

利用多数据处理方法提高 LIBS 谱信号质量

杨崇瑞¹, 汪家升², 盛新志², 娄淑琴¹

(1. 北京交通大学 电子信息工程学院, 北京 100044; 2. 北京交通大学 理学院, 北京 100044)

摘要: 基于分段光谱特征值提取法和小波变换算法等多个数据预处理方法, 分别针对分段基线差异及光谱噪声等严重影响激光诱导击穿光谱(LIBS)信号质量的主要影响因素, 开展光谱信号预处理研究。基于实验室 LIBS 实验装置, 通过实验验证, 基于多通道光谱仪不同波段光谱特征值提取, 提出了一种简单易行的多组数据中特征值点连接的方法, 有效地提高了 LIBS 光谱信号的基线平直度, 并得出以小波变换算法进行 LIBS 谱线信号去噪的最佳算法参数。在上述工作的基础上, 使用基于误差反向传播的人工神经网络方法, 实现了纯铜和不锈钢等物质种类的有效识别, 研究结果表明, 综合利用多数据处理方法进行 LIBS 技术中光谱信号处理可以有效提高谱线分析和识别的质量。

关键词: 激光诱导击穿光谱; 基线校正; 谱线识别; 小波算法; 人工神经网络

中图分类号: O433.4 **文献标志码:** A **文章编号:** 1007-2276(2014)11-3807-06

Technique to improve quality of LIBS spectrum signal based on multiple data processing methods

Yang Chongrui¹, Wang Jiasheng², Sheng Xinzhi², Lou Shuqin¹

(1. School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China;

2. School of Science, Beijing Jiaotong University, Beijing 100044, China)

Abstract: Based on multiple signal process methods, such as segmented spectral feature extraction and wavelet transform algorithm, the pre-spectrum signal treatment technique was investigated to decrease the difference of segmented spectral baseline and lower the spectral noise, and thus the signal quality in laser-induced breakdown spectroscopy (LIBS) was improved. Based on extracting the characteristic value in different spectral bands of multi-channel spectrometers, a simple method was presented to connect the characteristic value in different segmented data and effectively flats the signal baseline. Through analyzing the experimental data, the wavelet transform was used to lower the noise and obtain the optimum parameters. On the basis of the above work, artificial neural network based error back propagation was adopted to identify spectral line of the copper and stainless steel sample successfully. All the results illustrate that the utilization of multiple data processing method for spectral signal processing in LIBS technique can improve the quality of line's analysis and recognition.

Key words: laser-induced breakdown spectroscopy; baseline correction; line's recognition; wavelet transform; artificial neural network

收稿日期: 2014-03-07; 修订日期: 2014-04-01

基金项目: 国家自然科学基金(61177082, 61205074); 北京市自然科学基金(4122063)

作者简介: 杨崇瑞(1988-), 女, 硕士生, 主要从事光谱信号处理方面的研究。Email: 11120179@bjtu.edu.cn

导师简介: 娄淑琴(1965-), 女, 教授, 博士, 主要从事光纤传感、特种光纤方面的研究。Email: shqlou@bjtu.edu.cn

0 引言

激光诱导击穿光谱(Laser Induced Breakdown Spectroscopy, LIBS), 是一种基于原子发射光谱和激光等离子体发射光谱进行样品中元素成份分析的先进光谱技术^[1]。它具有高分辨率、高精度的特性, 不会破坏样品物质结构^[2], 可在线分析液体、固体、气体多种状态样品的优势, 在太空探测、材料成分分析、环境污染检测、文物鉴定、生物医学等各种民用和军事应用领域具有广泛地应用。随着激光和光谱技术的发展, LIBS 探测技术和检测装置的探测水平不断提高, 典型标志是 2012 年名为 ChemCam 的 LIBS 系统作为“好奇号”火星探测器的岩石成份分析仪成功登陆火星。

