

西林瓶内氧气含量在线检测的基线快速校正方法

朱高峰^{1*}, 朱红求², 钱灏², 朱剑平³

¹湖南人文科技学院信息学院, 湖南 娄底 417000;

²中南大学信息科学与工程学院, 湖南 长沙 410083;

³楚天科技股份有限公司研发中心, 湖南 长沙 410600

摘要 构建了一个开放单光路短光程激光检测平台,提出了一种基于线型的 Levenberg-Marquardt(L-M)算法,以洛伦兹线型和高斯线型的二阶导数作为 L-M 算法的拟合函数,对氧气体积分数为 0%~21% 的样本进行校正。结果表明:在基于洛伦兹线型的 L-M 算法校正中,氧气含量-二次谐波峰值的拟合系数最高,为 0.9995;以洛伦兹线型为优化依据,对含氧气体积分数为 1% 的样品进行了多次测量,校正处理前后预测的氧气体积分数最大偏差分别为 0.38% 和 0.22%,预测的均方根误差分别为 0.25% 和 0.16%。

关键词 光谱学;基线校正;Levenberg-Marquardt 算法;氧气含量;波长调制光谱

中图分类号 O433

文献标识码 A

doi: 10.3788/CJL201946.0211001

Fast Baseline Correction Method for Online Detection of Oxygen Content in Vials

Zhu Gaofeng^{1*}, Zhu Hongqiu², Qian Hao², Zhu Jianping³

¹School of Information, Hunan University of Humanities, Science and Technology,
Loudi, Hunan 417000, China;

²School of Information Science and Engineering, Central South University, Changsha, Hunan 410083, China;

³Research Center, Truking Technology Company Limited, Changsha, Hunan 410600, China

Abstract A laser detection platform with an open single path and short optical distance is constructed, and a Levenberg-Marquardt (L-M) algorithm based on linetype is proposed. The method uses second order derivative of Gaussian linetype and Lorentz linetype as the fitting function of L-M algorithm. The samples with oxygen volume fraction of 0%-21% are corrected. The results show that the correlation coefficients of oxygen content and peak is the highest in L-M algorithm correction based on Lorentz linetype (0.9995). The sample with oxygen volume fraction of 1% is measured several times on the basis of Lorentz linetype. Compared with the measured results before and after the correction, the predicted maximum deviations of oxygen volume fraction are 0.38% and 0.22%, respectively, and the root mean square errors of prediction are 0.25% and 0.16%, respectively.

Key words spectroscopy; baseline correction; Levenberg-Marquardt algorithm; oxygen content; wavelength modulation spectroscopy

OCIS codes 300.1030; 300.6360; 300.6380

1 引 言

为确保西林瓶内药品的稳定性和无菌性,制药工艺中常采用抽真空或充氮气的方法使药品与外界隔绝。对西林瓶密封完整性进行检测的传统方法(微生物挑战法、色水法等)都属于破坏性抽样测试。

我国 2015 版无菌药品生产质量管理规范(GMP)检查指南规定,对于密封的产品须进行定期密封性测试和趋势分析。《欧盟药品 GMP 指南》明确规定玻璃或塑料安瓿容器应当进行 100% 整体性测试,USP (United States Pharmacopoeia, 美国药典) 1207.2 中规定应采用激光技术来定量测试玻璃瓶

收稿日期: 2018-09-05; 修回日期: 2018-10-11; 录用日期: 2018-10-24

基金项目: 国家自然科学基金(61773403, 61873282)、湖南省教育厅科学研究项目(17A109)、湖南省技术创新引导计划(2018ZK4053)

* E-mail: hnrkzgf@163.com

体中氧气、水蒸气等参数。

激光技术中的波长调制光谱(WMS)技术是指受调制的激光器输出一定范围波长的激光来扫描气体的某条吸收谱线,通过检测其透射光强来识别气体的成分及性质,可以实现非破坏性原位快速、高精度检测,被广泛应用于气体体积分数和温度的检测^[1-5]。本课题组前期已成功构建了开放的单光路、短光程 WMS 平台^[6-8],用来检测密封西林瓶内氧气的体积分数(含量),通过优化系统参数、光路结构、数据处理算法等,在很大程度上抑制了系统噪声的影响,但析取的二次谐波信号仍存在基线偏移。系统仪器噪声、系统主要参数选择优化和数据处理算法效果、玻璃瓶壁引起的残留光学噪声、随机开放环境的空气背景干扰等,导致氧气含量-二次谐波峰值的对应关系存在不同程度的非线性,从而影响氧气含量测量的精度和稳定性。

