

Fourier 变换和小波变换处理分析化学信息的比较研究

鲍伦军 莫金垣* 唐祖英

(中山大学化学系 广州 510275)

摘要 小波变换是近年来才引入到分析化学数据处理中的新的数值计算方法,它是由 Fourier 变换发展而来.本文详细比较了两种方法的性质、算法特征以及对分析数据的处理效果. B 样条小波算法简单,无须对数据进行预处理,计算时间短,占用内存少,计算精度高.尤其是对含高噪声分析信号处理优势突出,在信噪比低至 0.2 时仍可获得满意结果.

关键词 Fourier 变换,小波变换,比较研究

当今分析化学家的采集数据的方式多为等时间间距记录下来的数值数据.在采集过程中,由于 A/D 转换的硬件及其他因素不可避免地要混有随机噪声,有时只能采到相当有限的数据点,尤其在低浓度样品分析、信号峰较小时,噪声常常对有用信息的提取造成极大的麻烦.采用化学计量方法进行数据处理不失为行之有效的办法^[1].

化学计量学方法多种多样,各有优劣. Fourier 变换方法是近年来分析化学家们研究较活跃的领域^[2]. 对于分析化学家来说 Fourier 变换技术并不象通讯工程领域的研究人员那样熟练,而且 Fourier 变换及其逆变换是一个比较复杂的数值计算方法. 占用内存大、耗时长. 快速 Fourier 变换虽然在运算速度上得到较大的提高,但是其算法本身的某些特性同样也制约了它的应用. 我们将近年来在数学领域最活跃的研究方法——小波变换引入到化学计量学的数值计算中^[3], 并与传统的 Fourier 变换方法进行了比较研究. 总结概括了 Fourier 变换、加窗 Fourier 变换和小波变换在分析化学信号处理应用中的特征及其应用前景.

1 Fourier 变换及其算法特征

假定我们欲分析处理的信号是时间函数 $f(t)$, 通常其连续 Fourier 变换及其逆变换的公式是:

$$\begin{cases} F(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-i\omega t} dt \\ f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) e^{i\omega t} d\omega \end{cases} \quad (1)$$

对于等时间间距的 N 个离散数据点 f_0, f_1, \dots, f_{N-1} 其离散 Fourier 变换为:

$$\frac{1}{N} \sum_{j=0}^{N-1} f_j \exp\left(-\frac{2i\pi k j}{N}\right) = F_k \quad k = 0, 1, \dots, N-1 \quad (2)$$

逆变换为:

$$\sum_{j=0}^{N-1} F_k \exp\left(\frac{2i\pi kj}{N}\right) = f_j \quad j = 0, 1, \dots, N-1 \quad (3)$$

原始数据 f_j 的虚部为零, 其 Fourier 变换是频率变量 ω 的复函数, 其离散形式有实部和虚部组成 $F_k = R_k + il_k$, 其中:

$$\begin{cases} R_k = \frac{1}{N} \sum_{j=0}^{N-1} f_j \cos\left(\frac{2\pi kj}{N}\right) \\ I_k = \frac{-1}{N} \sum_{j=1}^{N-1} f_j \sin\left(\frac{2\pi kj}{N}\right) \end{cases} \quad (4)$$

时序数据由下式得出:

$$f_j = \sum_{k=0}^{N-1} R_k \cos\left(\frac{2\pi kj}{N}\right) - I_k \sin\frac{2\pi kj}{N} \quad j = 0, 1, \dots, N-1 \quad (5)$$

基于以上算法, 我们可以根据需要对原始数据进行平滑^[4]、插值^[5]、滤波^[6]、拟合^[7]及提高分辨率^[8]等运算. 由(1)式可以看出 Fourier 变换是处理整个时域 $-\infty \sim +\infty$ 范围内的信号, 无法进行时频局部化分析. Gabor 在 1946 年首先提出了窗口 Fourier 变换^[9], 用来对信号进行时频局部化分析. 但是一旦窗口函数 $g(t)$ 确定以后, 其窗口是固定不变的, 无法根据频率的变化进行调节. 小波变换正是在此应用要求下发展起来的^[10].

