

一种 X 射线荧光光谱数据批处理新方法的研究

易龙涛 刘志国 陈曼 王锴 彭诗琪 赵为刚 何佳霖 赵广翠

北京师范大学核科学与技术学院射线束技术与材料改性教育部重点实验室, 北京 100875

摘要 论述了一种高精度、快速的能量色散 X 射线荧光(EDXRF)光谱自动分析方法,包括平滑、背底扣除和寻峰等计算过程。该方法不仅可用于单幅谱图的分析,也可用于大批量数据的批处理分析。对实验室测得的微区 X 射线荧光光谱数据的测试结果表明,该方法同常用的高斯拟合方法相比,结果基本一致,但算法更加简单,能够有效地识别出弱峰。在当前主流配置计算机上利用该方法进行多文件连续分析时,单个文件的平均用时小于 1 s,能够有效缩短数据分析时间、提高工作效率。

关键词 光谱学;能量色散 X 射线荧光;批处理;双向平滑;三次样条函数;对称零面积变换

中图分类号 O433.4 **文献标识码** A **doi:** 10.3788/LOP51.073001

A New Automatic Analysis Method for X-Ray Fluorescence Spectrometric Qualitative Analysis

Yi Longtao Liu Zhiguo Chen Man Wang Kai Peng Shiqi Zhao Weigang
He jialin Zhao Guangcui

Key Laboratory of Beam Technology and Material Modification of Ministry of Education, College of Nuclear Science and Technology, Beijing Normal University, Beijing 100875, China

Abstract A method is proposed to automatically analyze energy dispersive X-ray fluorescence (EDXRF) spectrum fast and accurately. Smoothing of spectrum, background subtraction and peak searching are used in the method. It not only can be applied to analyze single spectrum, but also can be used for analyzing a large number of spectra continuously. The method is evaluated by using the spectrum analyzed by micro-X-ray fluorescence in laboratory. Compared with Gaussian fitting, the result of the symmetric zero area method is basically the same but this method has the advantage of short computing time. It can obtain a good result in the case of the spectrum with big noise and well ability of picking out the weak peaks. On ordinary PC, the average time of analyzing a single file is less than 1 s while analyzing hundreds of files using this method. It can obviously reduce the work time and improve work efficiency.

Key words spectroscopy; energy dispersive X-ray fluorescence; batch processing; double smoothing; cubic spline function; symmetric zero-area conversion

OCIS codes 300.6560; 200.4560; 260.2510

1 引言

X 射线荧光分析技术作为一种快速、准确、无污染的无损多元素分析技术在诸多领域得到了广泛的应用。尤其是能量色散 X 射线荧光(EDXRF),以其体积小、结构简单、自动化程度高且价格低廉的优点,迅速发展成为一种普遍的多元素同时分析手段。计算机科学的迅速发展,电子元件的大规模集成化和功能化等,为开发更加智能化、小型化的高性能 EDXRF 设备提供了必要的硬件支持。因此,数据分析方法就成了

收稿日期: 2014-01-13; 收到修改稿日期: 2014-02-24; 网络出版日期: 2014-06-06

基金项目: 北京师范大学自主科研基金重点项目(2012LZD07)

作者简介: 易龙涛(1988—),男,博士研究生,主要从事核电子学、数据分析等方面的研究。

E-mail: yilongtao@mail.bnu.edu.cn

导师简介: 刘志国(1968—),男,博士,教授级高工,主要从事 X 射线光学器件研发和应用、核电子学与自动控制等方面的研究。E-mail: liuzhiguo512@hotmail.com (通信联系人)

其中的重要因素。但目前的大部分方法,无论手动还是自动,均是针对单幅谱图^[1-5]。虽可以得到精确的分析结果,却无法应对诸如原位微区 X 射线荧光二维扫描成像分析^[6]中的大批量复杂数据分析。这无疑成了开发新型高性能 EDXRF 设备的重要制约因素。因此,研究一种新的快速自动分析方法就显得十分必要。