目前,许多人对光谱技术中的基线校正进行了研究:张锐等^[9]提出了基于无吸收谱线区域检测谐波的背景扣除和非线性校正方法,该方法可以减小系统噪声的干扰,但仍存在剩余基线的影响;Werle等^[10]提出了先采集背景气体的谱线信号,再采集目标气体谱线吸收信号进行扣除的方法;Persson等^[11]和 Masiyano 等^[12]通过改进光路或气室结构进行非线性处理,目的是减小基线的影响,但该方法只适用于有气体吸收池的系统。在数据预处理过程中进行基线校正的方法也较多,目前主要有小波变换^[13-14]、正交信号处理^[15-16]等,但运算复杂,实时性不强。Levenberg-Marquardt (L-M)拟合算法同时具有梯度法和牛顿法的优点,且收敛速度快。本文利用洛伦兹线型和高斯线型的二阶导数作为吸收光谱的二次谐波表达式,以峰值高度、半峰宽度和峰值位置作为优化参数,采用 L-M 算法自适应调整阻尼因子来实现二次谐波的基线快速校正。

2 L-M 算法的原理

根据谱线展宽机理可知,洛伦兹线型和高斯线型描述的是直接吸收光谱的形状。因此,可以对选定线型求各阶导数,作为校正各次实测谐波信号的标准函数,所需校正信号的关键特征变量为峰值高度、半峰宽度和峰值位置。

在 WMS 系统中,采用 L-M 算法解决氧气含量-二次谐波峰值的非线性问题,其优化目标可表示为

$$E(x) = \min_{x \in \mathbf{R}^n} \|f(x) - M(x)\|^2, \quad (1)$$

式中: \mathbf{R}^n 为实数域 \mathbf{R} 上的 n 维标准欧氏空间; x 为对应拟合的二次谐波特征值变量,其初始特征变量 x_0 可由实测的二次谐波波形计算得到; $f(x)$ 为标准的参考线型表达值; $M(x)$ 为实际测量的二次谐波值。通过反复迭代可使目标函数 $E(x)$ 最小,此时的 x 值即为最终的迭代参数。

算法中的迭代表式为 $x_{k+1} = x_k + (\mathbf{H} + \mu\mathbf{I})^{-1}\mathbf{J}^T f$,其中: k 为迭代次数; \mathbf{J} 为 x 的 Jacobi 矩阵; \mathbf{I} 为单位矩阵;Hessian 矩阵 $\mathbf{H} = \mathbf{J}(x)^T \mathbf{J}(x)$; $(\mathbf{H} + \mu\mathbf{I})^{-1}\mathbf{J}^T f$ 为迭代步长,记为 h_{lm} ; μ 为阻尼因子,它能同时影响步长和搜索方向,这使得 L-M 算法无需进行线性搜索,大大减少了迭代次数。为了给出阻尼变化的依据,定义第 k 次迭代中的实际下降量 ΔE_k 为

$$\Delta E_k = E(x_k) - E(x_{k-1} + h_{lm}). \quad (2)$$

利用实际下降量的极性来实现步长的调节方向;当 $\Delta E_k < 0$ 时,说明 h_{lm} 不合适,需要缩小信赖域半径重新求解;当 $\Delta E_k > 0$ 时,说明模型与目标函数在信赖域的范围内有较好的相似度,可以适当增大信赖域的半径。当然,在实际工程中,可以通过对阻尼因子进行缩放来选取适当的信赖域。

理论上,当搜索优化找到全局极小值时才终止迭代,但本系统要实现实时在线快速响应,因此适当修改了终止迭代的条件,采用双重关口来判断迭代循环是否结束,这样可以显著减少迭代次数和计算时间。具体而言,当迭代点 x 的变化很小,低于某一阈值 ϵ 时,可以终止迭代步骤,表示为:

$$\|x_{\text{new}} - x\| \leq \epsilon, \quad (3)$$

式中: x_{new} 为新的迭代点。

同时,在迭代时间允许的范围内,设定最大迭代次数 k_{max} ,当迭代点变化未达到最小阈值但已达到最大迭代次数时也将终止算法处理过程。具体的 L-M 算法实现步骤如图 1 所示,经过不断调整,计算得到了曲线的最优化参数,完成了检测信号的基线快速校正。