2 小波变换

小波分析是近期兴起的一个新的数学分支, 它是调和分析、泛函分析、Fourier 分析、样条分析和数值分析的最完美的结晶^[11]. 它是 Fourier 分析发展史上的一次较大突破. 当今许多科学家都在关心并试图涉足这一领域. 其应用在信息处理中的算法简单易行且优势突出的. 我们可以通过十分简便的方法来介绍含噪分析信号的小波变换多分辨分析原理^[12~16].

设 ϕ 为 $L^2 = L^2(R)$ 的一个函数且 $\phi_{kj}(x) = \phi(2^k x - j)$. 对于每一个 $k \in Z$ 以 V_k 表示 $\{\phi_{kj}: j \in Z\}$ 的代数张的 L^2 闭包. 只要满足下列条件就可以说 ϕ 产生了多分辨分析:

- (i) $\dots \subset V_{-1} \subset V_0 \subset V_1$;
 - (ii) 闭包. $(\bigcup_{k \in Z} V_k) = L^2$;
 - (iii) $\bigcap_{k \in Z} V_k = 0$;
 - (iv) $\{\phi_{kj}: j \in Z\}$ 是 V_k 的无条件基
- (6)

假定 ϕ 产生了一个多分辨分析, 且在 L^2 中存在一个函数 Ψ , 如 $\{\Psi(\cdot - j): j \in Z\}$ 的代数张的 L^2 闭包 W_0 是 V_0 在 V_1 的正交补集. 因此 Ψ 相对于 ϕ 来说是基本小波. 由 Ψ 产生的小波空间 W_k 满足下列关系:

- (v) $V_{k+1} = V_k \oplus W_k, k \in Z$
 - (vi) $W_k \perp W_j, k \neq j$
 - (vii) $L^2 = \bigoplus_{k \in Z} W_k$
- (7)

由(vii)可以看出 L^2 中的每一个函数都有一个 $\{W_k, k \in Z\}$ 函数意义上的正交分解. 这也叫做小波分解. 在实际应用中, 应首先以足够大的 k 的 $f_k \in W$ 来近似 $f \in L^2$. 然后得到 f_k 的小波分解:

$$f_k = g_{k-1} + \dots + g_{k-l} + f_{k-l} \quad (8)$$

$g_i \in W_i, i = k-1, \dots, k-l, f_{k-l} \in V_{k-l}$ 所有成分之间均为相互正交的.

我们可以将信号 $f(t)$ 的小波变换定义为:

$$W_{f(b,a)} = |a|^{-1/2} \int_{-\infty}^{+\infty} h\left(\frac{t-b}{a}\right) f(t) dt, a \neq 0 \quad (9)$$

h 为小波母函数. 相应地, 离散小波变换定义为:

$$C_{m,n}(f) = \int_{-\infty}^{+\infty} h_{m,n}(t) f(t) dt \quad (10)$$

其中

$$h_{m,n}(t) = a_0^{-m/2} h(a_0^{-m} t - nb_0), \quad 0 < a_0 < 1, b_0 \neq 0 \quad (11)$$

可以看出小波变换对不同的频率成分(相应于 a_0^{-m})在时域上的取样步长(即 $a_0^m b$)是调节性的. 高频者(相应于小的 m 值)小, 低频者(相应于大的 m 值)大. 它可将信号分解成交织在一起的多种尺度成分, 并对于大小不同的尺度成分采用相应粗细的时域和频域取样步长, 从而能不断地聚焦到对象的任意微小细节. 因此小波变换被誉为数学显微镜^[12]. 小波母函数有多种: 正交小波基^[17,18]、斜交(样条)小波基^[13]、周期小波和多元小波^[19]等. 在此我们重点介绍斜交样条小波的算法及其在分析化学数据处理中的应用.