近年来, LIBS 的研究工作主要集中在其物理和光学理论基础、不同实验样品的影响特性、实验参数优化及最优分析精度等问题上, 较之拉曼光谱^[3]、近红外光谱^[4]等其他成熟的光谱技术, 光谱信号预处理技术有待进一步提高。光谱信号预处理, 一方面可以有助于提升硬件系统的技术指标, 降低系统的复杂度及成本; 另一方面, 还可以解决一些通过硬件优化无法解决的技术问题, 扩展系统在不同环境下的适用性。因此, 引起研究者的关注。对光谱基线校正、光谱降噪及谱线识别等光谱处理技术进行深入的研究, 对提高 LIBS 系统测试样品的谱线分析和识别质量具有重要的意义。然而, 目前适合 LIBS 光谱信号预处理的基线校正方案研究还较少; 在光谱降噪方面, GAO 等将小波变换用于紫外光谱的滤噪, 并把所得结果与移动窗口平均即傅里叶变换技术相比较, 表明了小波变换(WT)优于后两种方法^[5], Alsber 等将小波变换应用于红外光谱的滤噪, 并证实了小波变换在针对多种噪声的良好去噪效果^[6], 国内将小波变换应用于 LIBS 光谱数据去噪, 也实现了较好的性能提升^[7]; 在谱线识别方面, 已有将神经网络成功应用于定量检测^[8-9], 提高了物质定量检测准确度。目前, LIBS 光谱数据处理广泛研究的重点主要局限于以光谱滤波方法降噪, 光谱图像法基线平滑, 以及谱线重叠干扰分析等单一数据处理方法的光谱信号预处理技术, 已不足以有效地提高其谱线的分析和识别质量。

文中重点针对 LIBS 系统的光谱信号预处理技术开展研究, 综合利用多数据处理方法进行 LIBS 技术中光谱信号预处理, 以提高谱线的分析和识别质量。在以多通道光谱仪不同波段光谱特征值点提取和连接的方法进行光谱基线平直, 以及以小波变换算法进行 LIBS 谱线信号去噪等工作的基础上, 使用基于误差反向传播的人工神经网络方法, 实现了纯铜和不锈钢等物质种类的高精度有效识别。研究成果在 LIBS 技术领域具有一定的推广价值。

1 实验装置

文中建立的 LIBS 实验系统如图 1 所示, 系统由激光器、样品台、光谱仪和计算机等组成。其中激光器为小型化的固体脉冲激光器(Big Sky Laser), 输出波长为 1 064 nm, 脉冲宽度为 6 ns, 最大单脉冲能量超过 40 mJ(2~40 mJ 可调), 重复率 1~20 Hz 可调; 多通道光谱仪(HR2000, Ocean Optic)的光谱响应范围为 200~1 100 nm。

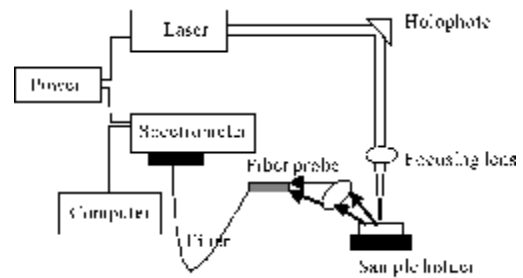


图 1 激光诱导击穿光谱实验系统

Fig.1 Experimental system of LIBS

激光器输出的高能量密度激光脉冲, 经过全反射镜反射, 再通过聚焦透镜聚焦到样品上, 激光聚焦区域的样品受到激光烧蚀, 温度会迅速升高, 固态样品均会超过气化点, 产生瞬态、高能态等离子体。等离子体在膨胀/碰撞的过程中会发射出离子、原子、原子团簇及分子等各种光谱。等离子体发射的广谱光谱经由光纤耦合进多通道光谱仪。计算机根据这些光谱的波长和强度, 对样品所含成份进行确认和量化分析。文中针对纯铜和不锈钢两种固体样品, 重点进行了相应光谱预处理的实验研究。