3 实验部分

3.1 实验装置

基于 WMS 技术的开放单光路短光程实验装置如图 2 所示,西林瓶放置在自由空间中,低频的锯齿波扫描信号由信号发生器 (AFG3022C) 产生,高频的正弦波调制信号由锁相器 (Model 7280) 产生,两者叠加后被送往激光控制器 (Model 6100),驱动中心波长为 760 nm 的分布式反馈 (DFB) 激光器发射

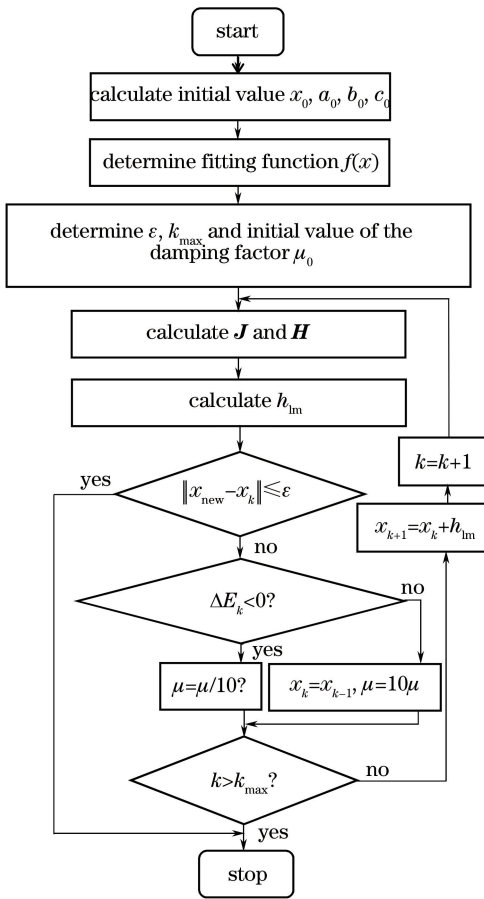


图 1 L-M 算法流程图

Fig. 1 Flow chart of L-M algorithm

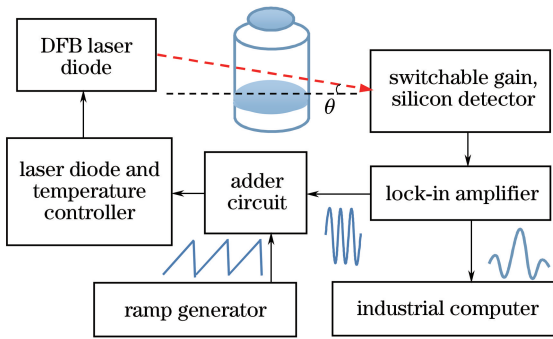


图 2 WMS 检测系统结构图

Fig. 2 Schematic of detection system based on WMS

激光,受调制的激光经过准直透镜组件,以一定角度穿透被测西林瓶的顶空部分(竖直方向相对于瓶壁的人射角 $\theta \approx 11^\circ$,能有效抑制玻璃瓶壁引起的干涉噪声的影响),透射光经聚焦透镜聚焦在光电探测器(PDA36A-EC)上,转换成电信号后被送往锁相器相敏检波,输出的二次谐波数据由通用接口总线接口卡(PCI-1671UP)送往工控机进行处理。

3.2 实验步骤

在气体压强为 1.01×10^5 Pa、温度为 296 K 的条

件下,采用氮气作为平衡气体,采集不同氧气含量的西林瓶(底部直径为 0.022 m)作为初始建模样本,选择优化好实验相关参数后,进行如下数据预处理:

步骤 1):对每个样本的二次谐波信号连续采集 20 个周期,进行粗大误差剔除及平均处理,即将 20 列采样数据去掉最大值和最小值后再进行平均,抑制随机噪声;

步骤 2):用改进的 Savitzky-Golay 滤波算法对数据进行平滑处理^[8],抑制系统周期噪声;

步骤 3):进行 L-M 算法的基线校正处理,使各谐波信号在同一水平基线上;

步骤 4):提取校正后的相应的二次谐波信号峰值;

步骤 5):扣除自由空间背景,将步骤 4)获得的谐波峰值减去标定中 0% 样本(包括自由空间中的氧气吸收)的谐波峰值,以消除自由空间的背景影响,得到瓶内实际氧气的吸收谱线特征峰值。

