第 29 卷 分析化学 (FENXI HUAXUE) 研究报告 第 11 期
2001 年 11 月 Chinese Journal of Analytical Chemistry 1246~ 1250

基于分形维数的分析信号自适应中值滤波

程翼宇* 余 杰 钟建毅
(浙江大学化学工程与生物工程学系, 杭州 310027)

摘 要 提出一种面向分析仪器谱图信号处理的分形维自适应中值滤波方法(AMeFFD)。该法延拓运用分形理论, 定义了相对点盒维数概念, 由此建立判定脉冲型噪声的特异性指标, 从而可自动调节中值滤波窗口宽度, 有效地滤除脉冲型噪声及其它类噪声。对仿真色谱信号及实测色谱图的处理结果表明: AMeFFD 法克服了经典中值滤波算法的缺陷, 无论在信号的均方根偏差还是谱峰偏差等指标上, 均明显优于后者, 能在确保谱峰不畸变的同时更有效地滤除脉冲型等常见噪声, 是处理化学谱图信号的有力工具。

关键词 分形理论, 分析仪器, 分析信号处理, 中值滤波, 相对点盒维数

1 引 言

滤除化学谱图中各类噪声是提高分析结果准确性的重要环节。近 30 年发展起来的现代滤波方法已能有效滤除许多类常见噪声, 但仍难以高保真地滤除化学谱图中出现的脉冲型噪声。中值滤波算法[1~2] (median filter, MeF) 尽管能有效滤除脉冲峰宽小于滤波窗口宽度的脉冲噪声, 但仍存在以下缺陷: (1) 滤窗宽度与谱峰信号失真度正相关, 在谱峰保真度和滤噪有效性两者间难以兼顾。(2) 不能根据脉冲噪声的尖窄自适应地调节滤窗宽度, 难以取得最佳滤噪效果。(3) 无法分辨信号峰与脉冲干扰峰。因此, 在滤噪同时易错失具有与脉冲噪声同数量级峰宽和峰高的信号谱峰, 引起严重失真。与中值滤波器相类似, 形态滤波器(morphological filter, MoF) [3] 也无法区分脉冲噪声与信号峰, 故只能对所有谱图数据进行相同的滤波处理, 这会使幅值和峰宽与脉冲噪声同数量级的小峰畸变严重。新近提出的改进型滤噪方法[4] 试图通过在谱峰不同区域人工选择不同结构元素尺寸来克服上述缺陷, 虽能减少信号失真, 但实用性不强。因此, 如何在确保谱峰不畸变的情况下滤除脉冲噪声一直是分析信号处理技术研究中的一个难题。

作者提出一种基于分形维数的自适应中值滤波算法(adaptive median filter based on fractal dimension, AMeFFD), 通过辨识脉冲型噪声, 实现自适应中值滤波。较之经典中值滤波法, 该方法具有更强的除噪能力和很高的谱峰保真度。

2 基本原理

2.1 定义与方法原理

根据分形理论[5], 若将面向连续模型提出的盒维数(box dimension) 作为分形维数, 可定义盒维数如下:

定义 1: 设 $A = \{a_n\}$ 是信号数据序列, 其中 $n \in Z$, $a_n \in Z$; 考虑 Z^2 中 k-正方形网:

$$[m_1 k, (m_1 + 1) k] \times \cdots \times [m_n k, (m_n + 1) k] \quad (1)$$

其中 $k \in N$, $m_1, m_2, \ldots, m_n \in Z$。设 $N_k(A)$ 为 k-正方形网与 A 相交的个数, 取不同的 k, 得到 $(k, N_k(A))$ 对, 作线性回归

$$\log_2 N_k = d \cdot \log_2 k + b \quad (2)$$

令 d 为 A 的整体盒维数。

若计算时仅取信号数据序列中的某个区段, 则所得的盒维数描述了信号的局部光滑性。通过拓展

2000-12-03 收稿; 2001-06-20 接受
本文系国家重点基础研究发展规划(973 计划, No. G1999054405) 及国家自然科学基金(No. 39970890) 资助项目

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

整体盒维数定义, 作者提出点盒维数概念如下:

