

基于极值统计的拉曼光谱信噪比评估方法与应用

王梓儒¹, 刘铭晖², 刘恩凯¹, 董作人², 蔡圣闻¹, 殷磊¹, 刘峰¹

1. 南京简智仪器设备有限公司, 江苏 南京 210046

2. 中国科学院上海光学精密机械研究所, 上海 201800

摘要 随着近年来便携式光谱仪技术的迅速发展, CCD 光谱仪相对于传统光谱仪在光谱收集方式上发生了很多变化: (1) 采集到的光谱对信号进行叠加积分, 传统信噪比评估方法无法通过单次检测获得探测器波动; (2) 对于谱图噪声(谱线随机波动), 由探测器响应随机波动和扫描重复误差转变为 CCD 探测器像素响应差异、探测器随机噪声和与光学系统分辨力有关的模式噪声。噪声类型发生改变, 导致原有的光谱质量评价方法适用性变差, 基于实测光谱提出更具适应性的光谱质量评价方法具有很强的现实意义。根据拉曼光谱仪检测器的变化, 对采集光谱信号的成分进行分析, 在该分析的基础上提出了 CCD 光谱仪的噪声模型假设, 根据该假设使用不同的信号极值点频率对不同的噪声进行像素分离, 并对噪声频率模式进行了数值模拟, 模拟结果与假设相符; 在此基础上提出并实验验证了通过谱线极值间距评估谱线噪声的拉曼光谱信噪比评估方法, 该方法包括以下两个步骤: (1) 通过采集多次实测光谱进行叠加, 叠加过程中对对应不同频次的光谱极值点数量进行统计, 得到统计结果后基于文中规律分离光谱仪中的环境噪声和暗噪声; (2) 应用上述分离结果, 对实测光谱中对应暗噪声的谱线极值点作统计平均, 再将该值应用于文中公式, 计算得到信噪比。该方法在进行了步骤(1)的前期准备后, 可以通过单张谱图评估 CCD 拉曼光谱仪的随机噪声, 并用于评估光谱的信噪比。基于光学构架相同、CCD 探测器不同的三个拉曼光谱系统进行实验, 采用该方法通过设定信噪比阈值对谱图质量进行控制, 获得了一致的光谱曲线; 基于该方法对同步叠加平均法进行信噪比拟合, 拟合优度达到 98%。该方法可用于拉曼光谱仪的性能评估和获取拉曼光谱谱图的质量实时控制。理论和实验表明: 对于基于 CCD 探测器的拉曼光谱仪器, 当确定样品和特征峰时, 可以基于此方法获得信噪比。该方法还可用于比对不同配置的拉曼光谱设备, 以及作为控制谱图质量一致性的标准, 并对基于拉曼光谱技术的智能鉴别系统的开发具有指导意义。

关键词 光谱学; 信噪比估计; 噪声模型; 拉曼光谱仪; 固定模式噪声

中图分类号: TH741

文献标识码: A

DOI: 10.3964/j.issn.1000-0593(2019)04-1080-06

引言

近年来, 基于 CCD 光谱仪的便携拉曼光谱仪由于其快速性和对样品的无损性, 得到了大力发展。拉曼光谱应用中弱信号检出能力是评价该类设备的核心指标。而相同样品特征信号的信噪比可作为检出能力的量化指标^[1-3]。

传统光谱检测器是基于移动光栅原理单点检测器扫描光谱, 据此前人提出了若干种评估信噪比的方法^[4-5]。而对比于传统拉曼光谱仪, CCD 光谱仪在光谱收集方式上有所改变, 扫描光谱的噪声也随之发生变化^[6-9]。噪声类型发生改