2 LIBS 谱信号预处理

由于多通道光谱仪的使用, LIBS 光谱信号在各波段由于摄入光量和各通道特性的不同, 在波段间

重叠区域会出现光谱基线不在同一条水平线的问题。基线校正近年来已成为谱线分析的关键技术之一,如近红外光谱分析中,针对不同样品及偏差来源,已有标准归一化(SNV)、相乘性散射校正(MSC)、微分小波变化等^[10]方法应用于基线校正。但适合 LIBS 光谱信号预处理的基线校正方案研究还较少。另外,LIBS 谱信号中存在来源于激光诱导过程中产生的较强的连续背景噪声辐射,光纤探头采集到的噪声,以及模/数、数/模转换等过程带来的附加噪声,这些噪声的存在影响准确确定光谱峰值的位置、谱峰强度等,需要进行处理,以便准确地获得实验数据。对采集到的光谱信号进行消噪优化,以提取有效信号,有助于光谱线位置及峰强度等光谱信息的准确确定,并且进一步提高该技术的应用性能指标以及数据自动分析处理能力。

2.1 分段特征值提取法校正光谱基线

图 2 为纯铜样品(图(a)~(c))和不锈钢样品(图(d)~(f))在 190~500 nm 波段的 LIBS 光谱,其跨越了 HR2000 型多通道光谱仪的 3 个波段,3 个波段因光纤耦合摄入光量和通道特性的差异,基线不在同一条水平线上。

多通道光谱仪输出的光谱信号,尽管会因不同波段耦合摄入光量和通道特性的差异而出现基线的

差异问题,但各波段的基线,尤其是交叠区的基线是应该在同一条直线上。因此,代表不同波段输出光谱信号大概率值的特征值点应该是幅度一致的。基于此,文中首先针对多通道光谱仪不同波段输出的光谱信号进行各自的特征值点的提取,然后将各波段光谱信号的特征值点幅度对齐,进而据此处理整个光谱范围的光谱信号,提高光谱基线的平直度。

首先,将 LIBS 光谱信号数据平均分为 N 份:

$$\text{number} = \frac{\text{wavelength}}{N} \tag{1}$$

式中: wavelength 为 LIBS 光谱信号波长范围; number 为所分数据组数。

再提取各组数据的强度最小值:

$$\min\{A(a_j)\}_j, j=1, 2, 3, \dots, n \tag{2}$$

作为该组数据的特征值点。

然后,把每组数据中的特征值点连接起来,并且使用样条差值对曲线进行拟合,得到 LIBS 光谱基线函数 f_{baseline} 。

最后,基于基线函数 f_{baseline} ,对 LIBS 光谱原始数据 spec_{org} 进行基差校正处理,所得

$$\text{spec}_{\text{correc}} = \text{spec}_{\text{org}} - f_{\text{baseline}} \tag{3}$$