本文提出了一种算法简单的快速自动分析方法,可以实现对大批量复杂 EDXRF 谱线数据的连续自动分析。

2 数据分析

EDXRF 设备测得的光谱主要由背景谱、特征 X 射线谱及其逃逸峰、脉冲堆积等构成。要准确分析出 EDXRF 光谱,需经过平滑滤波、背底扣除、寻峰等过程,其中寻峰的准确性直接决定了分析结果的准确性。

2.1 双向平滑

2.1.1 右移平滑

类似于均值平滑,取数据点及其前面的 $m-1$ 个点的平均值作为数据点的平滑值,对谱线做第一次平滑,平滑后的谱线整体向右移动。右移平滑的数学表达式为

$$(x'_i, y'_i) = \begin{cases} \frac{1}{m} \sum_{j=i-m+1}^{j=i} (x_i, y_i), & m \leq i \leq n \\ \frac{1}{m} \sum_{j=1}^{j=i} (x_i, y_i), & 1 \leq i < m \end{cases}, \quad (1)$$

式中 m 为参与平滑的点数, n 为谱线的总数据点数, (x_i, y_i) 为原始谱线数据点, (x'_i, y'_i) 为右平滑后的平滑数据点。

2.1.2 左移平滑

为消除右移平滑造成的谱线移动,需对谱线进行一次反向平滑。类似于右平滑,取点 (x'_i, y'_i) 及其后面的 $m-1$ 个点的平均值作为该点的平滑值。左移平滑的数学表达式为

$$(X_i, Y_i) = \begin{cases} \frac{1}{m} \sum_{j=i}^{j=i+m-1} (x'_i, y'_i), & 1 \leq i \leq n-m+1 \\ \frac{1}{m} \sum_{j=i}^{j=n} (x'_i, y'_i), & n-m+1 < i \leq n \end{cases}, \quad (2)$$

式中的 (X_i, Y_i) 为左平滑后的数据点,即最终的平滑数据。

图 1 和图 2 给出了以实验室三维共聚焦设备测自油画的微区 EDXRF 光谱进行平滑处理的效果,其中 X 射线光管采用的是英国 Oxford 公司 XTG UltraBright Microfocus X-ray Source 光管,探测器采用的是美国 Amptek 公司 X-123 detector 探测器。由图 1 可知,平滑后的谱峰高度较原始谱有所降低,原始谱线中的“毛刺”基本消除,谱线的数据趋势得到了很好的反映,且没有出现峰位移动。需要指出的是, m 需根据数据的

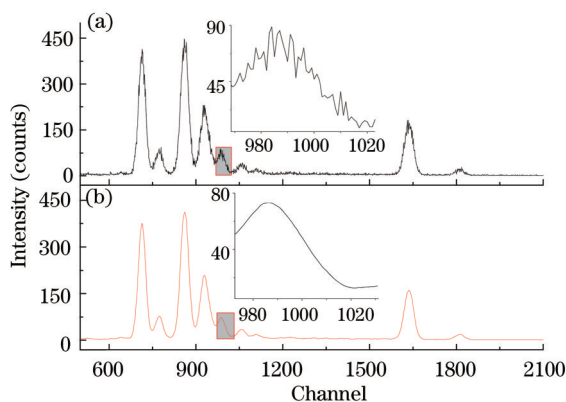


图 1 (a) 原始 EDXRF 谱线; (b) 15 点平滑后的谱线
Fig.1 (a) Original EDXRF spectra; (b) smoothed spectra by 15 points

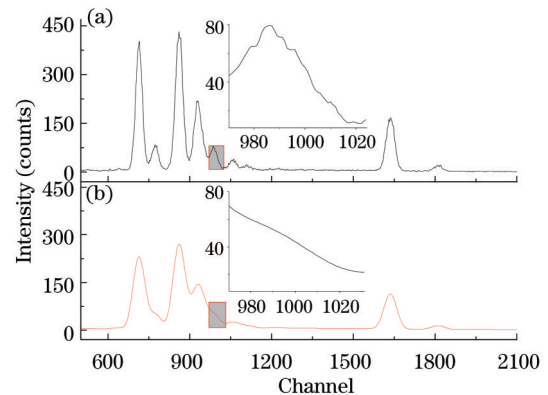


图 2 (a) 3 点平滑后的谱线; (b) 40 点平滑后的谱线
Fig.2 (a) Smoothed spectra by 3 points; (b) smoothed spectra by 40 points

