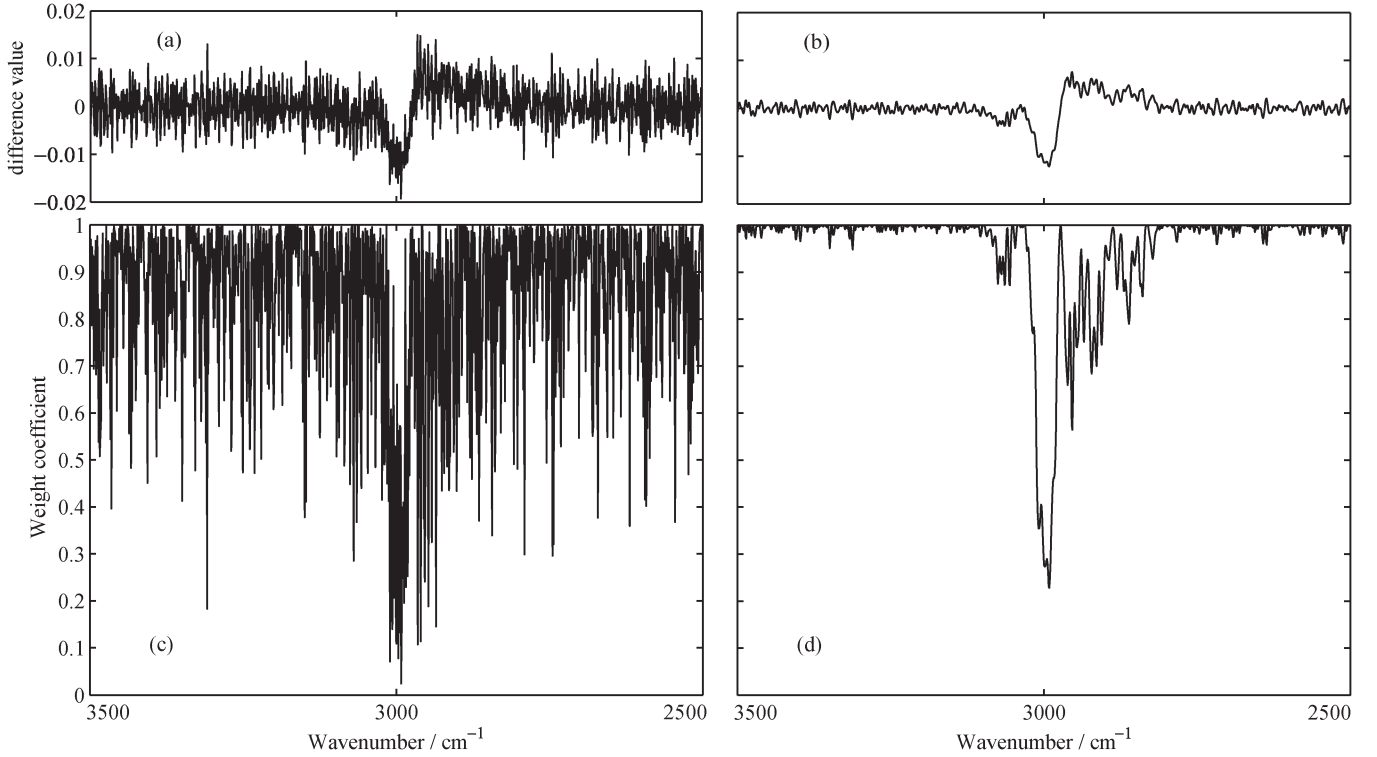
Fig. 8. Weighted values become more accurate with slight Gaussian smoothing. (a) First-order difference of SNR = 100 from 3500 to 2500 cm^{-1} without smoothing. (b) $f'_\sigma = (f \otimes G_\sigma)'$ with width $\sigma = 2$. (c) Weight coefficient of (a). (d) Weight coefficient of (b).

very sensitive to the noise. After iteration of 456, the converged spectrum deconvoluted by RATR is shown at the bottom of Fig. 7(c). The result seems more resolved and less noisy than that by the TR method.