定义 2: 设 d 为谱图中分析信号曲线上某区段的整体盒维数, 令该数值为此区段中心点 n_c 的点盒维数。为方便计, 各区段中数据点个数取为奇数, 则各点的点盒维数值构成了信号的点盒维数曲线。

由于脉冲型噪声的幅值远大于白噪声, 与小信号谱峰幅值同数量级; 而脉冲峰宽通常只有几个采样点, 明显窄于信号谱峰, 故脉冲型噪声的曲折度较白噪声与信号谱峰大得多。也就是说, 脉冲型噪声数据点的点盒维数值必定显著大于信号谱峰数据点的点盒维数值。由此, 可根据点盒维数值大小对脉冲型噪声作特异性检测, 从而开辟滤除这类噪声的有效途径。

在计算点盒维数时, 我们采用非常规区段(即与脉冲噪声峰宽同数量级的区段) 来分割谱图曲线, 从而使点盒维数值具有仅在脉冲型噪声对应的位置上大于 1 而在其它数据点等于 1 的特性。显然, 利用这一特性易于辨识脉冲型噪声并对其位置进行准确判定。为了区别于常规区段下计算的点盒维数, 本文将非常规区段下计算的点盒维数称为相对点盒维数。

2.2　基于分形理论的自适应中值滤波算法

AMeFFD 算法的基本思想是: 依据上述脉冲型噪声特异性检测原理, 自适应地调节中值滤波窗口宽度, 对含有脉冲型噪声的区域采用较宽的滤波窗口, 增大滤波强度; 而对其它不含脉冲型噪声的区域则调窄滤波窗口, 以避免造成有效信号的失真。本算法的实现步骤如下:

步骤 1. 计算初始谱图曲线的相对点盒维数分布　通常取每个非常规区段为 5 个数据点(x_1, x_2, . . . , x_5), 按前面定义 1 中的公式(1)、(2) 计算各区段的整体盒维数 d, 并根据定义 2 将 d 值作为中心点的相对点盒维数值 d_c, 从而可获得谱图曲线上各区段的相对点盒维数分布。

步骤 2. 辨识脉冲型噪声　设定相对点盒维数阈值。当谱图曲线上数据点的相对点盒维数值大于阈值时, 判断其含有脉冲型噪声, 从而获得脉冲噪声的位置及宽度信息。该阈值在理论上应设为 1, 为确保判断可靠性, 一般可设为 1. 02。

步骤 3. 分段自适应中值滤波　根据步骤 2 的辨识结果, 自动地调节中值滤波器窗口宽度。对含脉冲型噪声的区段, 选用较宽的滤波窗口(取脉冲噪声宽度的 4 倍) ; 而对于其它区段, 则采用较窄的滤窗(宽度取为 5 个数据点) , 从而实现整个谱图曲线的分段滤噪。

3　实验部分

3.1　含噪声色谱图的仿真方法

使用计算机仿真方法模拟含脉冲型噪声和白噪声的色谱谱图, 其中大、小信号峰用高斯函数模拟, 两个脉冲型噪声由尖峰三角函数模拟, 白噪声则由计算机内随机白噪声函数模拟, 仿真谱图见图 1a。上述仿真方法通过调用 Matlab 5. 3 软件包中 signal processing 工具箱的库函数实现。

3.2　含噪声色谱图的实测方法

3.2.1　实验仪器与材料　SP501N 气相色谱仪(鲁南化工仪器厂), 色谱数据工作站(自研)。茅台酒样品: 市售茅台酒。

3.2.2　色谱条件　HP-1 毛细管色谱柱 0. 20 mm × 25 m; 汽化室温度 220 ℃; 柱温采用程序升温, 50 ℃保持 5 min, 然后以 5 ℃/min 升温至 200 ℃; 检测器: FID, 230 ℃; 载气: N_2, 分流比: 50∶1; 燃气: H_2 50 mL/min; 助燃气: 空气 400 mL/min。

3.3　AMeFFD 算法的实现

本算法源程序采用 Matlab 语言编写, 运行于本实验室 Pentium Ⅲ300 兼容机的 Matlab 5. 3 环境下。

4　结果与讨论

4.1　仿真谱图的滤噪结果分析

采用含噪信号相对于理想信号的偏差均方根值 RMS 以及谱峰偏差等指标来比较和评价滤波结果。RMS 的表达式为:

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

$$\mathrm{RMS}=\sqrt{\frac{\sum_{i=1}^{N}(s_a(i)-s_o(i))^2}{N}} \tag{3}$$

其中, $s_o(i)$ 和 $s_a(i)$ 分别为理想信号和含噪信号采样点, N 为信号序列的项数。

谱峰偏差指标包括峰位偏差、峰高偏差和峰面积偏差, 其表达式分别如下:

$$\Delta P=\max\{|P_{a,i}-P_{o,i}|\} \tag{4}$$

$$\Delta H=\max\{\frac{|H_{a,i}-H_{o,i}|}{H_{o,i}}\} \tag{5}$$

$$\Delta A=\max\{\frac{|A_{a,i}-A_{o,i}|}{A_{o,i}}\} \tag{6}$$

其中, $P_{o,i}$、$H_{o,i}$和$A_{o,i}$分别为第 i 个谱峰的真实峰位置、峰高和峰面积, 而 $P_{a,i}$、$H_{a,i}$和$A_{a,i}$则为滤噪后该谱峰的位置、峰高和峰面积。

通过计算可知, 本仿真实验所产生的色谱图在滤波前的 RMS 值为 564。

图 1a 为仿真所得含脉冲型噪声和白噪声的色谱谱图, 其相对点盒维数分布如图 1b 所示, 其中每个分形区段取为 5 个数据点。如图所示, 脉冲型噪声的位置, 对应着图 1b 中相对点盒维数值大于 1 的阶跃峰, 其峰宽较所对应的噪声峰宽略大, 为脉冲型噪声峰宽与分形区段长度之和。因分形区段长度已知, 故脉冲噪声峰宽可准确获得, 这就表明本文方法确能有效地辨识脉冲型噪声的位置及峰宽, 从而自动调节中值滤波器窗宽。图 2a 和图 2b 分别示出了 MeF 法和 AMeFFD 法的滤波结果。用 MeF 法滤波后的 RMS 值随着滤窗宽度 n 的增加而逐渐减小, 在 $n=23$ 时取得最小值为 117, 即此时的滤噪效果最佳; 但与此同时, 谱图中信号小峰平顶化失真严重, 表明 MeF 法难以兼顾谱峰保真度和滤噪有效性, 这就是中值滤波算法无法克服的缺陷。而图 2b 表明, AMeFFD 法的滤波结果相当理想, 可在有效滤除脉冲型噪声和白噪声的同时, 避免平顶化失真。

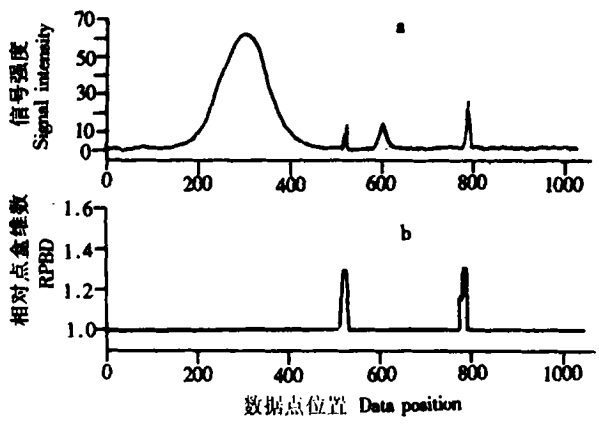


图 1 含脉冲型噪声和白噪声的仿真色谱图(a)及其相对点盒维数分布(b)

Fig. 1 The simulated chromatogram with impulsive noise and white noise (a) and its distribution of relative point box dimension (RPBD) (b)