3 B 样条小波变换快速算法

根据 Mallet 快速算法, B 样条小波的快速算法^[13]可由下式导出:

$$\begin{cases} f_j = \sum_m c_m^j \phi_{j,m}, j = k-l, \dots, k, \\ g_j = \sum_m d_m^j \Psi_{j,m}, j = k-l, \dots, k-1 \end{cases} \quad (12)$$

此处 $f_j = f_{j-1} + g_{j-1}, j = k-l, \dots, k, \phi_{j,m}$ 为尺度函数, $\Psi_{j,m}$ 为样条小波函数, l 为截止频率. 其中:

$$\phi(x) = \sum_{j=0}^m 2^{-m+1} \binom{m}{j} \phi(2x-j) \quad (13)$$

$$\Psi(x) = L_{2m}^{(m)}(2x-1) \quad (14)$$

$$L_{2m}(x) = \sum_{j \in Z} a_j N_{2m}(x+m-j) \quad (15)$$

N_m 为 m 阶 B 样条函数.

$$\begin{cases} c_l^{j-1} = \sum_n a_{n-2l} c_n^j, & j = k-l+1, \dots, k, \\ d_l^{j-1} = \sum_n b_{n-2l} c_n^j, & j = k-l+1, \dots, k. \end{cases} \quad (16)$$

其中 a_n, b_n 可由下式求出:

$$a_n, b_n = O(|r_m|^{-|n|/2}), n \rightarrow \pm \infty \quad (17)$$

且满足下列条件:

$$\phi_m(2t-l) = \sum_n a_{l-2n} \phi(t-n) + \sum_n b_{l-2n} \Psi(t-n) \quad (18)$$

由此我们可以得到同(8)式的多分辨分析. 相当于把 f_k 分解为频率不超于 2^{k-l} 的模糊项 f_{k-l} 与频率介于 2^{k-1} 与 2^{k-l} 之间的带通项 g_{k-1}, \dots, g_{k-l} , 如下列:

$$\begin{array}{ccccccc}
 L^2 \approx V_k & = & W_{k-1} & \oplus & W_{k-2} & \oplus \cdots \oplus & W_{k-l} \oplus V_{k-l} \\
 \downarrow & & \downarrow & & \downarrow & & \downarrow & \downarrow \\
 f_k & = & g_{k-l} & + & g_{k-2} & + \cdots + & g_{k-l} & + f_{k-l} \\
 \downarrow & & \downarrow & & \downarrow & & \downarrow & \downarrow \\
 \omega \leq 2^k & & 2^{k-l} \leq \omega \leq 2^k & & 2^{k-2} \leq \omega \leq 2^{k-l} & & 2^{k-l} \leq \omega \leq 2^{k-l+1} & \omega \leq 2^{k-l}
 \end{array} \quad (19)$$

利用小波变换多分辨分析可将原始信号 $f(t)$ 按照不同频率带通分解, 然后再根据不同的分析目的截取不同段的成分, 进行重构. 让有用信息高度保真地在时域中表示分解和滤除的多频率信息.

4 实验样例及结果讨论

4.1 Fourier 变换的性质及其应用局限

Fourier 变换仅确定了 $f(t)$ 在整个时间域上的频率特征, 构成标架的三角函数永远是振荡的. 因此 $f(t)$ 的信息分布在每个系数 R_k, I_k 中, 只有了解了 $f(t)$ 的全部频域信息, 才能确定 $f(t)$. $f(t)$ 中的奇异点在频域中与所有系数 R_k, I_k 有关, 误差将渗透到每个系数中, 要想完全滤除, 至今仍无较好办法. 如图 1 所示. 我们用计算机模拟产生端点非零的指数函数半微分理论曲线, $\frac{d^{1/2}}{dx^{1/2}} \left(\frac{1}{1+\exp(x)} \right)_{[x=-10 \sim +10]}$, 分别进行 Fourier 变换和逆变换的计算. 可以通过比较图 1, 观察滤波处理后的结果峰(曲线 2)与理论曲线(曲线 1), 可以清楚看出奇异点对 Fourier 变换的影响. 因此在进行变换前, 必须将数据点处理成首尾为零的周期函数, 否则末尾点的突变将波及到整个频率域.

Fourier 变换在滤除噪声的过程中对低频噪声很难处理. 虽然能通过增加采样点数来部分克服之, 但程度非常有限, 而且随着采样点数的增加计算量将呈几何级数地增加; 减低截止频率, 进行“二次滤波”^[4], 可以使部分低频噪声滤除, 但是对原峰形的影响很大, 使峰高明显减低, 由于 Fourier 变换本身的特性, 要求采样点时间函数 $f(t)$ 必须是频带有限函数, 即 $|\omega| > \omega_m$ 时, $f(t) = 0$. 这里 ω_m 是 $f(t)$ 的最高频率成份, 且采样间隔要满足 $\Delta T \leq \frac{1}{2\omega_m}$. 否则会出现频谱叠混^[20,21].

