

宽谱调谐二极管激光吸收光谱的数据预处理研究

高楠¹ 杜振辉^{1*} 齐汝宾¹ 李金义¹ 周涛² 周魁² 汪曩¹

(¹ 天津大学精密仪器与光电子工程学院, 天津 300072)
² 天津津航技术物理研究所, 天津 300308)

摘要 二极管激光吸收光谱以温度扫描实现快速宽谱调谐,其原始数据不便于传统的处理及运算,适当的预处理是进行后续数据处理的基础。在分析了温度调谐数据特征的基础上,利用 Savitzky-Golay 滤波器去除了波长序列中温度检测噪声的干扰,选择单调的波长区间对信号幅值进行三次样条插值,以固定的波长点对应采样点,得到一组位置固定的幅值序列用于后期处理。此方法将两列检测数据合并为一列,相同的起始点和固定的波长位置不仅将多周期平均后的信噪比提升为原始数据的 6 倍,也使多组数据的比较及评价成为可能,一致的波长分辨率保证了线形检测的精确度,减小了吸收峰的频带宽度,有助于后续的处理、运算以及评价。

关键词 光谱学;调谐二极管激光吸收光谱;数据预处理;Savitzky-Golay 滤波器;三次样条插值

中图分类号 O433.4 **文献标识码** A **doi**: 10.3788/AOS201232.0430004

Data Preprocessing of Broad-Spectrum Tunable-Diode-Laser Absorption Spectroscopy

Gao Nan¹ Du Zhenhui¹ Qi Rubin¹ Li Jinyi¹ Zhou Tao² Zhou Kui² Wang Yan¹

(¹ College of Precision Instruments and Opto-Electronics Engineering, Tianjin University, Tianjin 300072, China)
² Tianjin Jinhang Institute of Technical Physics, Tianjin 300308, China)

Abstract The diode-laser absorption spectroscopy has been tuned with fast speed and wide wavelength range by temperature scanning. The system original data does not suit for traditional processing and operation, therefore the proper data preprocessing is the basis of the follow-up operation. The original data is processed according to the characteristics of temperature tuning signal. The temperature detection noise in wavelength sequence is removed by Savitzky-Golay filter, then cubic spline interpolation is applied to obtain amplitude sequence with fixed wavelength position for follow-up operation. The result sequence, which has both information of amplitude and wavelength as the same starting point and fixed wavelength position, not only has 6 times higher signal-to-noise ratio than that of original signal by multi periods averaging, but also makes the comparison and evaluation between different signals possible. The consistent wavelength resolution ensures the precision of lineshape detection and decreases the frequency bandwidth of absorption peaks. The experimental results indicate that the data preprocessing provided is beneficial to follow-up operation, calculation and evaluation in such systems.

Key words spectroscopy; tunable-diode-laser absorption spectroscopy; data preprocessing; Savitzky-Golay filter; cubic spline interpolation

OCIS codes 300.6170; 300.6260; 200.4560

1 引言

调谐二极管激光吸收光谱(TDLAS)技术自 20

世纪 70 年代由 Hinkley 等^[1,2]提出以来,由于其高灵敏、高分辨和快速响应等特点,伴随着可调谐半导

收稿日期: 2011-10-26; 收到修改稿日期: 2011-11-14

基金项目: 国家 863 计划(2006AA06Z410)、天津市自然科学基金(10JCZDJC22700)和精密测试技术及仪器国家重点实验室探索课题(PILT1107)资助课题。