即为经过基线校正处理的光谱。

图 3、图 4 为使用上述方法针对纯铜、不锈钢样品 197~500 nm 波段 LIBS 光谱所进行的基线校正后的光谱,可见取得了很好的基线校正效果。

造成光谱基线不平的背景光谱漂移,在一般情况下都是具有相对稳定趋势的线性或者低频非线性

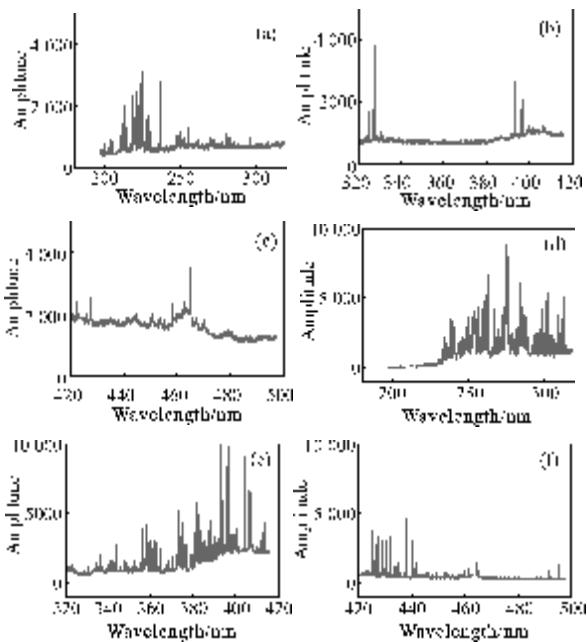


图 2 原始光谱图

Fig.2 Original spectral graph

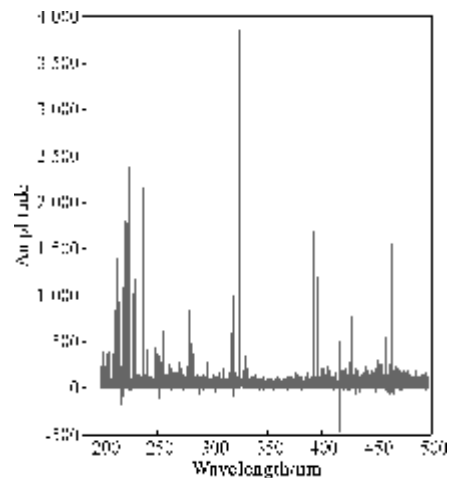


图 3 纯铜样品谱线进行基线校正后的图像

Fig.3 Spectral graph of copper after baseline correction

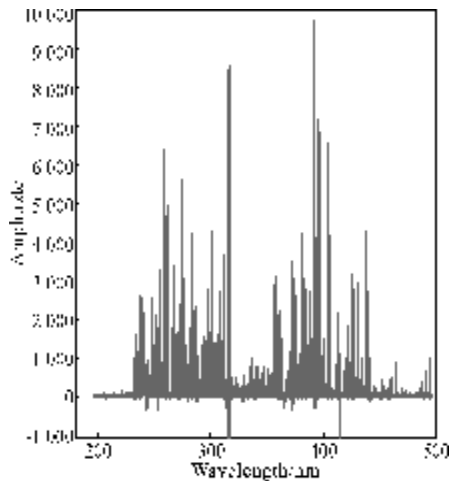


图 4 不锈钢样品谱线进行基线校正后的图像

Fig.4 Spectral graph of stainless steel after baseline correction

的。因此,文中方法对此类基线不平的问题均可进行有效的抑制和校正,而且对原始光谱信号影响很小,是一种有效的 LIBS 光谱信号的预处理技术。

2.2 小波变换算法光谱图像降噪技术

现在已有很多种信号降噪方法,如卡尔曼滤波法^[11],维纳滤波法,减谱法^[12]等。由于光谱信号主要是表现为尖峰和突起的非平稳信号,并且小波变换(Wavelet Transform, WT)算法具有良好的时频局域性、选基灵活性和去相关性等优点,可以区分高频信号和噪声成分,因此成为光谱信号降噪滤波的首选算法。

在小波域内的同一频段上,由于小波变换的时频局域性,光谱中的尖峰被压缩在仅少数的时段上,因此其对应的小波系数幅值较大;而噪声的小波系数是分布在所有时段上的,其小波系数幅值较小,两者的区别为同一频段不同时段上小波系数幅值的不同,从而用阈值法可以实现分离。即先行对含噪信号进行小波分解,获得各个频段的小波系数之后,然后使用相应的阈值算法对其进行处理,根据新的小波系数对小波进行重构,从而获得去除噪声以后的光谱信号。

光谱降噪效果,一般使用峰值信噪比(PSNR)来衡量。

$$\text{PSNR} = 10 \lg \frac{255^2}{E\{[X'(x,y) - X(x,y)]^2\}} \quad (4)$$

式中: $X(x,y)$ 表示原始图像; $X'(x,y)$ 表示去噪后的图像。

此节,仅以纯铜为例,将其已进行过基线校正的光谱信号进行 N 层小波分解,需要选择的参数有小波基、分解层数及阈值计算方法,这些参数对去噪效果的影响如图 5 与表 1。图 5 即为在不同小波基(包括常用的 db 系列, sym 系列,以及 coif 小波基)情况下,随着分解层数的改变而得到的 PSNR 值的变化曲线,其中 PSNR 越小,意味着消除噪声的均方值越大,降噪效果越好。表 1 为不同阈值计算方法和不同小波基情况下的 PSNR 值(单位为 dB, $N=6$)。