当信噪比较小时, 采样点数就必须越大, 采样点数少, 在进行 Fourier

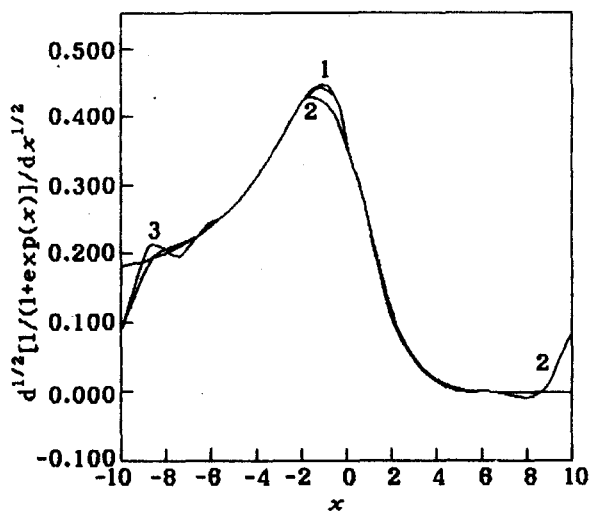


图 1 Fourier 变换和小波变换中奇异点的影响

1—理论曲线; 2—逆 Fourier 变换曲线, 截断系数为 17;

3—小波变换曲线, $l=3$

变换平滑后的结果峰形很不理想, 尤其是首尾部的波动较大. 在进行 Fourier 变换前将数据修

正成首尾为零的周期函数,可使结果得到明显改善,但是当信噪比进一步减小时,实际的端点值与理论曲线的差距越来越大,即使采用边界点修正方法也会引入较大误差.

快速变换运算虽然使运算速度提高较大,但其要求数据点数 N 必须是 2 的幂,不足者需用零值补偿技术补齐,增加了计算量.而且该算法只能 2^n 个输入产生 2^n 个输出,无法根据需要任意调节,在这一点上也同样增加了一些不必要的计算量.经典的 Fourier 变换还有一个固有的缺点,即在时域中没有任何分辨变换, $f(t)$ 在任何有限频段上的信息都不足以确定在任意小范围内的函数.

4.2 样条小波的理论特性

基于样条的多分辨分析及其尺度函数与小波函数,除具有简洁的解析表达式外,还具有线性相位或广义线性相位.对于一个实函数 $f(t) \in L^2(R)$,如果 $F(\omega) = \pm |F(\omega)| e^{i a \omega}$ ($a \in R$) 称作具有线性相位;如果 $F(\omega) = A(\omega) e^{i a \omega + b}$ ($a, b \in R$) 称作广义线性相位.因此尺度函数 ϕ_m 和小波函数 Ψ_m 当 m 为偶数时具有线性相位;当 m 为奇数时具有广义线性相位.可以证明,一个 L^2 函数 $f(t)$ 具有广义线性相位,当且仅当它关于某点 $a \in R$ 是对称的或反对称的.(广义)线性相位在滤波器设计中是经常需要的.当 $\phi(x), \Psi(x)$ 有线性相位时,其共轭 $\overline{\phi(x)}, \overline{\Psi(x)}$ 也有线性相位.因此用 $\{\phi(x), \Psi(x), \overline{\phi(x)}, \overline{\Psi(x)}\}$ 作为信号分解和合成时才不会引起失真.

4.3 两种方法的异同之处

从样条小波原理看,其平滑方法与 Fourier 平滑方法的原理亦相同.若噪声和有用信号的谱重叠严重,或由于样条间隔较大时,也会发生有用信息谱重叠,这种情况下,小波和 Fourier 两种平滑效果均受影响.另外,当设置的频率分界值太大,则有噪声进入;当设置的频率分界值太小,又把有用信号和噪声信号的谱同时去掉,在这两种情况下也均使小波和 Fourier 平滑处理后的信号与真实的有用信号相比,有严重畸形.以上是小波平滑和 Fourier 平滑方法所共有的特点,即原理和处理信号的条件是相同的.两方法之间的差别主要在于求有用信号谱时对处理信号的边界修正和数值计算方法,以及由这些而产生的处理结果的差异.