作者简介: 高楠(1982—),男,博士研究生,主要从事光谱技术及仪器等方面的研究。E-mail: dimflame@tju.edu.cn

导师简介: 汪曩(1955—),男,教授,博士生导师,主要从事分析仪器方面的研究。E-mail: wangyan@tju.edu.cn

* 通信联系人。E-mail: duzhenhui@tju.edu.cn

体激光器的发展,近年来已在工业及环境检测等领域得到广泛应用。传统的 TDLAS 技术利用注入电流的变化实现激光器波长的扫描^[3],由于实现简单且具有很高的调谐速度,此方法目前已成熟应用于各种气体的检测中^[4~6]。由于气体检测中常用的分布反馈(DFB)激光器存在安全电流阈值,电流调谐技术的波长扫描范围通常在 0.2~0.5 nm 之间,传统的 TDLAS 技术一般只针对一种气体的单一吸收峰进行扫描^[7],不仅降低了激光器的使用效率,也带来了实际应用中的局限性,主要反映在气体吸收峰位置寻找以及多吸收峰检测上。电流调谐技术可用来同时检测谱线位置非常接近的多种物质,如 CO₂ 和 CO,但每种气体均只有一个吸收峰被检测,此方法也可在温度测量中检测同种气体的相邻吸收峰^[8]。在这种情况下,相邻的吸收峰在激光器线宽较宽或气体压强较高时可能产生重叠,从而影响谱线特征提取过程,此时就需要更宽的调谐范围覆盖更多的吸收线,不仅提高被测物吸收峰选择的灵活度,利用多峰计算的方法提高系统的抗干扰能力,同时降低系统对激光器线宽和检测压强的要求。

激光器扫描范围过窄可通过设定工作温度解决,激光器的发射波长由注入电流和工作温度同时决定,利用这一特性,可使系统先设定多个工作温度点,并分别测量得到粗略的吸收谱线,在吸收较强的峰值位置固定工作温度进行电流调谐检测^[9],或将不同波长位置的谱线拼接得到较宽的吸收谱线进行多吸收峰检测^[10]。上述方法中,需要将工作温度稳定在固定值后进行电流调谐,在波长稳定点需要较高精度时,这一步骤通常需要较长时间,并且操作过程比较复杂,不利于快速实时检测。在本课题组之前的工作中,为了实现激光器的快速宽谱调谐,可对激光器中的热电制冷器件(TEC)施加驱动电流,使其工作温度在一定范围内扫描,从而实现其发射波长的宽谱扫描,DFB 激光器的温度调谐范围一般可达 4 nm,这使得系统的多种成分辨识及多参数检测成为可能^[11]。在上述需求下,温度调谐系统对检测信号中吸收峰位置以及吸收线形均有较高的精度要求,因而检测信号的波长对应问题将变得很重要,传统的电流调谐技术通常只检测气体吸收峰的幅值而非吸收位置,很少关注波长对应问题^[12]。

由于调谐方式的不同,电流调谐和温度调谐的检测信号存在差异,主要体现在波长对应过程上。电流调谐信号只有一列幅值,其采样点位置与扫描幅度对应,从而可以直接换算成波长;而温度调谐信

号由幅值序列和经过负温度系数热敏电阻(NTC)输出幅值换算得到的波长序列组成,幅值和波长位置一一对应。由两列数据组成的温度调谐信号,在读取和存储过程中比电流调谐信号更复杂,由于各周期信号的起始点并不重合,在后续处理过程中如粗大误差剔除及大量数据评估系统稳定性时多有不便,而非线性的波长采样率使得一段数据前后的波长精度不一致,影响吸收峰定位及线形计算精度。基于上述问题,有必要对温度调谐系统的检测信号进行预处理,使其更接近电流调谐信号,以便用已有的方法进行后续的处理及运算。本文首先介绍了温度调谐的原理及过程,在分析了温度调谐数据的特征后,将波长序列中 NTC 检测噪声滤波去除,在单调的波长区间内对幅值进行插值,得到一组位置固定的幅值序列用于后期处理,并将预处理结果与电流调谐信号进行比较和评价。

2 TDLAS 温度调谐技术数据特征

通常,光谱分析中会采用多次数据平均的方法去除噪声以提高检测极限,由于 TDLAS 的采集信号只有幅值和时间信息而不含波长信息,光谱吸收峰位置和线形受激光的波长扫描范围影响。在电流调谐技术中,激光器注入电流锯齿波的重复性好,且注入电流对波长的影响小,电流调谐速率约为 0.01 nm/mA,可通过测定激光器调谐特性曲线并加以拟合得到^[13],如图 1 所示,其中注入电流为偏置电流叠加扫描信号幅值。系统工作时,锯齿波扫描信号被同时采集,以最低点作为起始点截取完整周期的信号,在存储和读取时可以将多周期信号对齐存为一个矩阵,而根据图 1 对应得到的波长序列只需存为一个向量。