Fig. 8(a) shows the first-order derivative f' of the noisy spectrum [Fig. 6(a), SNR = 100] from 3500 to 2500 cm^{-1}. Weighting curve $w(f')$ [shown in Fig. 8(c)] is computed according to (4). Being built on the spectral data themselves, the weighting curve $w(f')$ seems to be sensitive to noise. It may result in heavy noise points being regarded as the steep points to preserve [shown in Fig. 7(b)]. However, $w(f'_\sigma)$ does not suffer this problem. Fig. 8(b) shows the convolution of G_σ and f'. Fig. 8(d) shows the more accurate weight curve $w(f'_\sigma)$. It shows the first iteration of the algorithm, which converges to Fig. 5(c) at last. We can conclude that RATR is more robust than ATR, particularly at low SNR. It can explain why the weight curve computed by the RATR method is more accurate than that by the ATR method.

Had the correct weight curve [Fig. 8(d)] of the spectrum data been known, spectral deconvolution [Fig. 7(c)] would have been facilitated. Clearly, the overlapped spectra could have been split based on the accurate weight curve. Furthermore, spectral denoising could have been facilitated without oversmoothing peaks. Conversely, had adequate spectral deconvolution been accomplished, successful feature parameters [Fig. 5(c)] would have been much easier to achieve. Spectral feature extraction and deconvolution are, therefore, tightly coupled tasks: The solution of either problem would become fairly straightforward given that of the other.

There are three effects of the smoothness adaptive term. First, the effect of the term is the restoration of the steep points and
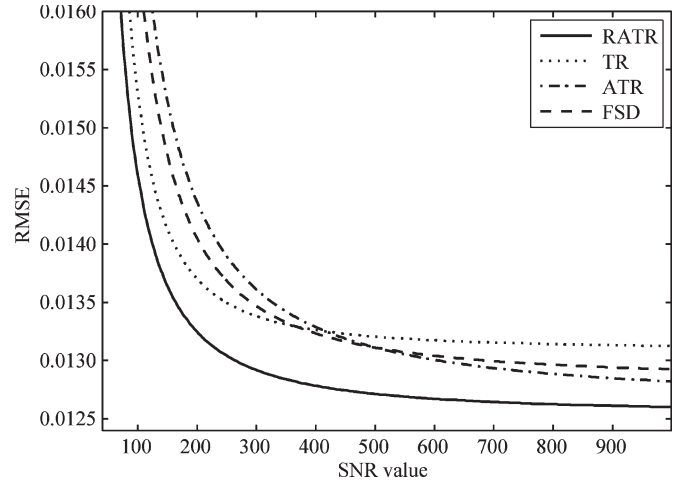

Fig. 9. Comparison of the RMSE values of FSD, TR, ATR, and the proposed method RATR in all SNR conditions. The lower RMSE values imply improved performance.

extraction of the bandwidth of spectral feature. Second, it can distinguish the noise region from flat and steep regions. Third, it helps improve the robustness of the method to noise. Three positive effects highlight the advantages of the proposed RATR method.

E. Relationship of the Models

The proposed method has been validated by direct comparisons with other methods at the merit of RMSE. The RMSE values of the all-level SNR spectra are plotted in Fig. 9.

