

【引用格式】蔡庸军, 周志明, 隋峰. 基于高斯函数拟合的离子迁移谱反卷积算法[J]. 数字海洋与水下攻防, 2025, 8(1): 82-87.

基于高斯函数拟合的离子迁移谱反卷积算法

蔡庸军, 周志明*, 隋峰

(中船重工安谱(湖北)仪器有限公司, 湖北 宜昌 443003)

摘要 离子迁移谱 (Ion mobility spectrometry, IMS) 常用于一些气体或是易挥发性物质的痕量检测, 在海洋环保、工业过程控制和国防安全方面有广泛的应用。其区分物质的关键在于不同物质离子迁移谱谱图保留时间的差异。然而由于实际样品中组分复杂, 导致不同物质的谱峰交叠在一起难以区分。因此提出了一种基于高斯函数拟合的反卷积算法, 利用高斯函数逐步分解光谱, 通过降低每一级高斯函数的标准差使谱峰更加尖锐、展宽变窄, 以此实现谱峰的分离, 提高光谱分辨率。

关键词 离子迁移谱; 算法; 反卷积

中图分类号 TH761

文献标识码 A

文章编号 2096-5753(2025)01-0082-06

DOI 10.19838/j.issn.2096-5753.2025.01.011

Deconvolution Algorithm for Ion Mobility Spectrometry Based on Gaussian Function Fitting

CAI Yongjun, ZHOU Zhiming*, SUI Feng

(AlphaPec Instrument Co., Ltd., Yichang 443003, China)

Abstract Ion mobility spectrometry (IMS) is commonly used for trace detection of gases or volatile substances, and has wide applications in marine environmental protection, industrial process control, and national defense security. The key to distinguish substances lies in the difference in retention time of ion mobility spectrometry of different substances. However, due to the complex composition of actual samples, the overlapping peaks of different substances make it difficult to distinguish. Therefore, a deconvolution algorithm based on Gaussian function fitting is proposed, which gradually decomposes the spectrum using Gaussian function. By reducing the standard deviation of each level of Gaussian function, the spectral peaks become sharper and wider, so as to achieve peak separation and improve spectral resolution.

Key words ion mobility spectrometry; algorithm; deconvolution

0 引言

离子迁移谱 (Ion Mobility Spectrometry, IMS) 是基于不同气相离子在弱电场中迁移率的不同从而进行区分的一种化学分析技术^[1]。这一技术在海

洋环境监测与军事防御领域展现出独特的优势, 其低检出限与快速响应特性使之成为监测海洋大气中污染物及快速识别军事威胁物质 (如化学战剂) 的重要工具^[2]。该技术能够有效检测并区分海洋上空大气中的各类有害气体, 包括但不限于二氧化

收稿日期: 2024-09-27

作者简介: 蔡庸军 (1984-), 男, 硕士, 高级工程师, 主要从事精密光学仪器总体及软件设计研究。

*通信作者: 周志明 (1995-), 男, 硕士, 工程师, 主要从事化学分析中的软件算法研究。

硫、氨气及氮氧化物等。此外, 该技术还能精确分析海水蒸发产生的挥发性有机物 (VOCs): 烷烃衍生物、烯烃衍生物、醇类、酯类、芳香族化合物、有机酸等^[3-4], 为海洋污染源的追踪及生态评估提供科学依据。

离子迁移谱数据横坐标为迁移时间, 纵坐标为检测器的响应强度, 通过寻峰算法^[5]对谱峰位置进行解析以实现对待测组分的定性分析。然而实际测试过程中由于样品通常成分复杂, 导致各自谱峰交叠, 对通常寻峰造成困难。为了解决该问题常采用反卷积的算法对数据进行处理。反卷积算法的目的是寻找卷积方程的解, 而对于类高斯函数的谱峰, 反卷积算法能够使谱峰更加尖锐, 从而能够分离各峰位置。反卷积算法在信号处理、数字通讯、地震波分析、语音分析、图像处理等方面有着广泛的应用^[6], 常用的反卷积算法包括盲均衡算法、反馈网络自适应算法、最大似然算法和神经网络算法^[7-10]。需要说明的一点是经过卷积变换的数据已经损失掉了一部分信息, 反卷积算法不可能完美复原数据。尽管如此, 反卷积算法仍是提升光谱分析精度和分辨率的重要手段。