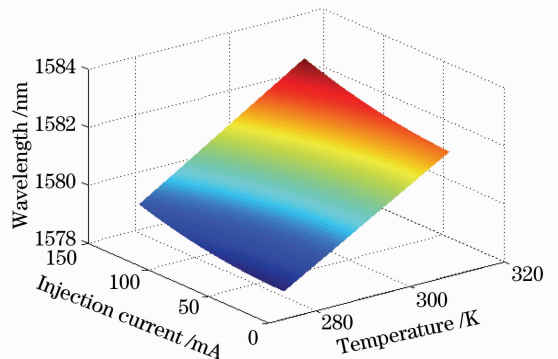


图 1 激光器发射波长与注入电流和工作温度的关系
Fig. 1 Relation between laser emission wavelengths with different injection currents and operating temperatures

温度调谐中,TEC 温度变化一个周期即扫描一周,检测信号和 NTC 输出同时进入采集卡。图 2 为 CO₂ 温度调谐原始信号,与电流调谐技术相同,包含幅值和时间信息。图 3 为 NTC 输出经过换算得到的波长序列,两图的采样点在时间上一一对应。

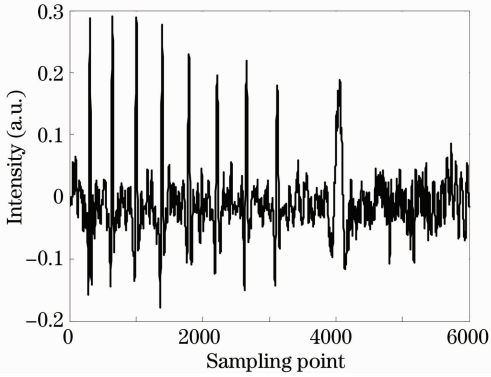


图 2 CO₂ 温度调谐二次谐波原始信号

Fig. 2 Original CO₂ second harmonic signal of temperature tuning

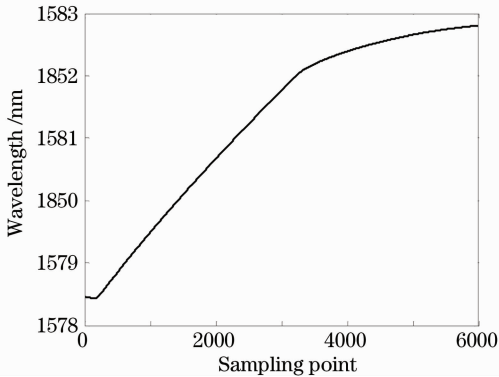


图 3 NTC 输出经过换算得到的波长序列

Fig. 3 Wavelength sequence calculated by output of NTC

图 3 中波长序列在前半段斜率较大,后半段上升平缓,且自身差异远大于电流调谐,这是由激光器根据 TEC 驱动电流的升温过程决定的,由此造成了不同采样点波长分辨率的不一致,也就是波长检测精度不同,从而影响吸收峰定位及线形计算精度。此外,采样点和波长对应不均匀导致了原始信号吸收峰在形态和频率上的差异,图 2 中后段的第 9 峰与前 8 峰明显不同,这种现象不利于信号的频域分析。

在读取和存储数据时,幅值和波长可分别存为两个矩阵,与电流调谐过程相比略为复杂,更严重的问题在于,由于不同数据的起始点可能不重合,在多周期平均时会将采样点位置差异带入计算结果,从而影响检测精度。与电流调谐信号不同,温度调谐

信号的波长序列不具重合性,如图 4 所示,200 个周期的波长序列起始点在 20 pm 范围内波动。原因有如下几点:1)激光器驱动器的温度设定精度比 NTC 检测精度低,即使在相同设定的情况下,检测到的实际工作温度也很难保持更高精度的重复性;2)受温度延时效应以及起始点选取的影响,激光器实际的起始温度与设定值会产生一定偏离;3)NTC 的检测噪声会被带入波长换算过程,进而对起始点的重合性产生比较细微的影响。从图 1 中可看出,激光器温度调谐速率要大于电流调谐速率,约为 0.1 nm/K,因此检测温度的波动在换算成波长时被放大。波长序列起始点不重合为不同周期数据的横向比较增加了难度,在使用多周期数据平均去噪、粗大误差剔除以及评估系统稳定性时存在诸多不便。