变, 导致原有的光谱质量评价方法适用性变差。

本文针对 CCD 光谱仪建立了噪声模型, 并进行了数值模拟, 与实验结果相符, 在此基础上提出一种 CCD 拉曼光谱仪信噪比评估方法, 通过单次检测光谱对拉曼光谱信噪比估计, 该方法可用于拉曼光谱仪性能评估, 以及对拉曼光谱谱图的质量控制。通过对三种不同 CCD 传感器构成的拉曼光谱系统进行实验, 利用本文提出的信噪比算法设定相同信噪比阈值, 令三种拉曼光谱仪获得了一致的谱线质量, 此外还基于同步叠加平均法进行了信噪比提升的实验, 理论与实验结果一致。

收稿日期: 2018-02-06, 修订日期: 2018-07-10

基金项目: 国家自然科学基金项目(61475165, 61775225, 61535014)资助

作者简介: 王梓儒, 1995年生, 南京简智仪器设备有限公司硕士研究生 e-mail: lf@ss-raman.com; 37283196@qq.com

1 拉曼光谱信噪比评估方法理论与仿真

1.1 CCD 光谱仪噪声分析

与移动光栅光谱仪采集方法不同, CCD 检测光谱仪采用多通道检测设备, 光谱噪声发生了以下两方面的变化:

(1) 采集到的光谱对信号进行叠加积分, 传统信噪比评估方法无法通过单次检测获得探测器波动;

(2) 谱图噪声由探测器响应随机波动和扫描重复误差转变为由 CCD 探测器像素响应差异引起的探测器随机噪声和与光学系统分辨力有关的模式噪声。

针对 CCD 光谱仪噪声特点, 本文假设拉曼光谱仪返回信号包含三个部分: ①样品光谱信号 S; ②每个像素暗电流随机波动引起的探测器像素噪声 $Noise_H$; ③由光信号载入的与光学系统分辨力相关的模式噪声 $Noise_L$ 。由噪声产生方式来看, $Noise_H$ 的频率应高于 $Noise_L$ 。而在实际的拉曼光谱应用中, $Noise_H$ 和 $Noise_L$ 都会对特征峰识别造成干扰, 但考虑到背景扰动是随环境而变化的, 且在单次检测中拉曼光谱有可能存在弱特征峰与其混淆造成信噪比不稳定或误判, 因此应设法将两种噪声进行分离。

1.2 噪声和模式噪声的像素分离

CCD 暗噪声与光信号无关, 在光谱曲线上表现为像素间的强度波动, 环境噪声导致的输入光信号波动, 经过光学系统后受光学分辨力的影响, 在光谱曲线上频率更低; 另一方

面其时域波动相对于环境光波动具有更高频率, CCD 暗噪声在时域上具有极高频率, 假设其与环境光波动引起的杂散光波动在时域上存在差异, 可以根据信号单位时间平均功率比的变化进行区分。

图 1 模拟了两个随时间波动频率比为 1 : 1 和 1 : 10 的两个随机信号, 图中示出了单位时间噪声平均功率及其比值随积分时间延长而变化的曲线。理论模拟表明, 当两个随机信号叠加时, 如果频率相同, 则随着积分时间的延长两个随机信号的平均强度比不变, 如图 1(a) 所示; 而当两个信号频率存在差异时, 低频信号与高频信号单位时间平均功率比将逐渐增加, 如图 1(b) 所示。

采用数值模拟的方法随机生成伪光谱, 其中包括“拉曼光谱信号”, 单像素随机信号构成的“暗噪声”和受光学分辨率限制的“模式噪声”。通过伪光谱数据叠加模拟积分时间延长噪声信号平滑的过程。暗噪声信号在叠加过程中随机生成。假设模式噪声频率远低于暗噪声的波动, 极限情况下杂散光在光谱上产生的波动的模式不变, 使用周期为七个像素的锯齿波模拟, 统计了每次叠加后的光谱图的四个像素间隔以内的极值点个数, 并统计其随光谱叠加次数增加的变化, 其中暗噪声和模式噪声的幅度比为 1 : 1。