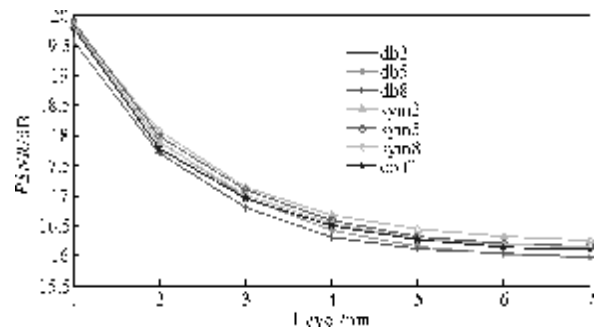


图 5 不同分解层数和不同小波基情况下的 PSNR(dB)值

Fig.5 Value of PSNR with different level N and different wavelets

表 1 不同阈值计算方法和不同小波基情况下的 PSNR 值($N=6$)Tab.1 Value of PSNR with different threshold computer methods and different wavelets ($N=6$)

Method	Wavelet				
	db5	db8	sym5	sym8	coif3
Rigrsure	22.678 4	20.877 1	22.064 8	21.941 8	21.657 1
Sqtwolog	16.033 6	16.035 7	16.201 6	16.179 4	16.148 0
Heursure	22.678 4	20.877 1	22.064 8	21.941 8	21.657 1
Minimaxi	17.442 91	17.356 8	17.636 2	17.590 2	17.673 8

实验结果表明,在同样的分解层数及阈值计算方法的情况下,db 系列的小波基效果较好;在同样小波基及阈值计算方法的情况下,分解层数越大,降噪效果越好,但鉴于在 $N=5$ 之后 PSNR 值随着 N 增大变化并不明显,并且计算也会随之变得更复杂,故一般选择 $N=5$;在小波基及分解层数相同的情况下, sqtwolog 阈值计算方法效果最好。故按照实验结果选取,小波变换算法的最佳参数应为 db5 小波基,分解层数 5,以 sqtwolog 阈值计算方法进行降噪处理,所取得的降噪效果如图 6 所示。

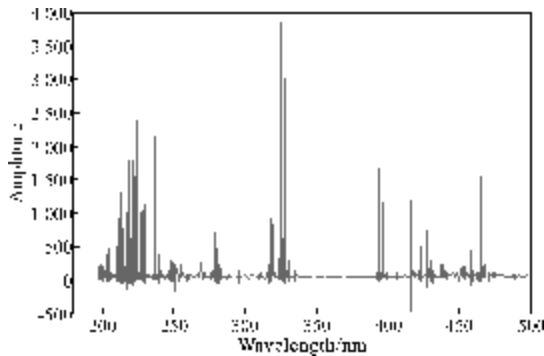


图 6 小波降噪后的光谱图

Fig.6 Spectral graph after decreasing the noise

3 实验及结果分析

目前,在利用 LIBS 进行物质分类时,采用较多的是主成份分析法(PCA)^[13]和神经网络(ANN)两种方法进行光谱数据分析处理。PCA 是种描述变量之间关系的多元分析技术,常用于不同物质的类别分析。LIBS 光谱数据经常会存在谱峰重叠等问题,使用 PCA 分析法需要进行复杂的处理,才能获得较为精确的结果。ANN 是一种模仿动物神经网络行为特征进行分布式并行信息处理的数学模型,具有非线性映射、自适应学习和稳健等特点。因此,文中采用基于梯度下降法的误差反向传播神经网络(BP-ANN)应用于 LIBS 定量分析中。

图 7 分别为 BP-ANN 的原理和结构。如图(a)所示,输入量 x_i 与权值 w_i 相乘并求和,然后减去阈值 b ,将差值通过传递函数转换从而得到输出 n 。图(b)为整个神经的网络结构,共包括 3 层,分别为输入层、输出层和隐层。BP-ANN 法根据网络的实际输出与期望输出之间的误差平方和,通过误差的反向传播,利用梯度下降法迭代调整神经元之间的权值和阈值,当收敛到预设的目标误差时结束训练。