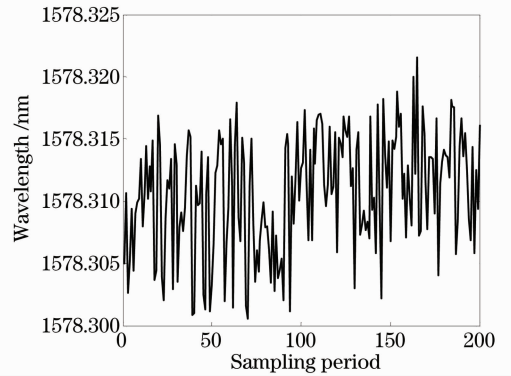


图 4 200 周期波长序列起始点

Fig. 4 Starting points of 200 periods wavelength sequence

由于温度调谐数据具有上述特征,系统在获得数据后有必要对其进行适当的预处理,使其可按照传统的电流调谐数据处理流程进行运算,以保证系统在两种调谐方式间切换时更改最少的参数设置和数据流程。

3 数据预处理及实验结果

与电流调谐数据相比,温度调谐数据不具备幅值和波长序列可分离、各周期信号采样点的波长重复性以及周期内波长分辨率一致等特征,这是本文的数据预处理过程旨在解决的主要问题。

数据预处理过程包括如下流程:数据起点和终点的选取、波长采样率的确定、插值起点选取、波长序列平滑滤波、插值过程和结果分析。

原始数据的波长起始点不一致,而波长序列的终点也在一定范围内波动,趋势同图 4 中曲线。因此要选定一个波长范围,使得结果数据的起点在所有原始数据起点之上,终点在所有原始数据终点之

下,并留出一定余量以保证每一段原始数据都能完全覆盖此波长范围。根据分析大量实验数据,可确定本文测得检测信号的实用波长范围为 1578.33~1582.63 nm,宽度为 4.3 nm。

波长采样率指固定波长范围内的采样点数,原始数据在前后两段波长采样率差异较大。图 5 为波长序列的一阶差分,表明其相邻两点的波长间隔,幅值的倒数即为该位置的波长采样率,在 3500 点前后波长采样率变化剧烈,两段相差数倍之多。结果信号的波长采样率过大会增加处理时间和无用数据量,而波长采样率过小会使信号形状失真从而影响特征值的提取精度。以单一完整吸收峰为例,包含二次谐波峰谷值以及两侧基线的宽度约为 300 pm,波长采样率设为 500 point/nm,此时吸收峰由 150 点描述,与电流调谐数据近似。

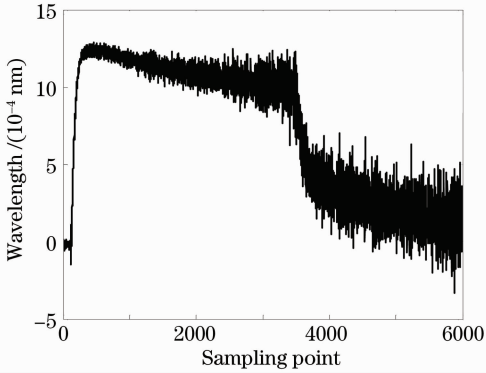


图 5 波长序列的一阶差分

Fig. 5 First-order difference of wavelength sequence

由于在插值运算过程中波长序列必须单调递增,而图 5 中一阶差分在首尾两端分别出现了负值,因此有必要对波长序列进行处理以保证其完成插值运算。图 5 中起始部分的负值表示激光器温度在大幅度上升前有短暂的下降阶段,因而其波长序列出现极小值,插值的起点要在极小值出现后,并且在结果数据的起始点之前以保证其插值精度,如图 6 所示,其中椭圆区域为插值起始点的选择区域,不同周期数据极小值点会在一定范围内波动,插值起始点应高于所有原始数据的波长起始点。

图 5 曲线尾端的负值来自 NTC 的检测噪声,在波长序列上升平缓处尤为明显。针对 NTC 的检测噪声,本文采用 Savitzky-Golay 滤波器对波长序列进行平滑滤波,Savitzky-Golay 滤波器可以根据设定的阶数和窗宽对数据曲线进行最小二乘拟合,用窗口中心点拟合值代原始数据并移动窗口实现平滑滤波,在滤除高频噪声的同时最大程度地保留了