As the SNR $\rightarrow \infty$, the value of RMSE decreases; that is to say, the recovered spectrum approximates to the original clear
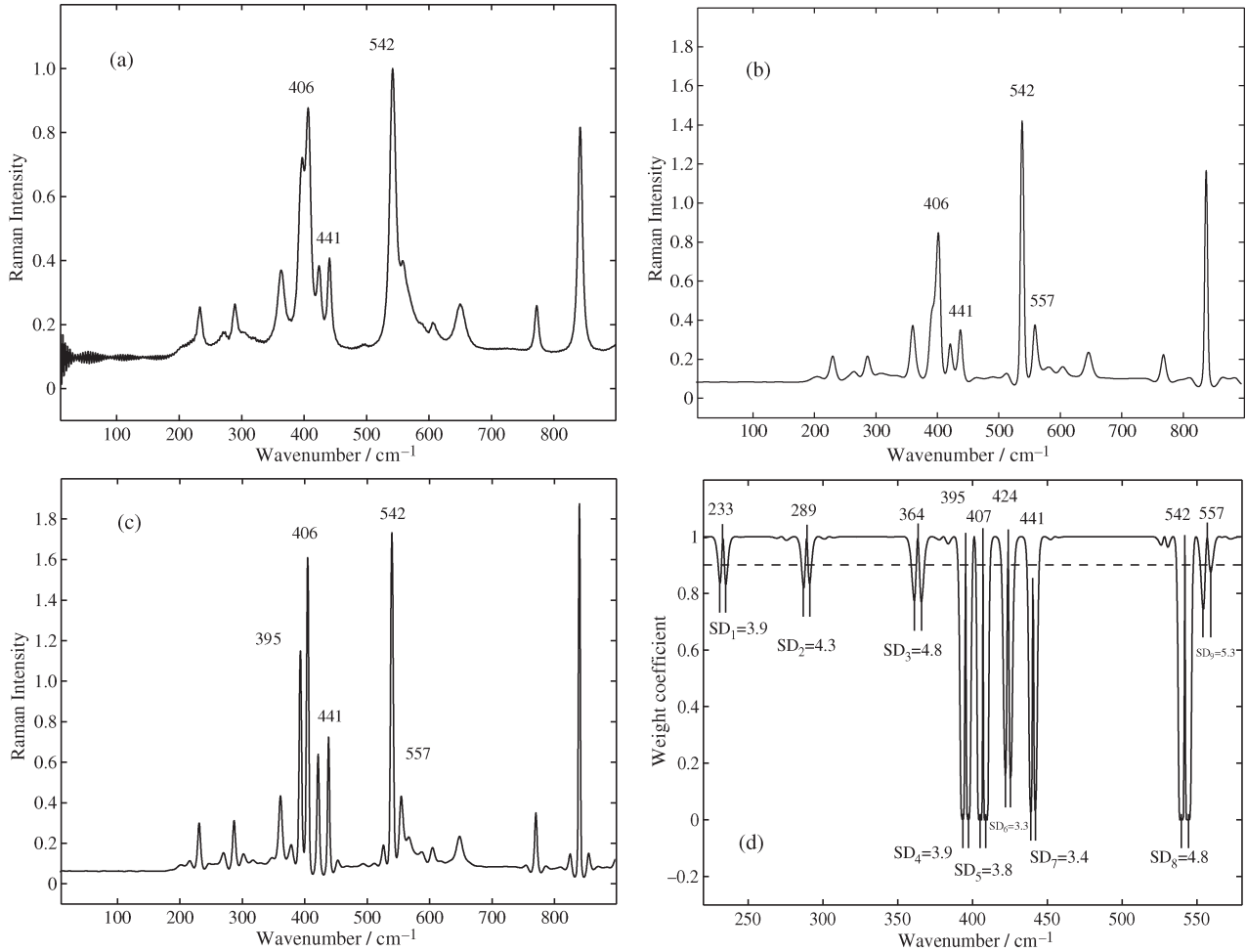
Fig. 10.   (a) One-thousand-six-hundred-long Raman spectral data of (D+)-glucopyranose from 900 to 10 cm^{-1}. (b) Result of TR for kernel width $= 9$ cm^{-1}. (c) Result of RATR method. (d) Feature-parameter extraction according to feature curve (weighting coefficient curve).

spectrum. On the contrary, if the SNR $\to 0$, the noise becomes dominant, resulting in an increase of the RMSE and invalidation of all the methods. It can be seen that the high SNR spectral signal is helpful to recover the spectra. On one hand, with the same SNR, whereas the SNR > 400, the order of compared methods of RMSE value is TR, FSD, ATR, and RATR. On the other hand, with the SNR < 400, that is ATR, FSD, TR, and RATR. Therefore, we can conclude that the proposed method outperforms the ATR, FSD, and TR methods at the abilities of peak narrowing and noise suppression.

Moreover, we also analyze the relationship among the proposed methods. It is easy to see that the model ATR can be regarded as a special case of the model RATR when σ equals zero. In Fig. 9, when noise free, the narrowing ability of RATR will equal that of ATR. When $k \to \infty$, both RATR and ATR reduce to TR. The reason is that the weighting function will become a fixed constant one. Obviously, the behavior of the filter depends on k and the diffusion function, on the diffusion process. In our experiments, we set $k = 0.02$.