对不同氧气含量的样本(各 15 支)重复上述步骤进行处理,获得对应的 15 个特征峰值,进行算术平均后再进行最小二乘拟合,建立氧气含量-二次谐波峰值反演模型。

4 结果与讨论

4.1 基于两种线型校正的波形对比

图 3 表示氧气体积分数分别为 0%、1%、2%、4% 的西林瓶的谐波检测结果,该检测信号已通过粗大误差剔除、多周期平均、平滑滤波等常规数据处理(未进行基线校正及背景扣除),但基线起伏仍较严重,且各基线水平与西林瓶内氧气含量无直接规律,导致氧气含量-二次谐波峰值对应关系呈现不同程度的非线性关系,这必然会影响到氧气含量反演的精度。

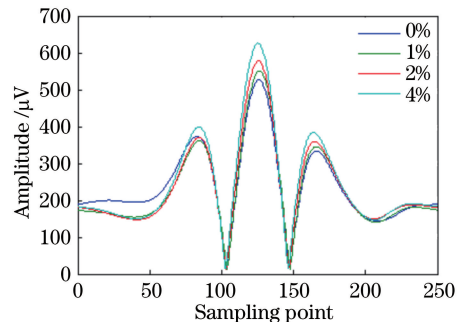


图 3 常规数据处理后的二次谐波波形

Fig. 3 Second harmonic signal waveforms after conventional data processing

分别利用两种线型对氧气体积分数为 0%、1%、2%、4% 的西林瓶检测信号进行校正,结果如图 4、5 所示,其中图 4 是利用洛伦兹线型进行校正的

结果,图 5 是利用高斯线型进行校正的结果。基于两种线型对不同氧气含量的谐波进行校正计算,得到的初始参数和优化后的参数如表 1 所示。

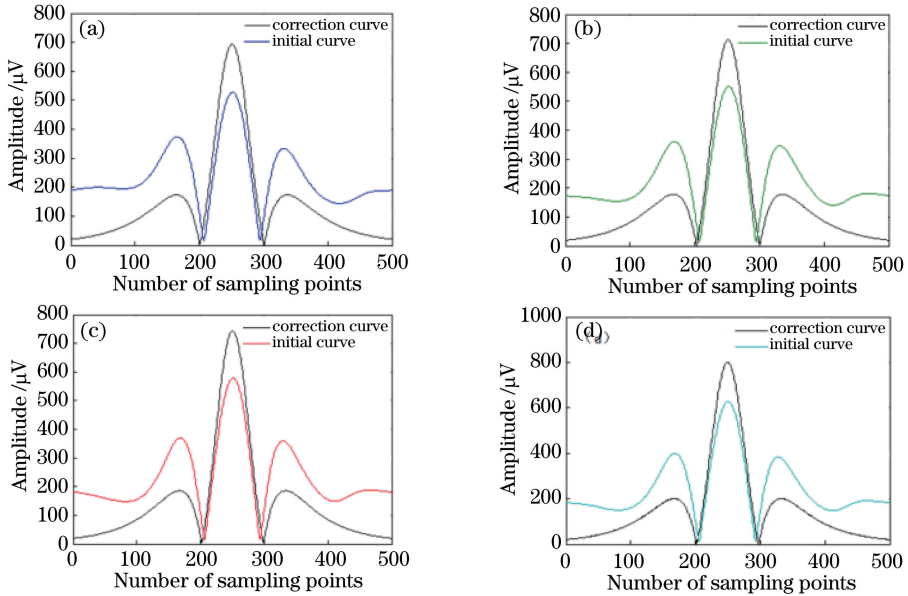


图 4 不同氧气含量下基于洛伦兹线型的二次谐波校正结果。(a) 0%;(b) 1%;(c) 2%;(d) 4%
Fig. 4 Second harmonic correction results of different oxygen contents based on Lorentz linetype.
(a) 0%; (b) 1%; (c) 2%; (d) 4%