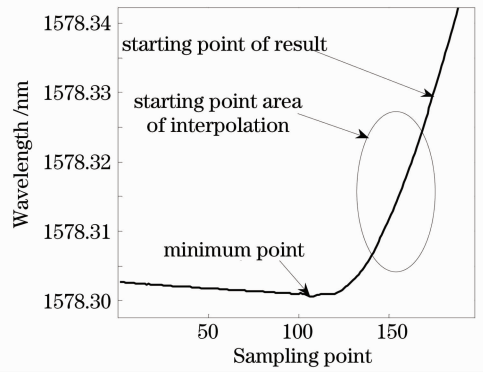


图 6 插值起始点选择区域示意

Fig. 6 Schematic of starting point area of interpolation 曲线特征^[14]。选择适当的阶数和窗宽,对图 3 中波长序列平滑滤波,对处理结果求一阶差分,如图 7 所示。与图 5 相比,曲线的变化趋势得到保留,NTC 检测噪声被滤除,曲线尾端已无负值点出现,表明波长序列单调递增,可作为位置序列对幅值序列进行插值。

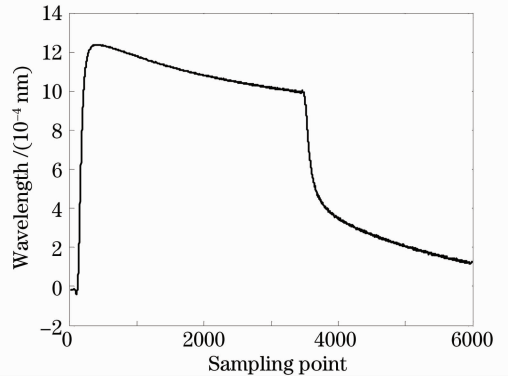


图 7 波长序列经 Savitzky-Golay 滤波器滤波后一阶差分图

Fig. 7 First-order difference of wavelength sequence processed by Savitzky-Golay filter

本文中原始波长序列的采样率要大于结果序列,不需要增补数据点,但在预设的固定位置上原始曲线往往没有数据点,为得到这些位置的信号幅值,需要对波长序列和幅值序列进行插值。在诸多插值类型中,三次样条插值适用于本文中的数据,其采用全局化分段插值的方法,能够避免高次多项式全局整体插值带来的数据震荡,具有很好的数值稳定性和收敛性,同时也保证了结果曲线的平滑度。

以原始波长序列和幅值序列作为参考,在 1578.33~1582.63 nm 区间每 2 pm 一点作为输入节点,对连续检测得到的 CO₂ 数据进行三次样条插值,将多周期处理结果存到同一矩阵。比较处理前后的 CO₂ 数据,如图 8 所示。

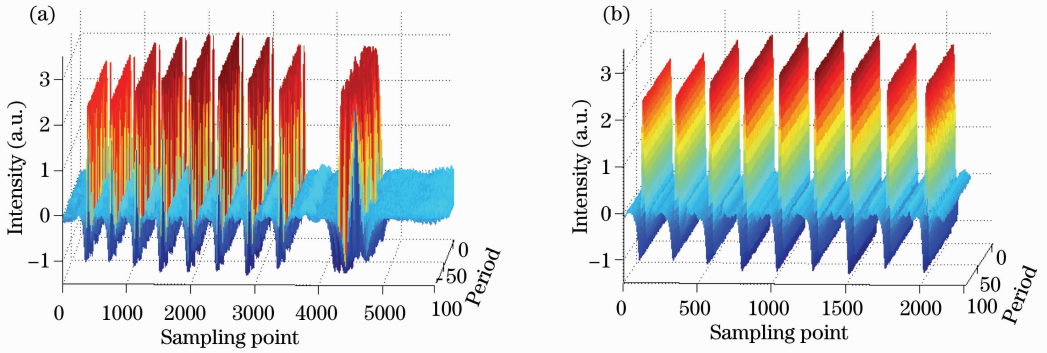


图8 多周期 CO₂ 数据处理前后结果比较。(a)处理前；(b)处理后

Fig. 8 Multi periods CO₂ signal before and after data processing. (a) Before data processing; (b) after data processing