在数据平滑的计算过程中,如前所述,边界修正问题是 Fourier 平滑能否有效的关键.而小波分析本身具有时频局部化特性,当离散数据起始和终止点不为零时,用小波变换处理这类信号,仅使该点附近有影响,不会传播到其它地方.如图 1 曲线 3 所示对数据直接进行处理效果也相当令人满意,这是 Fourier 方法无法相比的优越之处.选择分界频率 $l=3$ 后,仅进行一次小波样条频率分解,重构信号,即可求出经小波平滑处理后的波形,而且不存在 Fourier 变换的反反正反变换.

在进行数据处理前,我们无需分析数据的准确模型,也不必知道噪声的统计特征.而且,不论该系统是线性的还是非线性的,小波分析都能获得良好结果,还可显示带通部的准确信息.

4.4 两种方法对仿真高噪声信号的处理结果比较

小波分析中,采样间隔及点数对平滑结果的影响并不象 Fourier 方法那么显著,尤其是在处理信噪比较小的数据时,仍可获得较好结果,最低可达 $S/N=0.2$,这是 Fourier 方法所望尘莫及的,而且,通过一新的修正技术,还可改善之.我们选取曲线 $\frac{d}{dx} \left(\frac{1}{1 + \exp(x)} \right) \Big|_{[x=-10 \sim +10]}$ (峰值为 0.25),采样点数为 1024,加上信噪比为 0.2 的白噪声,作为原始数据进行滤波处理.如图 2 所示.清楚显示了不同截止频率($l=5, 4, 3, 2$)时的带通成分及模糊项.