离子迁移谱作为一种光谱分析技术, 其光谱曲线特征类似于高斯峰形叠加, 与红外光谱、拉曼光谱等方法相似。因此, 科研界探索了多种反卷积算法, 旨在提取光谱数据内在信息, 抑制噪声及干扰峰。

CANALES-RODRÍGUEZ^[11]利用非模型化反卷积技术于扩散谱磁共振成像 (DSI), 增强了图像角分辨率, 揭示了新的纤维结构, 降低了伪影峰及不确定性, 推动了生物医学成像。BUSLOV、

DONG、KATRAŠNIK 等人^[12-20]的工作则扩展了反卷积算法应用, 涵盖红外光谱、拉曼光谱到质谱分析。BUSLOV^[12]通过反卷积处理区分了 α -D-葡萄糖红外光谱的复杂成分; DONG^[13]用傅里叶自反卷积算法增强了光谱分辨率; KATRAŠNIK^[14]将反卷积算法应用于声光可调谐滤波器光谱仪, 虽提升噪声水平, 但显著提高光谱分辨率于声光可调谐滤波器光谱仪; EDWARDS^[15]将反卷积技术应用于拉曼光谱, 并解析了古人类与现代人皮肤化学成分差异; LOPES、BOBIN、KANIE 等人^[16-20]则分别将反卷积算法应用于胰腺癌诊断、质谱定量分析、质谱解离过程研究、有机酸成分识别及马尿提取物分析, 均提升了准确性及效率。

基于上述研究成果的启发, 本文创新性地采用了一种基于高斯函数拟合的反卷积算法, 该算法通过逐步迭代的方式, 将光谱数据精确分解为一系列高斯峰, 并通过逐步减小各高斯峰的标准差, 实现了光谱分辨率的显著提升。这一方法不仅保留了光谱数据的原始特征, 还有效抑制了噪声干扰, 为离子迁移谱及其它光谱分析技术提供了更为精确、高效的数据处理手段。

1 算法原理

在开展工作前, 需要对离子迁移谱图的光谱特征进行分析, 其原始谱图如图 1 (a) 所示, 其为倒峰, 且峰形基本符合高斯峰的特征。因此, 可以考虑利用高斯函数逐级分解光谱数据, 将特征峰分解为多个高斯峰, 这些高斯峰相互叠加共同构成原始数据, 并且通过调整每一级的高斯函数中标准差的值, 实现光谱的反卷积效果。