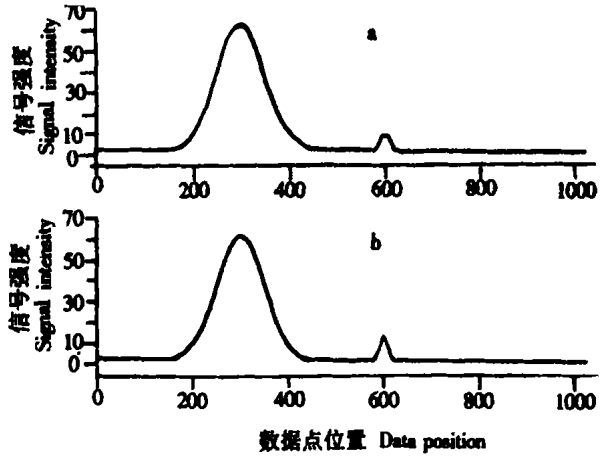


图 2 MeF 法(a)与 AMeFFD 法(b)的仿真滤噪结果对比

Fig. 2 Comparison of simulated denoising results with MeF (a) and AMeFFD (b)

上述两种方法滤波结果的比较与评价见表 1。表中数据显示, 在表征整体滤噪性能和谱峰保真度的各项指标上面, AMeFFD 法都远优于 MeF 法。

4.2 实测茅台酒色谱谱图滤噪结果分析

图 3a 和图 4a 为分析茅台酒样品时获取的含脉冲型噪声的典型色谱图片断, 图 3b 和图 4b 则是分别经 MeF 法和 AMeFFD 法处理后的滤波结果。从图中可以看出, MeF 法存在小信号峰平顶化失真问题, 而 AMeFFD 法则较好地避免了信号失真。经计算可得, MeF 法造成的小峰峰高偏差 ΔH 为 13%, 峰位偏差

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

ΔP 为±3 个点; 而 AMeFFD 法的峰高偏差为 0.4%, 峰位偏差 ΔP 为 1 个采样点, 均显著优于 MeF 法。

表 1　MeF 法和 AMeFFD 法的滤噪仿真实验结果比较

Table 1　Comparison of de-noising simulation experiments of filters with adaptive median filter based on fractal dimension (AMeFFD) and median filter (MeF)

滤波器类型 Filter type	RMS	峰位偏差 Peak position deviation ΔP		峰高偏差 Peak height deviation ΔH(%)		峰面积偏差 Peak area deviation ΔA(%)	
		小峰 Small	大峰 Big	小峰 Small	大峰 Big	小峰 Small	大峰 Big
MeF(n=23)	117.77	±6	±8	33	1	9	0.3
AMeFFD	76.02	±1	±1	0.46	0.3	0.2	0.3

对以上仿真色谱图和实测茅台酒色谱图的滤噪处理结果表明, AMeFFD 法较 MeF 法具有更佳的滤噪效果和更高的保真度。

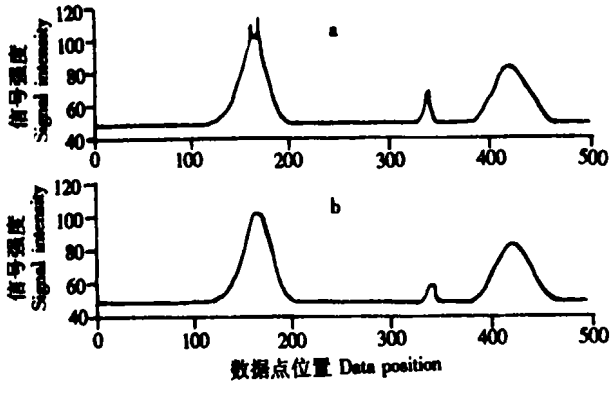


图 3　茅台酒的实测色谱图(a)与 MeF 法的滤噪结果(b)

Fig. 3　The practical chromatogram of Maotaijiu (a) and its denoising result with MeF (b)

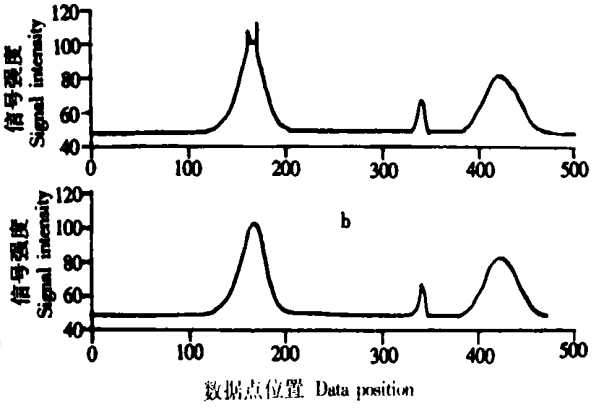


图 4　茅台酒的实测色谱图(a)与 AMeFFD 法的滤噪结果(b)