F. Application to Experimental Spectra

In this section, we show how the algorithm can be applied to experimental Raman spectra. Fig. 10(a) shows the Raman spec-

trum of (D+)-glucopyranose $(C_6H_{12}O_6)$ from 900 to 10 cm^{-1}. Observed Raman spectrum suffers the heavy noise from 200 to 10 cm^{-1}, and varying degrees overlap at the peaks 406 and 542 cm^{-1}. The result spectrum [Fig. 10(b)] was deconvoluted by TR with a 9-cm^{-1} Voigt kernel (equal to the FWHM at 441 cm^{-1}). The overlapped peak at 542 cm^{-1} can be resolved to 542 and 557 cm^{-1}, but the peak at 406 cm^{-1} cannot. The RATR method splits the 406 cm^{-1} to 395 and 406 cm^{-1}, and the 542 cm^{-1} splits the 406 cm^{-1} to 542 and 557 cm^{-1}. Compared with Fig. 10(b) and (c), it appears more flat from 200 to 10 cm^{-1}. It seems that the deconvolution result in Fig. 10(c) is more resolved than that in Fig. 10(b). The determined band positions are shown in Fig. 10(d), and the extracted feature parameters are shown in Table III.

Fig. 11 shows the deconvoluted results of the Raman spectrum of maltoriose $(C_{18}H_{32}O_{16})$. The FWHM of peak 1460 cm^{-1} is chosen as the width of the Voigt kernel, which equals 15.8 cm^{-1}. In both methods, five peaks are separated distinctly from the original peaks 1125 and 1080 cm^{-1}. It seems that the valleys of the proposed method [shown in Fig. 11(c)] are deeper than those of FSD [shown in Fig. 11(b)]. After 270 iterations, the spectral curve converges, and the weight curve is shown in Fig. 11(d). The spectral feature parameters automatically extracted are shown in Table IV.

TABLE III
PARAMETERS OF THE RECOVERED PEAKS IN FIG. 10

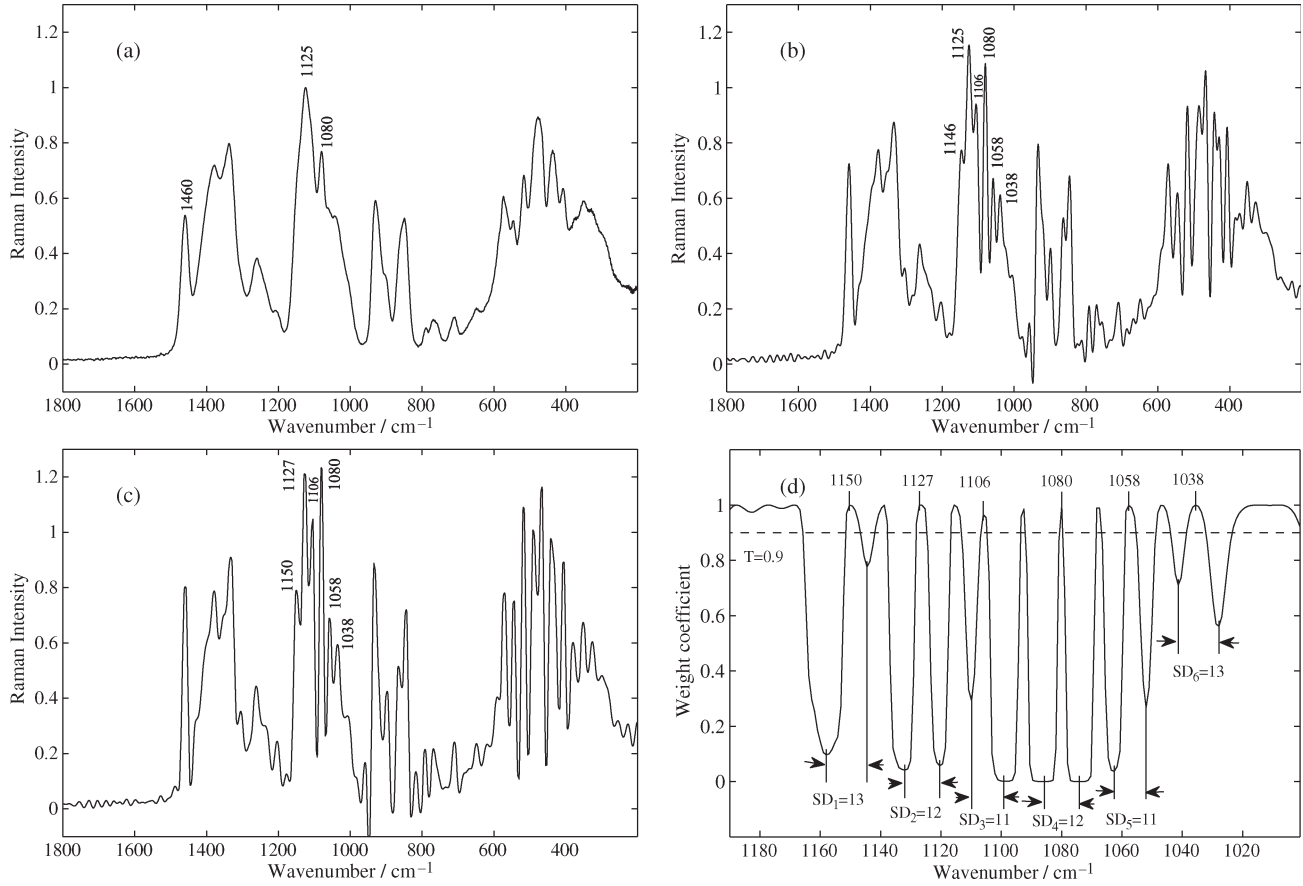| Original Position	233	289	364	406		424	441	542	
Split position	233	289	364	395	407	424	441	542	557
SD	3.9	4.3	4.8	3.9	3.8	3.3	3.4	4.8	5.3
FWHM	4.59	5.06	5.65	4.59	4.47	3.89	4.00	5.65	6.24