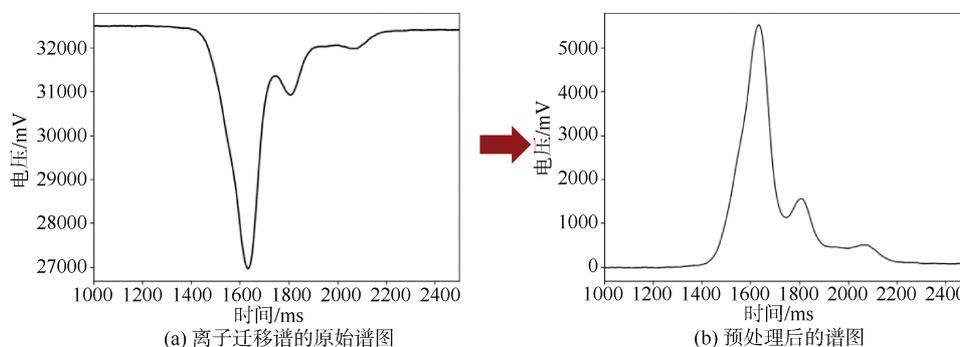


图 1 光谱数据的预处理

Fig. 1 Preprocessing of spectral data

- 1) 读取离子迁移谱图的数据 (1 列浮点数), 总共有 n 个点;
- 2) 将离子迁移谱数据的每个点减去离子迁移谱数据的第 1 个点并乘以 -1 , 记为散点 $S(n)$;
- 3) 对离子迁移数据从 0 开始编号, 一直到 $n-1$, 例如 0、1、2、...、 $n-2$ 、 $n-1$, 总共 n 个点, 记为散点 $N(n)$, 将 $N(n)$ 和 $S(n)$ 组成一个 $n \times 2$ 的矩阵 M , 矩阵第 1 列为 $N(n)$, 第 2 列为 $S(n)$;
- 4) 设置循环次数 L 和限制值 V ;
- 5) 通过寻峰算法计算 M 的一个最强峰位置, 得到峰的位置 b 和峰的强度 a , 如果 a 小于 V 则直接从第 9 步继续执行;
- 6) 在 M 中峰的位置 b 左右各取 W 个点, 组成新 $2W \times 2$ 矩阵 M_2 ;
- 7) 利用高斯函数拟合散点 M_2 , 并在列表 List 中记录该次循环中 a 、 b 以及拟合的高斯函数中标准差 std 的参数, 将 M 的第 2 列减去拟合的高斯函数, 并替换掉原先的矩阵 M ;
- 8) 循环计数加 1 (从 0 开始), 如果循环计数小于等于 L 则从第 5 步继续执行, 否则记录矩阵 M 为残差 R , 继续执行第 9 步;
- 9) 得到包含利用高斯函数逐步分解的参数 List 以及残差 R ;
- 10) 通过降低 List 中每一个 std 的值, 并利用 List 中的其它参数逐步计算每级高斯函数, 累加到残差中就能得到反卷积的数据。

2 算法的测试

利用高斯函数对谱图进行分解, 离子迁移谱图的原始数据均值在 30 000 左右, 且为倒峰, 不利于高斯拟合。因此, 在步骤 2 中将原始数据处理, 其基线位于 0 附近且特征峰均为正峰。

从变换后的谱图 (图 1 (b)) 中可以很轻易地找到谱图中的最大值点, 其坐标为 (1634, 5526)。在该点处前后各取 20 个点, 根据这些点利用最小二乘 (公式 (1)) 求解多项式二次函数 (公式 (2)) 的各项系数 a 、 b 和 c , 并利用二次函数的对称轴求值公式计算理论极值位置。

$$\text{Coefficient} = (M^T \times M)^{-1} \times M^T \times L \quad (1)$$

式中: M 为样本矩阵; L 标签矩阵; Coefficient 系

数矩阵。

$$y = a \times x^2 + b \times x + c \quad (2)$$

经过计算理论极大值坐标为 (1633.34, 5522.41)。此时根据高斯函数 (公式 (3)), 其中 A 和 B 以及通过二次函数的极值公式求得, 分别为理论极大值坐标的纵坐标和横坐标, 此时仅需知道 std^2 就可以得到高斯函数的所有参数, 而高斯函数并非线性函数, 无法直接用最小二乘的方式求解 std^2 , 需要对高斯函数进行变换。

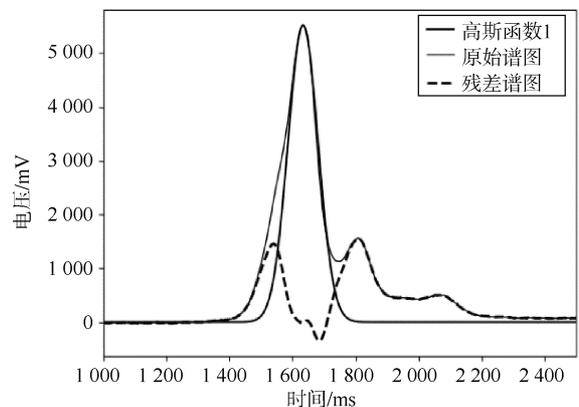
$$y = A \times e^{-\frac{(x-B)^2}{2 \times std^2}} \quad (3)$$