扰动程度确定,实验结果表明 m 选取 15~30 平滑效果较好。对于扰动较小的谱图 m 值可以取得较小。实验中具体表现为,若 m 偏小,则可能出现如图 2(a)所示的情况,虽可以反映出谱线的数据趋势,但无法较好地消除“毛刺”;若 m 过大,则可能出现如图 2(b)所示的情况,虽可以很好地消除毛刺,但会造成峰形的畸变。实验中,可通过手动选取离散度最大峰计算其方差的方式计算出 m 值。

2.2 三次样条扣除背底

通常,将背底看作分段光滑的曲线。剥峰法、正交多项式背景估计^[7,8]等传统的方法,主要通过人为干预的重复计算获取背底曲线。采用三次样条插值函数的方法单次获取背底曲线。为获得三次样条函数差值节点,给出以下关系式:

$$\nabla Y_i = \frac{Y_i - Y_{i-1}}{X_i - X_{i-1}} \quad (3)$$

若点 (X_i, Y_i) 满足:

$$\nabla Y_i < 0, \quad (4)$$

$$\nabla Y_{i+1} > 0, \quad (5)$$

则记为新的点 (X'_j, Y'_j) , 即三次样条函数的插值节点,且满足以下条件:

$$X'_j = X_i, \quad j = 1, 2, \dots, l, \quad (6)$$

$$Y'_j = Y_i, \quad j = 1, 2, \dots, l. \quad (7)$$

设区间 $[X'_j, X'_{j+1}]$ 上的函数为 $S(x)$, 由三次样条函数^[9-10]的定义,则有:

$$S(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \sum_{k=1}^{l-1} b_k (x - x_k)^3 / 6, \quad X'_j \leq x \leq X'_{j+1}, \quad (8)$$

式中 $b_k = S'''(x_k)$, $k=1, 2, \dots, l-1$, 其中 l 为插值点数。

实验中,采用三弯矩插值法^[8]求得 $[X'_j, X'_{j+1}]$ 上的 $S(x)$, 作为原始谱线的背底曲线。图 3 和图 4 分别给出了以 Amptek 公司提供的示例 EDXRF 谱线和实验室三维共聚焦设备测自油画的微区 EDXRF 光谱进行背底扣除的效果。图 3 中的厚背底谱线,扣除效果十分明显,峰高明显下降且无峰位移动,对厚背底上的弱峰有一定的突显效果,扣除效果优于剥峰法。图 4 中的薄背底谱线,峰高有所下降且无峰位移动,两种方法的扣除效果基本相同。相比于剥峰法,三次样条函数法无需重复计算,计算量明显减小。

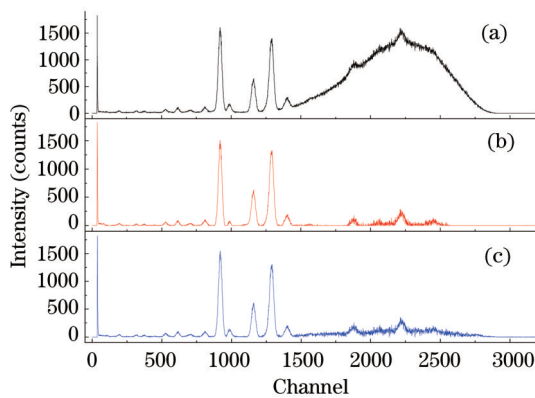


图 3 (a)原始的示例 EDXRF 谱线;(b)三次样条函数法扣背底后的谱线;(c)剥峰法 3000 次计算扣背底后的谱线

Fig.3 (a) Original sample EDXRF spectra; (b) background subtracted spectra by cubic splines method; (c) background subtracted spectra by stripping peaks method with 3000 times repetitive computation