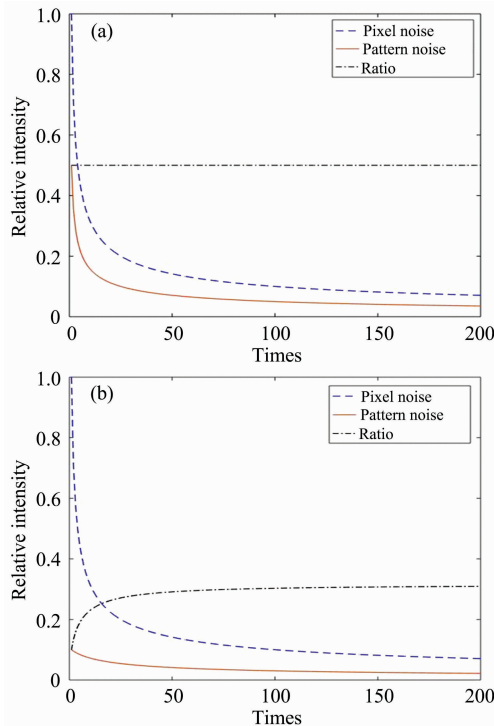


图 1 不同频率随机信号的平均功率随积分时间延长的变化关系

Fig. 1 The relationship between the average power of random signals with different frequencies and the extension of integration time

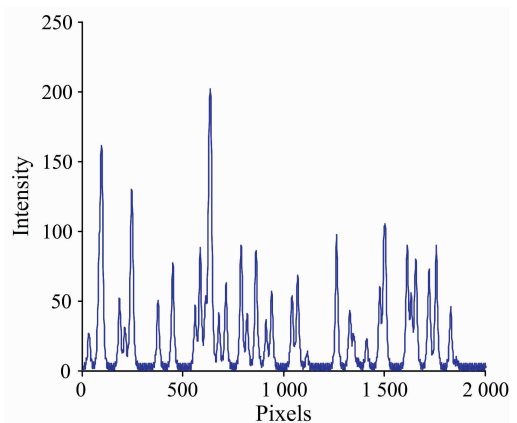


图 2 叠加杂散光信号的仿真光谱
Fig. 2 Simulation spectrum superimposed with stray light signal

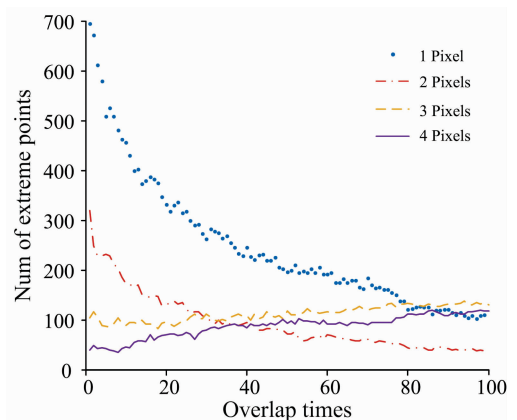


图 3 每次叠加光谱中间隔四个像素以内的极值点统计
Fig. 3 Statistics for the extreme points within 4 pixels in every overlap

由图 3 可知,随着叠加次数的增加,由于高频随机信号相对与低频信号衰减更严重,所引起的谱线波动被低频波动淹没,所以 1~2 像素间距的极值点数量减少,而间距为 3~4 个像素的极值点数量相应增加,表明 3~4 像素的波动中主要体现了低频的环境干扰信号的特征。基于这一规律可以通过实验确定光谱仪中环境噪声和暗噪声的分离条件。

1.3 信噪比评估方法

综上所述,本文提出将 CCD 光谱仪信噪比定义为

$$R_{SN} = S/Noise_H \quad (1)$$

其中 $Noise_H$ 定义为:谱图上不随光谱累积而增加的极值点极差值的统计平均,对应拉曼光谱仪暗噪声的平均功率。当采用一致样品时,用该方法评估单张谱图的信噪比,可体现由设备本身性能决定的拉曼光谱检出能力。

该信噪比评估方法主要分为两个步骤:

(1) 通过采集多次实测光谱进行叠加,叠加过程中对对应不同频次的光谱极值点数量进行统计,得到统计结果后基于上述规律分离光谱仪中的环境噪声和暗噪声;

(2) 应用上述分离结果,对实测光谱中对应暗噪声的谱线极值点作统计平均,将该值代入式(1)中的 $Noise_H$,计算得到信噪比。

2 采用不同配置光谱仪获得相同信噪比实验

2.1 实验光谱仪性能比较

实验采用光学构架相同、传感器不同的三台光谱仪,传感器型号分别为 ILX511B, S11639 和 S11510。实验中使用样品为酒精,激发光中心波长为 785 nm。

表 1 给出三种传感器的部分关键参数对比。

表 1 三种传感器参数

Table 1 Parameters of three sensors

传感器	类型	动态范围	量子效率(800 nm)/%
ILX511B	高灵敏度型	1 300	40
S11639	高灵敏度型	1 600	48
S11510	红外增强型	5 000	80

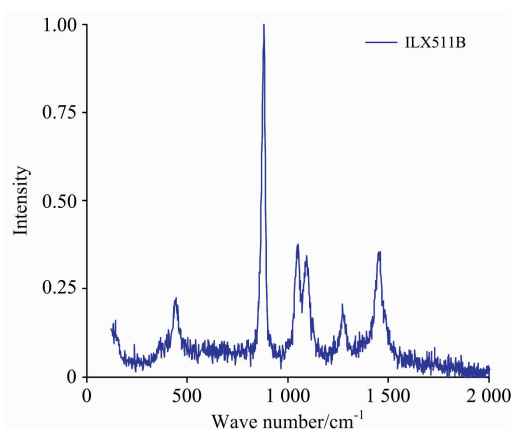


图 4 使用 ILX511B 收集的光谱

Fig. 4 Output spectrum of the spectrometer using ILX511B

如图 4—图 6 是三种设备相同条件下获得的拉曼光谱。

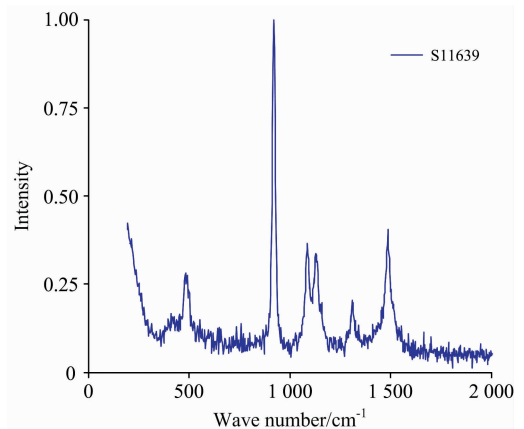


图 5 使用 S11639 收集的光谱

Fig. 5 Output spectrum of the spectrometer using S11639

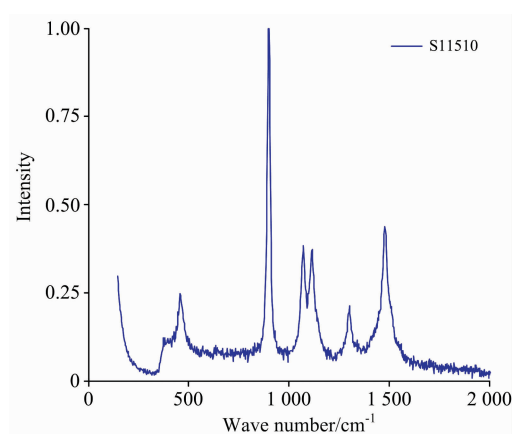


图 6 使用 S11510 收集的光谱

Fig. 6 Output spectrum of the spectrometer using S11510

2.2 $Noise_H$ 的计算

在相同条件下收集多次光谱进行叠加,统计叠加过程中谱线中毛刺信号的频率,对对应不同频次的光谱极值点数量进行统计。结果如图 7—图 9。