式中: A 为幅值; B 为对称中心位置; Std 为标准差。

对公式 (3) 两边进行一系列等价变换后得到公式 (4), 此时对于 std^2 , 等式两边为线性关系, 将附近的坐标点带入到 x 和 y 中, 构建矩阵 (算法实现中的 M_2), 并对该矩阵用最小二乘求解, 可以计算出 std^2 的值为 2 283.15, std 为 47.78, 此时得到该高斯函数中所有的必要参数。

$$2 \times (\ln(A) - \ln(y)) \times std^2 = (x - B)^2 \quad (4)$$

将原始数据减去上述高斯函数, 其残差如图 2 (a) 所示, 实现 1 次高斯函数分解, 并记录高斯函数的所有参数。对残差谱图重复上述步骤依次分解, 图 2 (b) 为高斯函数的 3 次分解, 将图中每一级高斯函数相加就得到重构数据。理论上来说, 分解次数如果足够多, 残差谱能够变为 $y=0$ 的一条直线, 可以不需要残差谱图, 但并没有必要这样做, 首先会产生极大的运算开销, 其次对于噪声的高斯分解也并没有意义, 因此保留残差并设置适当的分解次数更符合实际应用需求。



(a) 高斯函数分解谱图1次

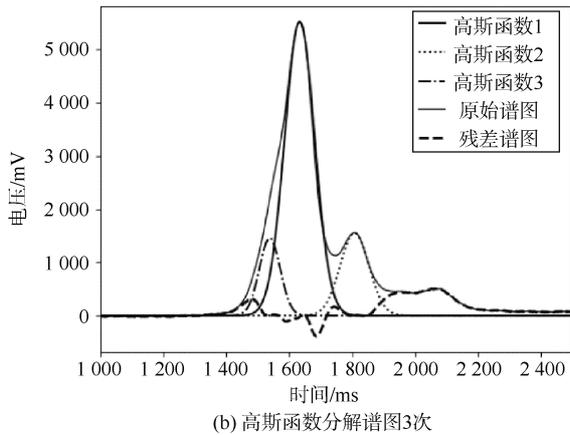


图 2 高斯函数分解谱图的过程

Fig. 2 Process of decomposing spectrum using Gaussian function

在上述步骤中对原始数据实现了高斯函数分解, 通过缩小每个分解出高斯函数中的 std 值能够实现原始数据的反卷积, 高斯函数的分解次数设置为 9 次, 将分解的高斯函数进行叠加得到重构数据 (rebuild)。通过图 3 可以看出: 随着 std 值的逐步减小, 重构数据中的特征峰逐渐分离和凸显, 反卷积效果良好。