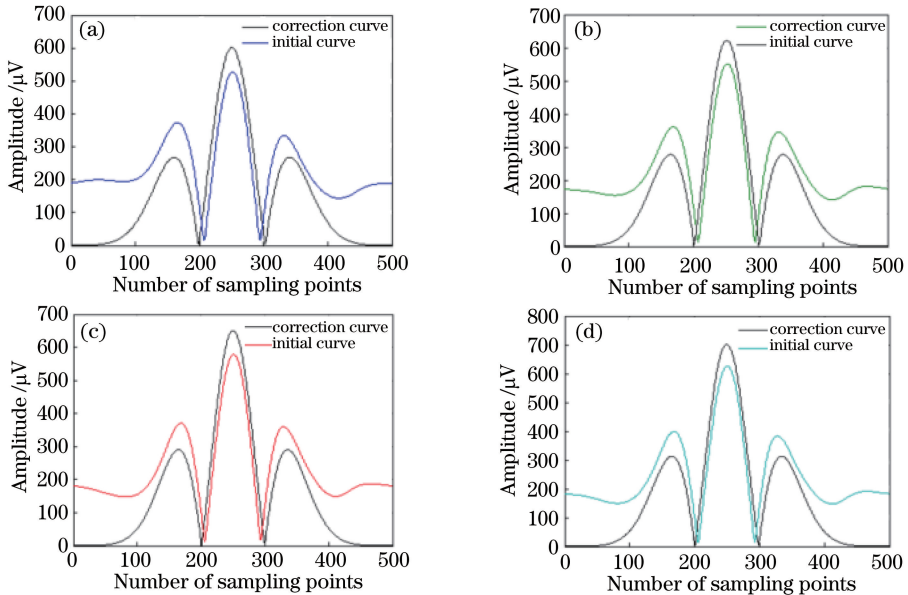


图 5 不同氧气含量下基于高斯线型的二次谐波校正结果。(a) 0%;(b) 1%;(c) 2%;(d) 4%
Fig. 5 Second harmonic correction results of different oxygen contents based on Gaussian linetype.
(a) 0%; (b) 1%; (c) 2%; (d) 4%

由图 4、5 可以看出,用洛伦兹线型或高斯线型对检测信号进行校正后,谐波信号的基线接近零坐标轴,且从一定程度上抑制了剩余幅度调制噪声引起的二次谐波波形左右不对称的影响,波形的失真得到显著改善。这说明利用基于两种线型的 L-M

算法进行非线性处理能够对测量曲线进行校正,在保留原有信号的同时抑制了系统残余光学干涉噪声、仪器及环境噪声引起的基线漂移现象。对比数据可以发现,校正前后的峰位不变,且峰值均有部分提升,但这不会影响实际的氧气含量测量结果,因为

表 1 基于两种线型校正的参数

Table 1 Parameters based on two kinds of linetype

Linetype	Volume fraction of oxygen /%	Initial parameter a_0, b_0, c_0	Corrected parameter $a_{opt}, b_{opt}, c_{opt}$
Gaussian linetype	0	$1.26 \times 10^6, 252.00, 68.17$	$1.62 \times 10^6, 250.61, 73.50$
	1	$1.24 \times 10^6, 253.00, 66.95$	$1.56 \times 10^6, 251.49, 70.83$
	2	$1.25 \times 10^6, 251.00, 65.73$	$1.58 \times 10^6, 250.29, 69.68$
	4	$1.35 \times 10^6, 251.00, 65.73$	$1.57 \times 10^6, 249.50, 69.02$
Lorentz linetype	0	$1.28 \times 10^{10}, 252.00, 83.50$	$1.91 \times 10^{10}, 251.21, 86.11$
	1	$1.25 \times 10^{10}, 253.00, 82.00$	$1.79 \times 10^{10}, 251.72, 84.16$
	2	$1.21 \times 10^{10}, 251.00, 80.50$	$1.79 \times 10^{10}, 250.38, 83.30$
	4	$1.31 \times 10^{10}, 251.00, 80.50$	$1.85 \times 10^{10}, 249.68, 82.46$

氧气含量-二次谐波峰值反演模型中的样本也都进行了基于线型的非线性处理,波形进行了整体趋势校正,所有样本峰值都得到了不同程度的提升,能改善原来氧气含量-二次谐波峰值的非线性关系。

4.2 氧气含量-二次谐波峰值拟合的对比

理论上,氧气含量与二次谐波的峰值应呈线性关系,针对不同氧气含量的西林瓶(氧气体积分数分别为 0%、1%、2%、4%、8%、12%、15%、21% 的样品各 15 支),进行上述数据预处理和多个特征峰值平均后,进行最小二乘拟合,建立氧气含量-二次谐波峰值反演模型。将数据进行最小二乘拟合,设 $y = kx + b$,其中 x 为西林瓶内氧气的体积分数, y 为对应的二次谐波信号峰值强度, k, b 为回归系数。非线性校正前后的氧气含量-二次谐波峰值拟合关系如图 6 所示。