Fig. 11. (a) Real raw Raman spectral data of maltoriose from 1800 to 200 cm^{-1}. (b) Deconvolution result by FSD [28]. (c) Deconvolution result by the RATR method. (d) Feature-parameter extraction according to feature curve (weighting coefficient curve).

TABLE IV
PARAMETERS OF THE RECOVERED PEAKS IN FIG. 11

Original Position		1125		1080		--(no peak)
Split position	1150	1127	1106	1080	1058	1038
SD	13	12	11	12	11	13
FWHM	15.31	14.13	12.95	14.13	12.95	15.31

The same spectrum has been deconvoluted with the high-order cumulation method by Yuan *et al.* [19], where the deconvoluted result seems rather infradeconvolution. The computing cost was about 26 s in Matlab 2010a.

V. CONCLUSION

In this paper, we have proposed a new model to recover the latent Raman spectrum and extract the feature parameters with RATR. The proposed method was validated by successful

simulated and experimental results. Three advantages of the proposed method are the following.

1) It can recover the latent Raman spectrum and extract feature parameters according to the weighting curve, such as peak positions, FWHM, and so on.
2) It can distinguish three types of points, flat, noise, and steep points, which are defined in this paper.
3) It is quite robust to random noise.

Moreover, the RMSE and peak distorted merits are employed for objective evaluation. The experiment results show that the proposed method RATR outperforms the traditional deconvolution methods both at the subjective and objective evaluations.

With our analysis and experiments, the structure informations (first order) of the Raman spectra have been fully utilized, and the robust function distinguishes three types of points, which suppress noise effectively as well as preserve details. The weighting curve becomes more and more accurate with the iterations. Then, we can extract the spectral feature according to the weighting curve. Both recovered Raman spectra and the feature parameters are used for interpreting the unknown chemical mixture. We will build a Raman spectral database in our further research.

APPENDIX I

In this Appendix, we give the relationship between the FWHM and SD. Gaussian-shape band can be expressed as

$$G(v) = G_{\max} \exp\left(\frac{-v^2}{G_{\text{width}}}\right)$$

where the symbols $G_{\max}$ and G_{width} denote the height and width of the band, respectively.