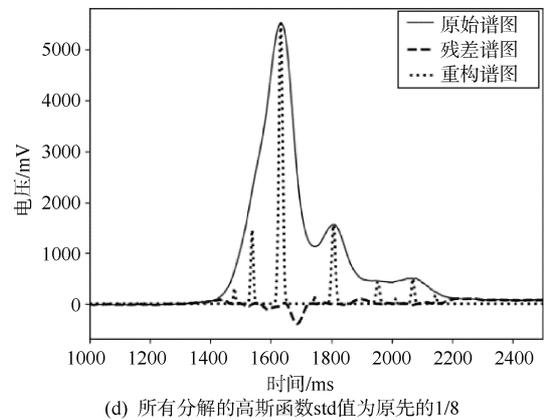
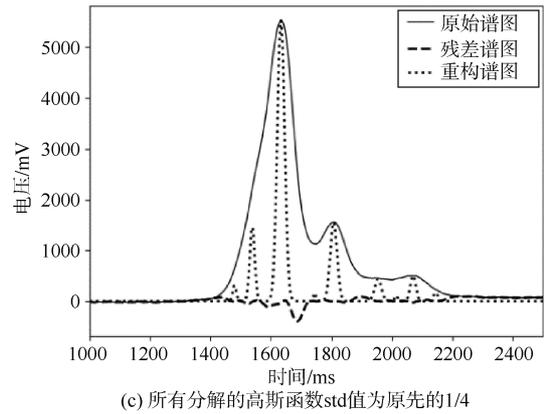
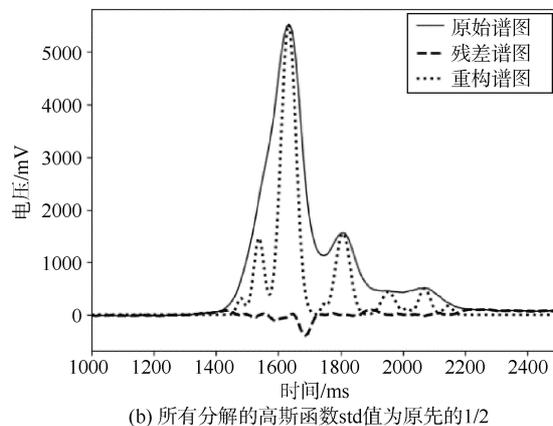
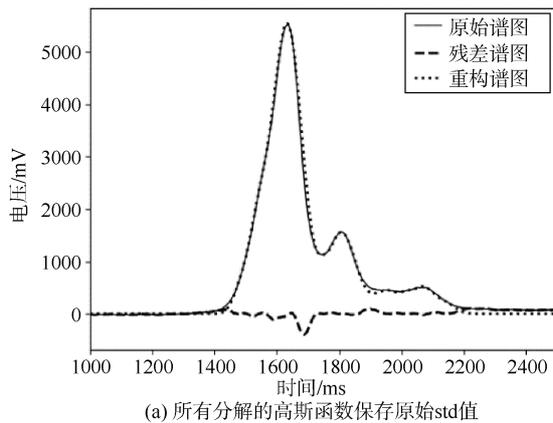


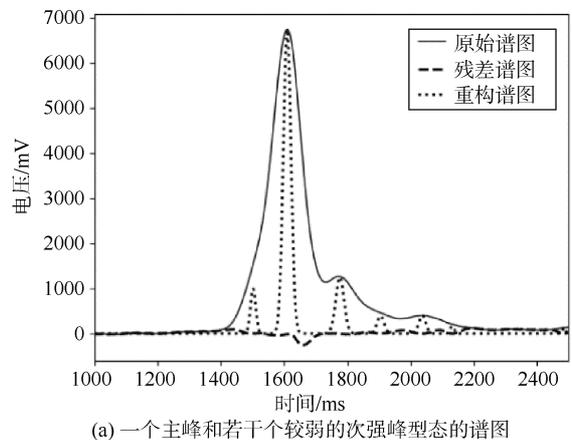
图 3 不同 std 值重构光谱的变化图

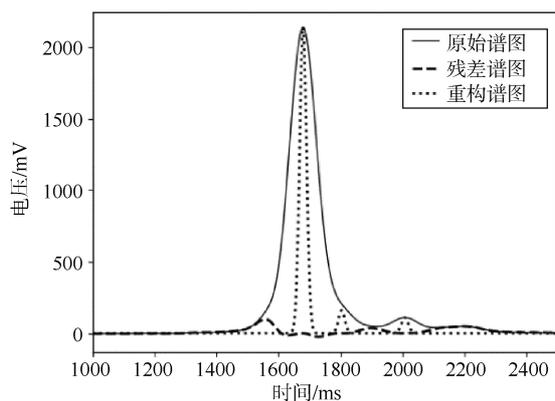
Fig. 3 Variation of reconstructed spectra with different std values



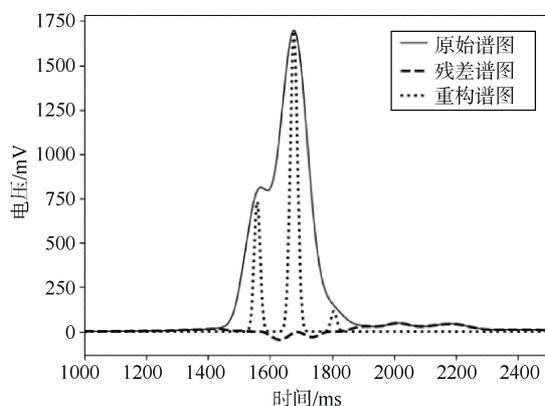
3 不同离子迁移谱图的测试效果

将该算法应用到不同的离子迁移谱图中, 将分解次数阈值设置为 9, 最大峰强限值设置为 100, 并将标准差缩小至原 1/4。图 4 (a) 中经过反卷积运算后主峰和次强峰凸显, 主峰位于 1 608, 次强峰为 1 512 和 1 785。图 4 (b) 经过反卷积运算后仅有 1 个位于 1 682 的主峰。图 4 (c) 经过运算后