预测分析所用网络结构采用 3 层 BP 神经网络,输入层节点数由输入数据的数量决定,共 1 层隐层,节点数为 20,输出共 2 个节点,60 组光谱数据分为 3 部分,其中 42 组为训练数据,9 组为校验数据,9 组为测试数据。最大迭代误差为 20 次。

文中对纯铜和不锈钢样品在同样环境下各取 30 组数据,将其进行如上预处理过程后,作为训练集输入网络。对 3 层的 BP 神经网络模型进行训练,确定最优的神经网络预测模型。神经网络中以 sigmoid 函数

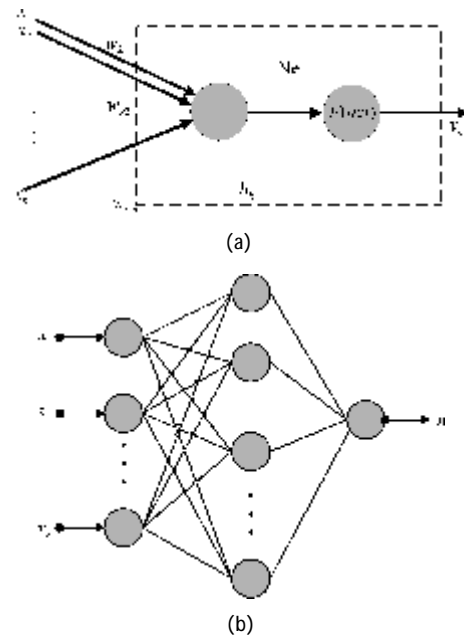


图 7 神经网络原理及结构图

Fig.7 Principle and architecture of artificial neural network

作为连接隐含层和输出层的传递函数,表 2 所示为训练次数对计算时间的影响,并且鉴于纯铜、不锈钢两种样品易于区分,其准确度未随着训练次数发生改变,如表 3 所示,实验中分类预测结果准确度达到了 100%。

表 2 训练次数对运算时间的影响

Tab.2 Impact of training times on computing time

Item	Value			
Training times	200	500	800	1 000
Realization time/s	0.2	0.26	0.3	0.3

表 3 ANN 模型的预测结果

Tab.3 Predicting results of ANN model

Sample kinds	Number of test samples	Correct identified number	Incorrect identified number	Accuracy	Error rate
Copper	30	30	0	100%	0%
Stainless	30	30	0	100%	0%

BP 神经网络的输入层节点数、隐含层节点数以及隐含层层数对网络的性能有较大影响。当输入层的节点数过少时,网络获取的信息量就会大大减少,从而导致网络模型产生缺陷;当输入层节点数过多时,网络训练的工作量就会很大。隐含层层数的增多也会增加网络的复杂度,大大提高权值以及网络的训练时间。所以需要通过反复试验,从而获得适合数据所需的网络参数。

4 结 论

近年来,通过光谱数据的预处理优化方法提高 LIBS 应用的精度已成为一个主要的技术发展方向。文中基于实验室 LIBS 实验装置,对纯铜和不锈钢样品开展了 LIBS 光谱的实验,结合实验开展了检测光谱的预处理技术研究。进行通过实验验证,得出以小波变换算法进行 LIBS 谱线信号去噪的最佳算法参数。基于多通道光谱仪不同波段光谱特征值提取,提出了一种简单易行的多组数据中特征值点连接的方法,有效地提高了 LIBS 光谱信号的基线平直度。在上述工作的基础上,使用基于误差反向传播的人工神经网络方法,实现了纯铜和不锈钢等物质种类的有效识别,研究结果表明,综合利用多数据处理方法进行 LIBS 技术中神经网络方法,实现了纯铜和不锈钢等物质种类的有效识别,研究结果表明,综合利用多数据处理方法进行 LIBS 技术中光谱信号预处理可以有效提高谱线分析和识别的质量,实现了对纯铜和不锈钢样品的分类,有效地提高了 LIBS 光谱法进行物质分类识别的可靠性。