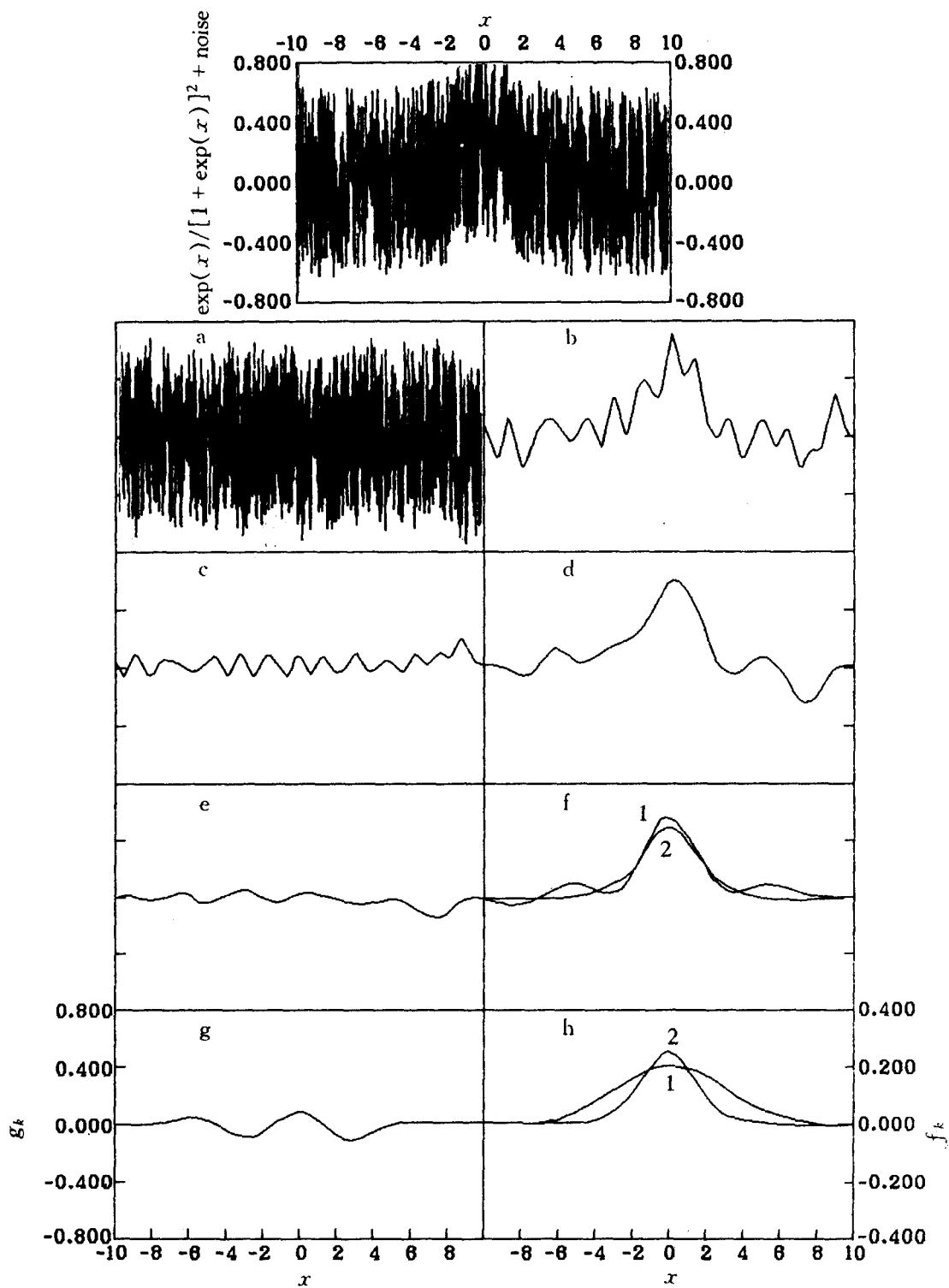


图2 含高噪声信号($S/N=0.2$)小波变换多分辨率分解

x —原始信号; a, c, e, g 和 b, d, f, h 分别为 $l=5, 4, 3, 2$ 时的带通和模糊项, 其中 f, h 的曲线 2 为理论曲线

图3清楚显示了Fourier变换对高噪声信号的处理效果. 理论曲线与图2相同, 加信噪比为0.5的白噪声. 当截断系数为10时, 滤出曲线仍有很大波动(曲线1), 截断系数为6时, 滤出结果严重偏离理论值(曲线2). 可见对于信噪比小于0.5的高噪声信号, Fourier方法确有些无能为力.

4.5 实验数据处理结果比较

图4数据为 $0.1 \text{ mol} \cdot \text{dm}^{-3}$ KCl 底液中还原 $10^{-8} \text{ mol} \cdot \text{dm}^{-3}$ 的 Pb(II) 时的示差脉冲溶出伏安记录, 采样点数为250, $\Delta E = 50 \text{ mV}$, 脉冲宽度为50ms, 脉冲周期为200ms, 分别用Fourier方法进行平滑(截断系数为10)及小波变换进行平滑($l=3$), 可以明显看出, 小波分析与Fourier分析的优劣.

4.6 结论

小波分析处理方法简单、计算时间短、占用内存少、计算精度高、重现性强, 处理结果为连续函数, 其离散值计算的点长可任意设定. 基于小波变换的其它数据处理如插值、实时滤波器的设计以及提高分辨率的计算等, 均与上述算法相似. 详细内容将另文报道. 可以说小波分析处理分析化学数据是具有较高的实用价值的, 值得推广.