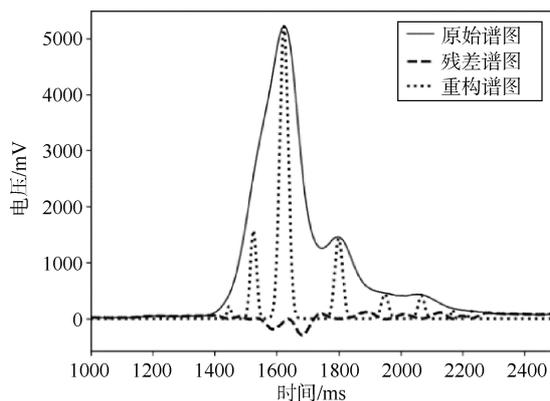




(b) 单个主峰型态的谱图



(c) 一个主峰和一个较强的次强峰型态的谱图



(d) 一个主峰和若干个较强次强峰型态的谱图

图 4 基于高斯函数拟合的反卷积算法对 4 种不同型态谱图的反卷积效果

Fig. 4 Deconvolution effect of deconvolution algorithm based on Gaussian function fitting on 4 different types of spectrograms

分离出一个位于 1 676 的主峰和位于 1 562 的一个次强峰。图 4 (d) 经过运算后分离出除了 1 627 的主峰外, 还分离出位于 1 514 和 1 796 的次强峰。

以上测试结果说明该算法的反卷积效果良好, 不同型态的谱图经过处理后均能够使原本谱峰凸显和分离, 有利于后续对谱图的解析。

4 结束语

反卷积算法根据高斯函数对谱图进行分解, 通过降低每个高斯函数的标准差值来实现反卷积效果, 文中展示了该算法对不同型态谱图的反卷积处理结果, 可显著提高谱图的分辨率, 排除谱图中的干扰。除此之外, 该算法与小波变换类似, 其还可用于谱图的降噪与数据的压缩; 在 python 中进行实现, 易于用 C 语言移植, 并可在仪器上位机或是下位机中实现; 有较好的反卷积效果, 有助于谱峰的识别, 增强仪器对于不同物质的分辨能力, 并降低噪声; 实现复杂度较低, 对运行环境要求不高, 便于不同平台和系统上的移植。

同时, 在测试过程中发现该算法存在一些限制: 1) 需要保证谱峰基线在 0 附近; 2) 该分解会出现一些倒峰。对于第 1 个问题, 当基线存在严重漂移时需要对谱图进行扣除基线的预处理。对于第 2 个问题, 由于目前算法无法对 2 个重叠度过高的高斯峰进行区分, 其次实际谱峰可能和理论的高斯峰之间存在差异。为了适应更广泛的应用场景, 可根据使用的具体情况对本文算法做进一步改进。

参考文献

- [1] BORSODORF H, MAYER T, ZAREJOUSHEGHANI M, et al. Recent developments in ion mobility spectrometry[J]. Applied Spectroscopy Reviews, 2011, 46 (6): 472-521.
- [2] ZIMMERMANN S, ABEL N, BAETHER W, et al. An ion-focusing aspiration condenser as an ion mobility spectrometer[J]. Sensors & Actuators B: Chemical, 2007, 125 (2): 428-434.
- [3] KARASEK F W. Plasma chromatography of the polychlorinated biphenyls[J]. Analytical Chemistry, 1971, 43 (14): 1982-1986.
- [4] KARASEK F W, KIM S H, HILL H H. Mass identified mobility spectra of p-nitrophenol and reactant ions in plasma chromatography[J]. Analytical Chemistry, 1976, 48 (8): 1133-1137.
- [5] 埃森门. 离子迁移谱[M]. 北京: 国防工业出版社, 2010.
- [6] JOHO M, SCHNITER P. Frequency domain realization of a multichannel blind deconvolution algorithm based on the natural gradient[J]. Journal of Computer Science and Engineering, 2003: 4988964.