从图8中可看出,处理前的各周期 CO₂ 原始数据在位置上并不重合,经过预处理后具有很好的位置重复性,在曲线的峰值和谷值处尤为明显。此结果与电流调谐数据具有相同的模式,可利用已有的技术对其进行后续处理,由于采样点对应固定波长,在计算过程中可以将间隔 2 pm 的采样点数作为衡量位置的单位,在结果输出时换算成波长单位。

在计算信噪比评价预处理结果时,以多周期平均结果的峰谷值作为信号,以峰值点位置的多周期标准差作为噪声,算得 CO₂ 检测信号 9 个吸收峰的信噪比。其中的噪声反映多周期平均过程带入计算结果的位置点差异,与信号对齐程度直接相关。表1为处理前后的信噪比及改善率,顺序为图8中由左至右 9 个吸收峰,改善率由两信噪比相除得到。

表1 预处理前后多周期平均结果信噪比及改善率

Table 1 Signal-to-noise ratio of multi periods averaging before and after data processing and its improvement rate

Signal-to-noise ratio before data processing /dB	Signal-to-noise ratio after data processing /dB	Improvement rate of signal-to-noise ratio
30.3	169.4	5.6
33.0	176.0	5.3
35.7	200.8	5.6
37.4	217.8	5.8
40.0	232.0	5.8
40.2	247.8	6.2
37.8	246.9	6.5
37.0	222.7	6.0
34.3	217.9	5.9
Average	214.6	5.9

从表1中可看出,温度调谐原始数据在多周期平均时将位置点的偏差带入计算结果,使得其信噪比较低,均值为 36.2,经过预处理之后信噪比有了显著提高,均值达到 214.6,平均改善率为 5.9 倍。

在已有的数据处理技术中,大量的背景信号可以用来评价系统稳定性,如果背景条纹随时间移动,吸收峰峰高的测量就会产生扰动,此时应将背景从信号中剔除^[15]。经此方法处理的温度调谐背景信号已与文献[15]中的模式相同,可用来观察系统稳定性,图9为 200 周期 N₂ 背景信号,采样时间为 1 h,此结果可直观的观察背景条纹的分布以及随时间的变化规律,为后续处理中方法和步骤的确定提供了依据。

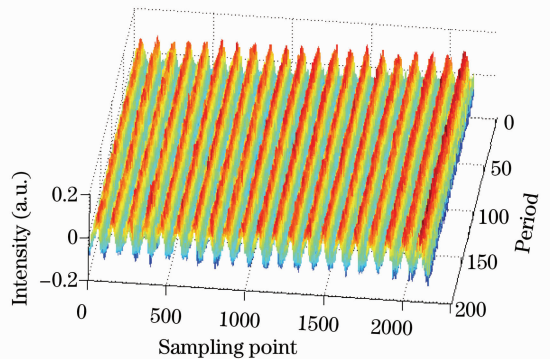


图9 200 周期 N₂ 背景信号处理结果

Fig. 9 Data processing results of 200 periods N₂ signal as background

对检测数据进行频域分析是解析信号和噪声频率范围的重要手段,波长采样率的变化会使信号各吸收峰频率不一致,从而干扰频域分析结果。经过预处理过程后的数据采样点代表波长信息,此时各吸收峰频率一致,频带更窄,更易于从原始数据中提取并分离出来。

4 结 论

分析了温度调谐数据的特征及形成原因,在此基础上利用 Savitzky-Golay 滤波器和三次样条插值

对原始数据进行预处理,使其在构成模式上更接近电流调谐数据,适用于传统的数据处理流程。对检测数据的处理结果表明,此方法可将多周期平均结果的信噪比提升为原来的6倍,有助于信号后续的运算处理及评价过程,便于分析、存储和观察,保证了线形检测的精确度,减小了吸收峰的带宽。