Fig. 4　The practical chromatogram of Maotaijiu (a) and its denoising result with AMeFFD (b)

5　结　　论

本文拓展了分形理论中的整体盒维数概念, 提出以相对点盒维数值作为辨识化学谱图中脉冲型噪声的特异性指标, 有效地解决了脉冲型噪声的识别和定位难题。在此基础上, 发展了一类面向分析仪器谱图信号处理的自适应中值滤波算法。对仿真色谱数据和实测色谱数据的处理结果表明: 本法既保留了经典中值滤波法的优点, 又克服了其在滤除脉冲型噪声的同时易造成信号小峰失真的缺陷, 具有高信噪比和高保真度的优点, 是化学谱图数据滤波处理的强有力工具。

References

1　Miao Huajian(缪华健), Hu Shangxu(胡上序). *Chem. J. Chinese Universities*(高等学校化学学报), **1995,** 16(7): 1020~1023

2　Gallagher N C. *IEEE Trans. Acoustic, Speech, Signal Processing*, **1981,** 29(6): 1136~1140

3　Wu Yongdong(吴永东), Ma Songde(马颂德). *Chinese J. Computers*(计算机学报), **1997,** 20(11): 989~995

4　Chen Feng(陈　峰), Hu Shangxu(胡上序), Yu Menghuai(俞蒙槐). *Chinese J. Anal. Chem.*(分析化学), **1999,** 27(7): 773~777

5　Li Houqiang(李后强), Wang Fuquan(汪富泉). *Fractal Theory and Its Applications in Molecular Science*(分形理论及其在分子科学中的应用). Beijing(北京): Science Press(科学出版社), **1993:** 1~52

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

An Adaptive Median Filter for Processing Analytical Signal Based on Fractal Dimension

Cheng Yiyu*, Yu Jie, Zhong Jianyi
(*Department of Chemical Engineering & Bioengineering, Zhejiang University, Hangzhou* 310027)

Abstract A novel adaptive median filter, named Adaptive Median Filter based on Fractal Dimension(AMeFFD), is proposed to process analytical signal. A new concept, Relative Point Box Dimension(RPBD), is developed from fractal theory and applied to identify the impulsive noise accurately. According to the results of identification, the window width of median filters (MeF) can be adjusted adaptively. Using the method, the impulsive noise and other common noise can be removed effectively. The filter algorithm has been used to process both simulatied chromatographic signal and practical chromatogram with impulsive noise and other noise. The results showed that, overcoming the drawback of MeF, it was much better than MeF in the RMS index and peak deviation. It has been proved that AMeFFD has the advantage of filtering noise more effectively without distortion of signal peaks and is a powerful tool for processing spectral signal.

Keywords Fractal theory, analytical instrument, analytical signal processing, median filter, relative point box dimension

(Received 3 December 2000; accepted 20 June 2001)

第八届全国电分析化学学术会议

(第Ⅰ 轮)

受中国化学会、中国仪器仪表这会的委托,由西北师范大学承办的第八届全国电分析化学学术年会将于 2002 年 8 月~9 月(具体时间待定)在兰州召开。会议的主题为:讨论新世纪里电分析化学的最新发展前景。主要内容为:(1)伏安法、极谱与极谱催化波;(2)电化学传感器及电化学免疫分析;(3)化学修饰电极与微电极;(4)电位法及其它电化学分析技术;(5)电分析化学联用技术;(6)纳米电化学;(7)电分析化学仪器的新发展及应用;(8)其它。

本次会议由第三世界科学院院士、中国科学院院士汪尔康先生担任大会主席,西北师范大学化学化工学院教授卢小泉博士担任大会秘书长。欲参加会议者请与卢小泉同志联系。

联系地址:兰州市 西北师范大学化学化工学院 邮编:730070

联系电话:0931-7971276(o),7973036(h)

电子信箱:luxq@nwnu.edu.cn

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net