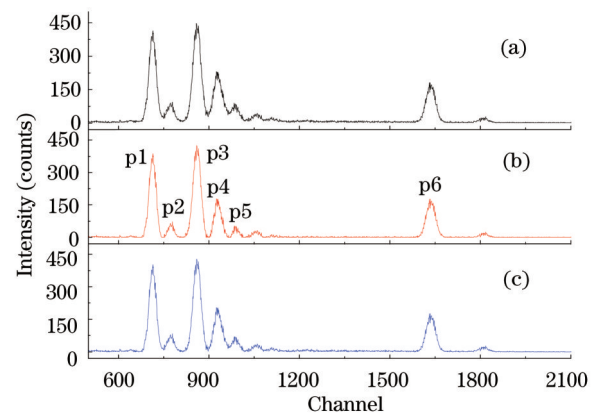


图 4 (a)原始的微区 EDXRF 谱线;(b)三次样条函数法扣背底后的谱线;(c)剥峰法 3000 次计算扣背底后的谱线

Fig.4 (a) Original micro EDXRF spectra; (b) background subtracted spectra by cubic splines method; (c) background subtracted spectra by stripping peaks method with 3000 times repetitive computation

2.3 对称零面积变换寻峰

对称零面积变换法^[11-13]是把面积为零的对称的“窗”函数与光谱数据进行褶积变换。可以证明,面积为

零的对称“窗函数”对线性基底的褶积变换结果将为零,只有在峰存在的条件下不为零,且变换值最大处为峰值所在。对称零面积变换寻峰的关键是找到合适的“类峰型窗口函数”,以达到最优寻峰效果。实验中对扣背底并平滑处理后的数据寻峰,选取高斯函数作为窗口函数。其具体公式如下所示:

$$C_i = G_i - d, \quad (9)$$

$$d = \frac{1}{w} \sum_{-n}^n G(i), \quad (10)$$

$$G(i) = \exp[-4 \ln 2(i/H)^2], \quad (11)$$

式中 $w=2n+1$ 为“窗”宽,即变换宽, H 为峰的最小半宽度,实验中分别取 11 和 3。针对 EDXRF 谱图中背底曲线上存在的极小弱峰的干扰,寻峰过程中设置了寻峰阈值,寻峰阈值由平均背底噪声值与最高峰值之比计算得出,寻峰阈值为 5%。

表 1 列出了分别用对称零面积变换寻峰,Matlab2012b 软件和 Origin8.5 软件对图 4(b) 寻峰的结果,其中 Matlab2012b 软件采用的是多峰高斯拟合寻峰,Origin8.5 采用的是手动选择谱峰。对比可以发现,三种方法的寻峰结果基本一致,多峰高斯拟合无法识别出 p5 峰。相比于常用的高斯拟合法,对称零面积变换法对弱峰和重叠峰具有更好的识别能力且算法简单。

表 1 EDXRF 光谱寻峰结果

Table 1 Results of peak searching of ERXRF spectra

	p1	p2	p3	p4	p5	p6
Symmetric zero area	712.0	773.0	860.0	928.0	987.0	1635.0
Origin8.5	713.0	774.1	860.2	929.5	985.1	1635.7
Matlab2012b	713.1	774.2	860.2	929.1	—	1636.0