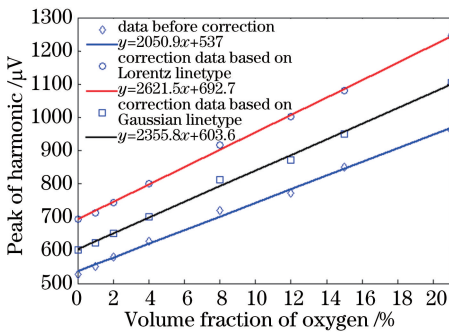


图 6 氧气含量-二次谐波峰值的拟合结果

Fig. 6 Fitting results of oxygen content and peak value of second harmonic

求得校正前及基于洛伦兹线型、高斯线型校正后的相关系数分别为 0.9978、0.9995、0.9983。可以发现,基于高斯线型校正的相关系数比校正前的相关系数略大,基于洛伦兹线型校正的效果最好,相关系数比校正前提高了 0.17%。这表明,在常温常压下,分子的运动较为剧烈,此时碰撞展宽更为突出,洛伦兹线型占吸收谱线线型的主导地位。基于洛伦兹线型校正后,所测西林瓶内氧气含量与二次谐波

信号峰值呈更好的线性关系,这将有助于更准确地预测氧气含量。

4.3 基于洛伦兹线型校正后的检测结果

以洛伦兹线型为优化依据对氧气体积分数为 1% 的西林瓶进行多次测量,对比基线校正处理前后的实测结果(图 7),可以看出:校正处理后,氧气体积分数预测结果的波动显著减小,预测的最大偏差分别为 0.38% 和 0.22%,同时校正后的预测结果分布更加集中,预测的均方根误差分别为 0.25% 和 0.16%,验证了本方法能有效提高系统测量的精度和稳定性。

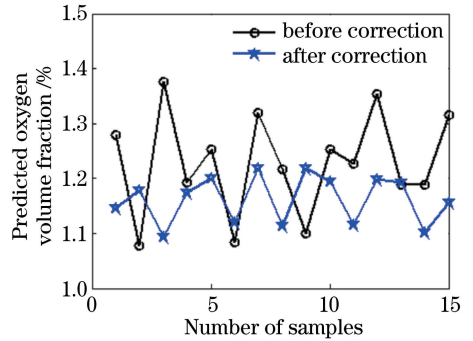


图 7 校正前后的氧气含量分布

Fig. 7 Oxygen content distribution before and after correction

5 结 论

利用 WMS 技术对 760 nm 附近的氧气吸收谱线进行扫描,析取玻璃瓶内氧气所对应的二次谐波信号峰值进行浓度反演。给出了 L-M 算法和数据预处理步骤,选取以高斯线型和洛伦兹线型的二阶导数作为校正的基准函数,论证了以洛伦兹线型为优化依据能有效提高系统测量的精度和稳定性,为 WMS 技术在线检测西林瓶内氧气含量提供了一种快速有效的基线校正方法。同时,该方法还在一定程度上抑制了剩余幅度调制引起的二次谐波波形左右不对称的影响,下一步工作将把波形左右严格对