Set $G(v) = H/2$. Wet get $v = \sqrt{G_{\text{width}} \ln 2}$ and FWHM $= 2\sqrt{G_{\text{width}} \ln 2}$.

For the SD

$$G'(v) = G_{\max} \exp\left(\frac{-v^2}{G_{\text{width}}}\right)\left(-\frac{2v}{G_{\text{width}}}\right)$$

and $G''(v) = G_{\max} \exp(-(v^2/G_{\text{width}}))((4v^2/G_{\text{width}}) - (2/G_{\text{width}})) = 0$. We get $v = \sqrt{G_{\text{width}}/2}$. SD $= 2\sqrt{G_{\text{width}}/2}$. Therefore, FWHM $= \sqrt{2 \ln 2} \times$ SD.

APPENDIX II

The E–L (8) is derived from the cost functional (6)

$$Q(f) = \frac{1}{2}\int_L (f \otimes g - y)^2 dv + \lambda \int_L w(f'_\sigma)|f'|dv$$

where $w(f'_\sigma) = \exp[-(f'_\sigma/k)^2]$, $G_\sigma = 1/\sqrt{2\pi}\sigma \exp[-v^2/2\sigma^2]$, and

$$f_\sigma(v) = \int f(v-z)G_\sigma(z)dv, \qquad f'_\sigma(v) = f' \otimes G_\sigma$$

$$= \int f'(v-z)G_\sigma(z)dv.$$

In the following derivation, let $f(v)$ be extended to the whole $\mathbf{R}$ with the Neumann boundary condition, meaning that the spectrum intensity values are uniform at the spectral boundary

$$F = \frac{1}{2}(f \otimes g - y)^2 + \lambda w_\sigma |f'|^2.$$

The E–L equation with respect to $\delta Q/\delta f = F_f - (\partial/\partial v)F_{f'} = 0$. For the first part

$$\frac{\partial F}{\partial f} = (f \otimes g - y) \otimes g(-v)$$

and the second part

$$\frac{\partial F}{\partial f'} = 2\lambda w_\sigma f'$$
$$\frac{\partial F_{f'}}{\partial v} = 2\lambda w_\sigma f''.$$

Combined with the first and second parts, we get

$$\frac{\delta F}{\delta f} = (f \otimes g - y) \otimes g(-v) - 2\lambda w_\sigma f''.$$

REFERENCES

[1] Á. L. Gallego, A. R. Guesalaga, E. Bordeu, X. Gonza, and A. S Lez, "Rapid measurement of phenolics compounds in red wine using Raman spectroscopy," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 2, pp. 507–512, Feb. 2011.
[2] S. Potgieter-Vermaak, N. Maledi, N. Wagner, J. H. P. Van Heerden, R. Van Grieken, and J. H. Potgieter, "Raman spectroscopy for the analysis of coal: A review," *J. Raman Spectrosc.*, vol. 42, no. 2, pp. 123–129, Feb. 2011.
[3] Q. Wang., D. D. Allred, and L. V. Knight, "Deconvolution of the Raman spectrum of amorphous carbon," *J. Raman Spectrosc.*, vol. 26, no. 12, pp. 1039–1043, Dec. 1995.
[4] V. Presser, "Metamictization in zircon. Part I: Raman investigation following a Rietveld approach: Profile line deconvolution technique," *J. Raman Spectrosc.*, vol. 40, no. 5, pp. 491–498, May 2009.
[5] R. D. B. Fraser and E. Suzuki, "Resolution of overlapping absorption bands by least squares procedures," *Anal. Chem.*, vol. 38, no. 12, pp. 1770–1773, Nov. 1, 1966.
[6] T. Sundius, "Computer fitting of Voigt profiles to Raman lines," *J. Raman Spectrosc.*, vol. 1, no. 5, pp. 471–488, Nov. 1973.
[7] R. K. Singh, S. N. Singh, B. P. Asthana, and C. M. Pathak, "Deconvolution of Lorentzian Raman linewidth: Techniques of polynomial fitting and extrapolation," *J. Raman Spectrosc.*, vol. 25, no. 6, pp. 423–428, Jun. 1994.
[8] S. A. Tatulian, "Attenuated total reflection Fourier transform infrared spectroscopy: A method of choice for studying membrane proteins and lipids," *Biochemistry*, vol. 42, no. 41, pp. 11 898–11 907, Oct. 1, 2003.
[9] F. Alsmeyer and W. Marquardt, "Automatic generation of peak-shaped models," *Appl. Spectrosc.*, vol. 58, no. 8, pp. 986–994, Aug. 2004.
[10] V. A. Lórenz-Fonfría and E. Padrós, "Curve-fitting overlapped bands: Quantification and improvement of curve-fitting robustness in the presence of errors in the model and in the data," *Analyst*, vol. 129, no. 12, pp. 1243–1250, Dec. 2004.
[11] J. K. Kauppinen, D. J. Moffatt, H. H. Mantsch, and D. G. Cameron, "Fourier transforms in the computation of self-deconvoluted and first-order derivative spectra of overlapped band contours," *Anal. Chem.*, vol. 53, no. 9, pp. 1454–1457, Aug. 1, 1981.
[12] J. K. Kauppinen, D. J. Moffatt, H. H. Mantsch, and D. G. Cameron, "Fourier self-deconvolution: A method for resolving intrinsically overlapped bands," *Appl. Spectrosc.*, vol. 35, no. 3, pp. 271–276, May 1981.