3 计算机处理

实验中,采用 delphi7 软件编写了光谱分析程序,数据处理流程如图 5 所示。

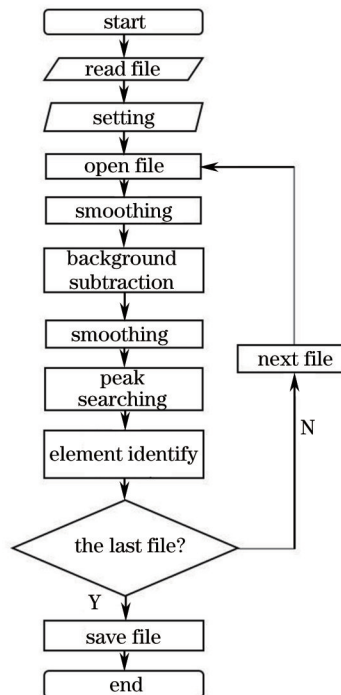


图 5 计算机处理流程

Fig.5 Flow chart of computer processing

整个分析程序由文件读取、平滑、背底扣除、寻峰、元素识别和文件存储七个模块组成。其中,寻峰模块包括峰位确定、峰边界确定和峰面积计算三部分。其运算过程如下:首先通过对称零面积变换法寻峰,然后由峰位向两侧寻找峰边界并计算峰面积等数据。设置包括平滑点数 m 、寻峰阈值和元素识别阈值三个参数的设定。文件读入之后,需手动设定以上三个参数,之后的整个分析过程由程序自动完成,无需人为干预。需要指出的是,整个过程使用同一组参数。图6分别给出了以剥峰法扣背底并以高斯拟合寻峰的传统方法和对称零面积变换法对图1中的原始谱图处理的效果图,两种方法的结果基本一致。表2列出了在当前主流配置计算机上批处理多个实验室三维共聚焦设备测自油画的微区EDXRF光谱文件的运算时间,可见单个文件的平均处理时间均小于1s,有效缩短了数据分析时间。

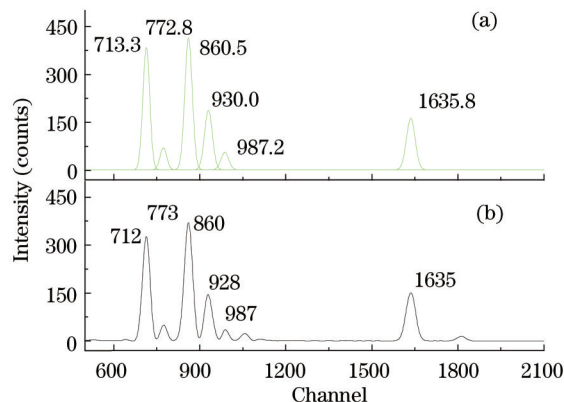


图6 (a) 传统方法;(b) 对称零面积变换法

Fig.6 (a) Traditional method;(b) symmetric zero area method

表2 批处理文件时间

Table 2 Time of batch multiple files

Files	50	100	200	500	1000
Operation time /s	29.76	43.04	79.76	182.16	358.66

4 结 论

通过对实际EDXRF光谱文件的分析表明,文中数据分析方法具有算法简单,计算精度高的特点。为验证分析结果的准确性,分别采用对称零面积变换法和高斯拟合方法对微区EDXRF光谱进行寻峰测试,结果基本一致,且能够有效地识别出弱峰。为验证对大批量文件的批处理分析能力,以实验室三维共聚焦设备测得的数据进行测试,结果表明能够有效缩短数据处理时间、提高工作效率。相比于传统方法,只需在开始阶段设定相应参数,整个中间过程无需人为干预。

实际应用方面,该方法已经应用于实验室微区X射线荧光二维扫面成像分析并取得了较好的效果。但需要指出的是,当同批次文件中包含测自不同实验条件且离散程度差异较大的文件时,分析结果的准确性有所降低。作为一种快速自动分析方法,该方法不仅可用于单个X射线荧光光谱的分析,也可以用于大批量光谱数据的分析,对开发高性能的X射线荧光分析设备有着实际的意义。

参 考 文 献

- Zhang Lianxiang, Fu Bin. Advances in X-ray fluorescence spectrometry [J]. Chinese Journal of Inorganic Analytical Chemistry, 2013, 3(3): 1-7.
章连香, 符 斌. X射线荧光光谱分析技术的发展[J]. 中国无机分析化学, 2013, 3(3): 1-7.
- Yang Mingtai, Tang Hui. The actualities and trend of energy dispersive X-ray fluorescence spectrometry [J]. Nuclear Electronics & Detection Technology, 2011, 31(12): 1307-1311.
杨明太, 唐 慧. 能量色散X射线荧光光谱仪现状及其发展趋势[J]. 核电子学与探测技术, 2011, 31(12): 1307-1311.
- Ji Ang, Tao Guangyi, Zhuo Shangjun, et al.. X-Ray Fluorescence Spectrometry [M]. Beijing: Science Press, 2003. 87-88.
吉 昂, 陶广义, 卓尚军, 等. X射线荧光光谱分析[M]. 北京: 科学出版社, 2003. 87-88.