- [7] RAU U, CORNWELL T J. A multi-scale multi-frequency deconvolution algorithm for synthesis imaging in radio interferometry[J]. *Astronomy & Astrophysics*, 2012, 532: 201117104.
- [8] FIORI S. A contribution to (neuromorphic) blind deconvolution by flexible approximated Bayesian estimation[J]. *Signal Processing*, 2001, 81 (10): 2131-2153.
- [9] 方标, 黄高明, 高俊. 多通道盲反卷积算法综述[J]. *信号处理*, 2013, 29 (6): 712-722.
- [10] FIORI S. Fast fixed-point neural blind-deconvolution algorithm[J]. *IEEE Transactions on Neural Networks*, 2004, 15 (2): 455-459.
- [11] CANALES-RODRÍGUEZ E J, ITURRIA-MEDINA Y, ALEMÁN-GÓMEZ Y, et al. Deconvolution in diffusion spectrum imaging[J]. *Neuroimage*, 2010, 50 (1): 136-149.
- [12] BUSLOV D K, NIKONENKO N A, SUSHKO N I, et al. Analysis of the results of α -d-glucose Fourier transform infrared spectrum deconvolution: comparison with experimental and theoretical data[J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 1998, 1425 (98): 94592341.
- [13] DONG Y, XIANG L B, ZHAO B. Theoretical study on Fourier self-deconvolution (FSD) of Fourier transform spectrum[J]. *Acta Photonica Sinica*, 2002, 31 (7): 841-850.
- [14] KATRAŠNIK J, PERNUŠ F, LIKAR B. Deconvolution in acousto-optical tunable filter spectrometry[J]. *Applied Spectroscopy*, 2010, 64 (11): 1265-1273.
- [15] EDWARDS H, FARWELL D W, WILLIAMS A C, et al. Novel spectroscopic deconvolution procedure for complex biological systems: vibrational components in the FT-Raman spectra of ice-man and contemporary skin[J]. *Journal of the Chemical Society Faraday Transactions*, 1995, 91 (21): 3883-3887.
- [16] LOPES P C, MOREIRA J A, ALMEIDA A, et al. Discriminating adenocarcinoma from normal colonic mucosa through deconvolution of Raman spectra[J]. *Journal of Biomedical Optics*, 2011, 16 (12): 127001.
- [17] BOBIN S, POPOT M A, BONNAIRE Y, et al. Approach to the determination of insulin-like-growth-factor-I (IGF-I) concentration in plasma by high-performance liquid chromatography-ion trap mass spectrometry: use of a deconvolution algorithm for the quantification of multiprotonated molecules in electrospray ionization[J]. *Analyst*, 2001, 126: 1996-2001.
- [18] KANIE O, KANIE Y, DAIKOKU S, et al. Multi-stage mass spectrometric information obtained by deconvolution of energy-resolved spectra acquired by triple-quadrupole mass spectrometry[J]. *Rapid Communications in Mass Spectrometry*, 2011, 25 (11): 1617-1624.
- [19] HALKET J M, PRZYBOROWSKA A, STEIN S E, et al. Deconvolution gas chromatography/mass spectrometry of urinary organic acids-potential for pattern recognition and automated identification of metabolic disorders[J]. *Rapid Communications in Mass Spectrometry*, 1999, 13 (4): 279-284.
- [20] DEMIR C, HINDMARCH P, BRERETON R G. Deconvolution of a three-component co-eluting peak cluster in gas chromatography-mass spectrometry[J]. *Analyst*, 2000, 125 (2): 287-292.

(责任编辑: 曹晓霖)