参 考 文 献

- 1 E. D. Hinkley. High-resolution inferred spectroscopy with a tunable diode laser[J]. *Appl. Phys. Lett.*, 1970, **16**(9): 351~354
- 2 J. Reid, J. S. Wehun, B. K. Garside *et al.*. High sensitivity pollution detection employing tunable diode lasers[J]. *Appl. Opt.*, 1978, **17**(2): 300~304
- 3 P. Werle, F. Slemr, K. Maurer *et al.*. Near- and mid-infrared laser-optical sensors for gas analysis[J]. *Opt. & Lasers in Engng.*, 2002, **37**(2-3): 101~104
- 4 B. L. Upschulte, D. M. Sonnenfroh, M. G. Allen. Measurements of CO, CO₂, OH, and H₂O in room-temperature and combustion gases by use of a broadly current-tuned multisection InGaAsP diode laser[J]. *Appl. Opt.*, 1999, **38**(9): 1506~1512
- 5 J. C. Nicolas, A. N. Baranov, Y. Cuminal *et al.*. Tunable diode laser absorption spectroscopy of carbon monoxide around 2.35 μm[J]. *Appl. Opt.*, 1998, **37**(33): 7906~7911
- 6 Kan Ruifeng, Liu Wenqing, Zhang Yujun *et al.*. Infrared absorption spectrometer of monitoring ambient methane[J]. *Acta Optica Sinica*, 2006, **26**(1): 67~70
 阚瑞峰, 刘文清, 张玉钧等. 基于可调谐激光吸收光谱的大气甲烷监测仪[J]. *光学学报*, 2006, **26**(1): 67~70
- 7 Gao Nan, Du Zhenhui, Tang Miao *et al.*. System parameters selection and optimization of tunable diode laser absorption spectroscopy [J]. *Spectrosc. & Spectral Anal.*, 2010, **30**(12): 3174~3178
 高楠, 杜振辉, 唐逸等. 可调谐二极管激光吸收光谱技术参数选择及优化[J]. *光谱学与光谱分析*, 2010, **30**(12): 3174~3178
- 8 Wang Jian, Huang Wei, Gu Haitao *et al.*. Gas temperature measurement with tunable diode laser absorption spectroscopy [J]. *Acta Optica Sinica*, 2007, **27**(9): 1639~1642
 王建, 黄伟, 顾海涛等. 可调谐二极管激光吸收光谱法测量气体温度[J]. *光学学报*, 2007, **27**(9): 1639~1642
- 9 V. Weldon, J. O'Goorman, P. Phelan *et al.*. H₂S and CO₂ gas sensing using DFB laser diodes emitting at 1.57 μm[J]. *Sensors and Actuators B*, 1995, **29**(1-3): 101~107
- 10 Weixiong Zhao, Xiaoming Gao, Lunhua Deng *et al.*. Absorption spectroscopy of formaldehyde at 1.573 μm [J]. *J. Quantit. Spectrosc. & Radiative Transfer*, 2007, **110**(2): 331~339
- 11 Zhenhui Du, Jinyi Li, Rubin Qi *et al.*. High-resolution absorption spectroscopy of carbon dioxide by NIR diode laser spectroscopy between 6320 and 6335 cm⁻¹ [C]. Proceedings of the 14th International Conference of Near Infrared Spectroscopy, 2009: 1047~1052
- 12 P. Werle. A review of recent advances in semiconductor laser based gas monitors [J]. *Spectrochimica Acta Part A*, 1998, **54**(2): 197~236
- 13 Gao Nan, Du Zhenhui, Qi Ruben *et al.*. Lineshape deviation analysis and correction method of tunable diode laser absorption spectroscopy [J]. *J. Optoelectronics · Laser*, 2011, **22**(6): 893~896
 高楠, 杜振辉, 齐汝宾等. 调谐二极管激光吸收光谱技术中的线形误差与校正[J]. *光电子·激光*, 2011, **22**(6): 893~896
- 14 A. Savitzky, M. J. E. Golay. Smoothing and differentiation of data by simplified least squares procedures [J]. *Anal. Chem.*, 1964, **36**(8): 1627~1639
- 15 P. Werle, P. Mazzinghi, F. D'Amato *et al.*. Signal processing and calibration procedures for in situ diode-laser absorption spectroscopy [J]. *Spectrochimica Acta Part A*, 2004, **60**(8-9): 1685~1705

栏目编辑: 李文喆