[13] V. A. Lórenz-Fonfría and E. Padrós, "Maximum entropy deconvolution of infrared spectra: Use of a novel entropy expression without sign restriction," *Appl. Spectrosc.*, vol. 59, no. 4, pp. 474–486, Apr. 2005.

[14] V. A. Lórenz-Fonfría and E. Padrós, "Method for the estimation of the mean Lorentzian bandwidth in spectra composed of an unknown number of highly overlapped bands," *Appl. Spectrosc.*, vol. 62, no. 6, pp. 689–700, Jun. 2008.

[15] P. B. Crilly, "A quantitative evaluation of various iterative deconvolution algorithms," *IEEE Trans. Instrum. Meas.*, vol. 40, no. 3, pp. 558–562, Jun. 1991.

[16] P. A. Jansson, *Deconvolution: With Applications in Spectroscopy*. New York: Academic, 1984.

[17] Y. Senga, K. Minami, S. Kawata, and S. Minami, "Estimation of spectral slit width and blind deconvolution of spectroscopic data by homomorphic filtering," *Appl. Opt.*, vol. 23, no. 10, pp. 1601–1608, May 1984.

[18] S. Sarkar, P. K. Dutta, and N. C. Roy, "A blind-deconvolution approach for chromatographic and spectroscopic peak restoration," *IEEE Trans. Instrum. Meas.*, vol. 47, no. 4, pp. 941–947, Aug. 1998.

[19] J. Yuan, Z. Hu, and J. Sun, "High-order cumulant-based blind deconvolution of Raman spectra," *Appl. Opt.*, vol. 44, no. 35, pp. 7595–7601, Dec. 2005.

[20] J. Yuan, "Blind deconvolution of x-ray diffraction profiles by using high-order statistics," *Opt. Eng.*, vol. 48, no. 7, pp. 076501–076505, Jul. 2009.

[21] R. Z. Morawski, A. Miekina, and A. Barwicz, "Combined use of Tikhonov deconvolution and curve fitting for spectrogram interpretation," *Instrum. Sci. Technol.*, vol. 24, no. 3, pp. 155–167, Aug. 1, 1996.

[22] A. Miekina, R. Z. Morawski, and A. Barwicz, "The use of deconvolution and iterative optimization for spectrogram interpretation," *IEEE Trans. Instrum. Meas.*, vol. 46, no. 4, pp. 1049–1053, Aug. 1997.

[23] A. N. Tikhonov and V. I. A. Arsenin, *Solutions of Ill-Posed Problems*. Washington, DC: Wiley, 1977.

[24] S. Twomey, "On the numerical solution of Fredholm integral equations of the first kind by the inversion of the linear system produced by quadrature," *J. ACM*, vol. 10, no. 1, pp. 97–101, Jan. 1963.