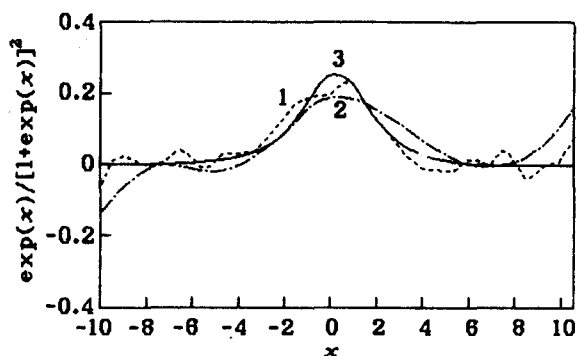


图3 Fourier方法对高噪声($S/N=0.5$)信号的处理效果
1—截断系数=10; 2—截断系数=6; 3—理论曲线

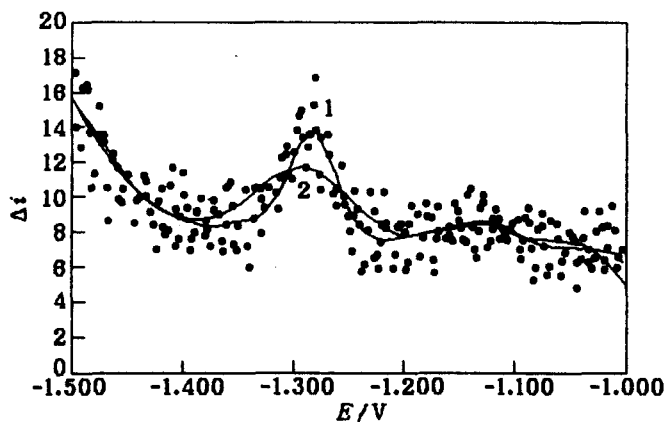


图4 实际样品数据处理结果
1—小波变换平滑结果; 2—Fourier变换平滑结果

参考文献

- 1 俞汝勤,“化学计量学导论”,湖南教育出版社,长沙,1991,第6页.
- 2 S. O. Engblom, *J. Electroanal. Chem.*, **1992**, 332, 73.
- 3 L. Yan, J. Mo, *Chinese Science Bulletin*, **1995**, 60, 1567.
- 4 E. E. Anbanel, J. C. Myland, K. B. Oldham, C. G. Zeski, *J. Electroanal. Chem.*, **1985**, 184, 239.
- 5 R. J. O'Halloran, D. E. Smith, *Anal. Chem.*, **1978**, 50, 1391.
- 6 J. W. Hayes, D. E. Glover, D. E. Smith, *Anal. Chem.*, **1973**, 45, 277.
- 7 J. C. Myland, K. B. Oldham, G. Zhu, *Anal. Chem.*, **1988**, 60, 610.
- 8 S. O. Engblom, *J. Electroanal. Chem.*, **1990**, 296, 371.
- 9 D. Gabor, *J. Inst. Elect. Engin. (London)*, **1946**, 93(II), 429.
- 10 邓东皋, 彭立中, 数学进展, **1991**, 3, 294.
- 11 王建忠, 数学进展, **1992**, 3, 289.
- 12 刘贵忠, 邸双亮, “小波分析及其应用”, 西安电子科技大学出版社, 西安, **1992**, 第88页.
- 13 C. K. Chui, J. Z. Wang, *Proc. of the Amer. Math. Soc.*, **1991**, 113(3), 785.
- 14 C. K. Chui, J. Z. Wang, *Trans. of the Amer. Math. Soc.*, **1992**, 330(2), 903.
- 15 S. G. Mallat, *IEEE Trans. Acoust. Speech Signal Process*, **1989**, 37, 2091.
- 16 S. G. Mallat, *Trans. of the Amer. Math. Soc.*, **1981**, 315, 69.
- 17 I. Daubechies, *Comm. Pure and Appl. Math.*, **1988**, 41, 909.
- 18 I. Daubechies, *IEEE Trans. on Inform. Theory*, **1990**, 36, 961.
- 19 C. de Boor, R. DeVore, *Trans. Amer. Math. Soc.*, **1983**, 276, 775.
- 20 R. N. Bracewell[美]著, 杨燕昌等译, “Fourier 变换及其应用”, 人民邮电出版社, 北京, **1986**, 第174页.
- 21 王正光, 周忠英, 侯伯亨, 李伯成, “数据采集与处理”, 国防工业出版社, 北京, **1985**, 第26页.

Comparative Study on Signal Processing in Analytical Chemistry by Fourier and Wavelet Transforms

BAO Lun-Jun MO Jin-Yuan* TANG Zu-Ying

(Chemistry Department, Zhongshan University, Guangzhou, 510275)

Abstract The Wavelet transform is a new numerical algorithm in processing analytical chemistry data. It is an advance of Fourier transform method. In this paper, we have compared the properties, algorithm and the results of processing data of Fourier and Wavelet transforms. The B-spline Wavelet's algorithm have a lot of advantages such, high-speed, low memory requirement and high-precision and not requiring data preprocessing so on. Especially, it is satisfying to process the high noise data even the signal-to-noise ratio is as low as 0.2.