- 4 Zhou Yijun, Zhang Linyan, Dai Ting. Study on qualitative analysis algorithm of energy dispersive X-ray fluorescence spectrometry [J]. Analytical Instrumentation, 2010, (6): 77-80.
周怡君, 张林艳, 戴挺. 能量色散X射线荧光光谱定性分析算法研究[J]. 分析仪器, 2010, (6): 77-80.
- 5 Zhang Wei, Zhang Yujun, Chen Dong. Quantitative analysis of nickel in soil samples using X-ray fluorescence spectroscopy [J]. Laser & Optoelectronics Progress, 2012, 49(1): 013002.
章炜, 张玉钧, 陈东. 土壤重金属镍元素的X射线荧光定量分析[J]. 激光与光电子学进展, 2012, 49(1): 013002.
- 6 Xu Tao, Luo Liqiang. Developments of micro-X-ray fluorescence spectrometer and applications [J]. Rock and Mineral Analysis, 2011, 30(3): 375-383.
许涛, 罗立强. 原位微区X射线荧光光谱分析装置与技术研究进展[J]. 岩矿测试, 2011, 30(3): 375-383.
- 7 Zhang Linyan, Dai Ting, Zhou Yijun. Study on a background subtraction method for energy dispersive X-ray fluorescence spectrum [J]. Chinese Journal of Analysis Laboratory, 2009, 28(12): 111-114.
张林艳, 戴挺, 周怡君. 能量色散X射线荧光光谱背景扣除方法的探讨[J]. 分析试验室, 2009, 28(12): 111-114.
- 8 Sun Huaao. EDXRFA Software Development And Algorithm Design [D]. Chengdu: Chengdu University of Technology, 2008. 25-52.
孙华傲. EDXRFA软件开发与算法设计[D]. 成都: 成都理工大学, 2008. 25-52.
- 9 Wu Boying, Wang Demin, Ding Xiaohua, *et al.*. Numerical Analysis Principle [M]. Beijing: Science Press, 2003. 134-141.
吴勃英, 王德民, 丁效华, 等. 数值分析原理[M]. 北京: 科学出版社, 2003. 134-141.
- 10 Gao S, Zhang Z, Cao C. On a generalization of cubic spline interpolation [J]. Journal of Software, 2011, 6(9): 1632-1639.
- 11 Pang Jufeng, Zheng Guifang, Hou Xiaofeng. The symmetric zero-area conversion method for peak-seeking [J]. Atomic Energy Science and Technology, 1987, 21(3): 270-279.
庞巨丰, 郑桂芳, 侯晓凤. 对称零面积变换法找峰[J]. 原子能科学技术, 1987, 21(3): 270-279.
- 12 Shang Fengjun, Wang Haixia, Zhou Rongsheng. The discussion and realization of seek-peak algorithm on X fluorescence spectrometry analysis [J]. Computing Techniques for Geophysical and Geochemical Exploration, 2000, 22(4): 364-368.
尚凤军, 王海霞, 周蓉生. 荧光光谱分析中寻峰算法的探讨及实践[J]. 物探化探计算技术, 2000, 22(4): 364-368.
- 13 Pang Jufeng, Zheng Guifang. A comparison of automatic peak search methods [J]. Acta Metrologica Sinica, 1985, 6(3): 213-220.
庞巨丰, 郑桂芳. 自动找峰方法的比较[J]. 计量学报, 1985, 6(3): 213-220.