[25] J. Ottaway, J. H. Kalivas, and E. Andries, "Spectral multivariate calibration with wavelength selection using variants of Tikhonov regularization," *Appl. Spectrosc.*, vol. 64, no. 12, pp. 1388–1395, Dec. 2010.

[26] L. J. Meng and D. Ramsden, "An inter-comparison of three spectral-deconvolution algorithms for gamma-ray spectroscopy," *IEEE Trans. Nucl. Sci.*, vol. 47, no. 4, pp. 1329–1336, Aug. 2000.

[27] R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE Trans. Image Process.*, vol. 5, no. 6, pp. 996–1011, Jun. 1996.

[28] J. van Trier and W. W. Symes, "Upwind finite-difference calculation of traveltimes," *Geophysics*, vol. 56, no. 6, pp. 812–821, Jun. 1991.

[29] R. K. Singh and A. L. Verma, "Estimation of slit width from observed Raman spectra," *J. Raman Spectrosc.*, vol. 28, no. 5, pp. 301–304, May 1997.

[30] P. L. Davies and A. Kovac, "Local extremes, runs, strings and multiresolution," *Ann. Stat.*, vol. 29, no. 1, pp. 1–48, Feb. 2001.

[31] P. B. Crilly, "Error analysis with deconvolution algorithms," *IEEE Trans. Instrum. Meas.*, vol. 42, no. 1, p. 78, Feb. 1993.

[32] V. A. Lórenz-Fonfría and E. Padrós, "The role and selection of the filter function in Fourier self-deconvolution revisited," *Appl. Spectrosc.*, vol. 63, no. 7, pp. 791–799, Jul. 2009.
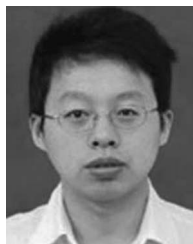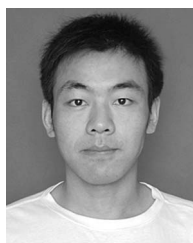
Luxin Yan (M'12) received the B.S. degree in electronic communication engineering and the the Ph.D. degree from Huazhong University of Science and Technology (HUST), Wuhan, China, in 2001 and 2007, respectively.

He is currently an Assistant Professor and Director with the Institute for Pattern Recognition and Artificial Intelligence, HUST. His research interests include multispectral image processing and pattern recognition.
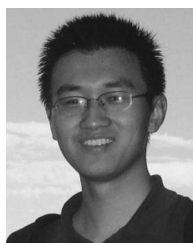
Yi Chang received the B.S. degree in automation from the University of Electronic Science and Technology of China, Chengdu, China, in 2011. He is currently working toward the M.S. degree in pattern recognition and intelligent systems in the Institute for Pattern Recognition and Artificial Intelligence, Huazhong University of Science and Technology, Wuhan, China.

His current research interests include image reconstruction and remote sensing image processing and application.

Houzhang Fang (S'12) received the B.S. degree in applied mathematics from Fuyang Normal College, Fuyang, China, in 2008 and the M.S. degree from the Huazhong University of Science and Technology, Wuhan, China, in 2010, where he is currently working toward the Ph.D. degree in the Institute for Pattern Recognition and Artificial Intelligence.

His research interests lie in image restoration, microwave image processing, and blind deconvolution.

Hai Liu (S'12) received the B.S. degree in computational mathematics in XiangFan University, Xiangyang, China, in 2008 and the M.S. degree in applied mathematics from Huazhong University of Science and Technology, Wuhan, China, in 2010, where he is currently working toward the Ph.D. degree in the Institute for Pattern Recognition and Artificial Intelligence.

His current research interests include spectral analysis, deconvolution theory, and pattern recognition.

Tianxu Zhang received the B.S. degree from the University of Science and Technology of China, Hefei, China, the M.S. degree in computer science from Huazhong University of Science and Technology (HUST), Wuhan, China, in 1983, and the Ph.D. degree in biomedical engineering from Zhejiang University, Hangzhou, China, in 1989.

He is currently a Professor and Director with the Institute for Pattern Recognition and Artificial Intelligence, HUST. His research interests include image processing, computer vision, pattern recognition, and medical imaging.