参考文献:

- [1] Kraushaar M R, Noll R, Schmitz H U. Slag analysis with laser-induced breakdown spectroscopy [J]. *Appl Spectrosc*, 2003, 57: 1282-1287.
- [2] Ma Demin, Ma Yanhua, Shu Rong, et al. Analysis of LIBS feasibility for in-situ detection of lunar soil components [J]. *Infrared and Laser Engineering*, 2007, 10(36): 656-658. (in Chinese)
马德敏, 马艳华, 舒嵘, 等. 激光诱导现场探测月壤成分的可行性分析 [J]. *红外与激光工程*, 2007, 10(36): 656-658.
- [3] Chen Chen. Data processing and database upbuilding of Raman spectroscopy in detecting water pollutants [D]. Wuhan: Huazhong Normal University, 2006. (in Chinese)
陈晨. 水体污染物拉曼光谱检测的数据处理与建库技术 [D]. 武汉: 华中师范大学, 2006.
- [4] Zhang Yaodong. Analysis and modeling of NIR spectrum based on intelligent algorithm [D]. Shenyang: Northeast University, 2006. (in Chinese)
张耀东. 基于智能算法的近红外光谱分析与建模 [D]. 沈阳: 东北大学, 2006.
- [5] Hussam A, Alauddin M, Khan A H. Solid phase microextraction: measurement of volatile organic compounds (VOCs) in Dhaka city air pollution [J]. *Journal of Environmental Science and Health Part A-Toxic/Hazardous Substances & Environmental Engineering*, 2002, 37(7): 1223-1239.
- [6] Acharya T, Chakrabarti C A. A survey on lifting-based discrete wavelet transform [J]. *Journal of VLSI Signal Processing*, 2006, 42(3): 321-339.
- [7] Jiang Meiyu. The application of wavelet transform in Laser-Induced Breakdown coal spectroscopy analysis [D]. Guangzhou: South China University of Technology, 2010. (in Chinese)
蒋梅域. 小波分析应用于激光诱导击穿煤质光谱分析 [D]. 广州: 华南理工大学, 2010.
- [8] Edilene C Ferreira, Debora M B P Milori, Ednaldo J Ferreira et al. Artificial neural networks for Cu quantitative determination in soil using a portable laser induced breakdown spectroscopy system [J]. *Spectrochimica Acta Part B*, 2008, 63(10): 1216-1220.
- [9] Shen Qinmei, Zhou Weidong, Li Kexue. Quantitative elemental analysis using Laser Induced Breakdown Spectroscopy and Neuro-Genetic approach [J]. *Chinese Journal of Lasers*, 2011, 38(3): 0315001-0315006. (in Chinese)
沈沁梅, 周卫东, 李科学. 基于遗传神经网络的激光诱导击穿光谱元素定量分析技术 [J]. *中国激光*, 2011, 38(3): 0315001-0315006.
- [10] Sun Yi, Du Zhenhui, Yin Xin, et al. Investigation on baseline correction in pollutant monitoring on-line by NIR spectroscopy [J]. *Spectroscopy and Spectral Analysis*, 2008, 28(10): 2282-2284. (in Chinese)
孙毅, 杜振辉, 尹新, 等. 近红外光谱气体在线分析中基线校正方法的研究 [J]. *光谱学与光谱分析*, 2008, 28(10): 2282-2284.
- [11] Chen Guanrong, Wang Jianrong, Leang Shieh S. Interval Kalman filtering [J]. *IEEE Transactions on Aerospace and Electronic Systems*, 1997, 33(1): 250-259.
- [12] Dnoho D L. Denoising by soft-thresholding [J]. *IEEE Transaction IT*, 1995, 41(3): 613-627.
- [13] Huang Jisong, Chen Qiaoling, Zhou Weidong. Laser induced breakdown spectroscopy for the determination of Cr and Sr in soil [J]. *Spectrosc Spect Anal*, 2009, 29(11): 3126-3129. (in Chinese)