称作为校正的优化指标之一,进一步完善数据处理算法,克服奇次谐波的影响。

参 考 文 献

- [1] Bolshov M A, Kuritsyn Y A, Romanovskii Y V. Tunable diode laser spectroscopy as a technique for combustion diagnostics[J]. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 2015, 106: 45-66.
- [2] Neethu S, Verma R, Kamble S S, *et al.* Validation of wavelength modulation spectroscopy techniques for oxygen concentration measurement[J]. *Sensors and Actuators B: Chemical*, 2014, 192: 70-76.
- [3] Nie W, Kan R F, Yang C G, *et al.* Research progress on the application of tunable diode laser absorption spectroscopy [J]. *Chinese Journal of Lasers*, 2018, 45(9): 0911001.
 聂伟, 阚瑞峰, 杨晨光, 等. 可调谐二极管激光吸收光谱技术的应用研究进展[J]. *中国激光*, 2018, 45(9): 0911001.
- [4] Wang M M, Dai W G, Yang H N, *et al.* Leakage detection of vials based on tunable diode laser absorption spectroscopy[J]. *Laser & Optoelectronics Progress*, 2017, 54(8): 083004.
 王明明, 戴伟国, 杨荟楠, 等. 基于可调谐半导体激光吸收光谱技术的西林瓶检漏方法[J]. *激光与光电子学进展*, 2017, 54(8): 083004.
- [5] Goldenstein C S, Almodóvar C A, Jeffries J B, *et al.* High-bandwidth scanned-wavelength-modulation spectroscopy sensors for temperature and H₂O in a rotating detonation engine[J]. *Measurement Science and Technology*, 2014, 25(10): 105104.
- [6] Zhu G F, Zhu H Q, Yang C H, *et al.* Parameter optimization of a short open optical path's oxygen concentration detection system based on WMS[J]. *Chinese Journal of Electronics*, 2017, 26(4): 720-724.
- [7] Zhu G F, Hu X, Zhu H Q, *et al.* The multi-beam interference suppression for measuring penicillin vial's oxygen concentration based on tunable diode laser absorption spectroscopy [J]. *Spectroscopy and Spectral Analysis*, 2018, 38(2): 372-376.
 朱高峰, 胡鑫, 朱红求, 等. 基于 TDLAS 检测西林瓶内氧气浓度的多光束干涉抑制方法[J]. *光谱学与光谱分析*, 2018, 38(2): 372-376.
- [8] Zhu G F, Zhu H Q, Yang C H, *et al.* Improved Savitzky-Golay filtering algorithm for measuring a pharmaceutical vial's oxygen content based on wavelength modulation spectroscopy[J]. *Journal of Optical Technology*, 2017, 84(5): 355-359.
- [9] Zhang R, Zhao X H, Hu Y J, *et al.* Background elimination method based on harmonic detection in no absorption spectral region [J]. *Acta Optica Sinica*, 2013, 33(4): 0430006.
 张锐, 赵学珏, 胡雅君, 等. 基于无吸收谱线区域检测谐波的背景消除方法[J]. *光学学报*, 2013, 33(4): 0430006.
- [10] Werle P W, Mazzinghi P, D'Amato F, *et al.* Signal processing and calibration procedures for *in situ* diode-laser absorption spectroscopy [J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2004, 60(8/9): 1685-1705.
- [11] Persson L, Andersson F, Andersson M, *et al.* Approach to optical interference fringes reduction in diode laser absorption spectroscopy [J]. *Applied Physics B*, 2007, 87(3): 523-530.
- [12] Masiyano D, Hodgkinson J, Tatam R P. Gas cells for tunable diode laser absorption spectroscopy employing optical diffusers. Part 2: integrating spheres[J]. *Applied Physics B*, 2010, 100(2): 303-312.
- [13] Guo Y, Zhao X H, Zhang R, *et al.* The noise filtering and baseline correction for harmonic spectrum based on wavelet transform [J]. *Spectroscopy and Spectral Analysis*, 2013, 33(8): 2172-2176.
 郭媛, 赵学珏, 张锐, 等. 小波变换应用于谐波谱线的噪声滤除与基线校正[J]. *光谱学与光谱分析*, 2013, 33(8): 2172-2176.
- [14] Hu Y G, Zhao Z Y, Wang G. Baseline correction and background elimination using wavelet transforms[J]. *Journal of Huazhong University of Science and Technology (Natural Science Edition)*, 2011, 39(6): 36-40.
 胡耀垓, 赵正予, 王刚. 基于小波的光谱信号基线校正和背景扣除[J]. *华中科技大学学报(自然科学版)*, 2011, 39(6): 36-40.
- [15] Fu W H, Hei Y Q, Li X H. UBSS and blind parameters estimation algorithms for synchronous orthogonal FH signals [J]. *Journal of Systems Engineering and Electronics*, 2014, 25(6): 911-920.
- [16] Liu X, Han L J, Yang Z L, *et al.* Application of orthogonal signal correction for near infrared spectroscopic calibration transfer of crude protein of straw silage [J]. *Chinese Journal of Analytical Chemistry*, 2012, 40(4): 596-601.
 刘贤, 韩鲁佳, 杨增玲, 等. 基于正交信号校正的秸秆青贮饲料粗蛋白近红外分析模型传递方法[J]. *分析化学*, 2012, 40(4): 596-601.