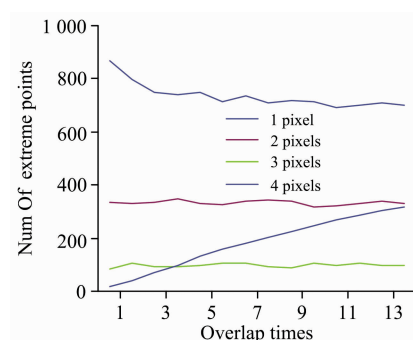


图 7 使用 ILX511B 的间隔为四个像素以内的极值点的统计

Fig. 7 Statistics for the extreme points within 4 pixels using ILX511B

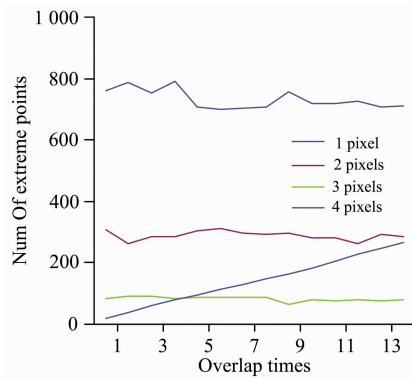


图 8 使用 S11639 的间隔为四个像素以内的极值点的统计

Fig. 8 Statistics for the extreme points within 4 pixels using S11639

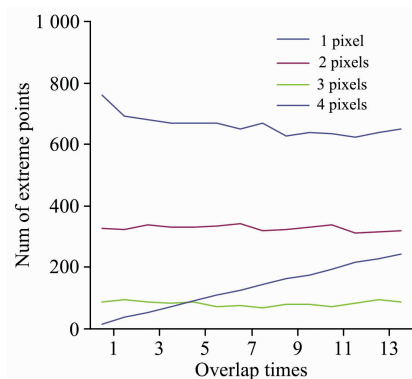


图 9 使用 S11510 的间隔为四个像素以内的极值点的统计

Fig. 9 Statistics for the extreme points within 4 pixels using S11510

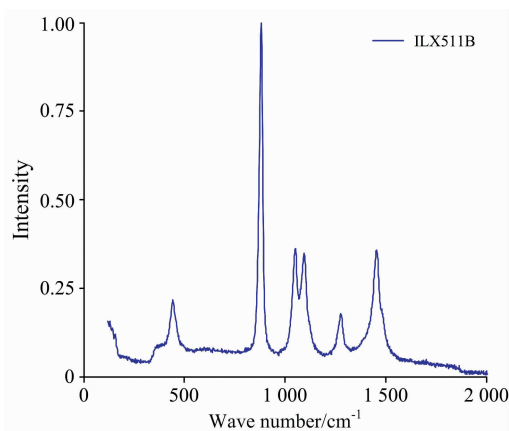


图 10 ILX511B 在信噪比评估为 200 时返回的光谱
Fig. 10 Spectrum returned at a signal-to-noise ratio of 200 using ILX511B

实验结果与 1.2 节中模拟情况一致。对于间距为 1~3 像素的毛刺信号,在信号叠加初期,信号极值点个数略微下降后渐趋稳定。当信号大到四个像素时,信号出现的概率开始与信号叠加相关,其随叠加次数增加而增加,显然该像素

间隔的极值点与 Noise_L 或样品光谱信息相关。由于设备采用相同的光学构架具有相同的光学分辨率,所以 3 台设备都在 4 像素间距时出现了随积分时间增加而增加的趋势。由于使用了相同的条件,这证明了 Noise_L 受光学系统本身特性影响。

基于此结果,根据式(1)选择所统计频率为 1~3 像素的所有毛刺信号强度作为噪声,酒精最高特征峰信号作为信号强度,获得设备信噪比参数。并以此参数控制 3 台设备,采用光谱叠加方式,根据设定的信噪比参数,由程序自动控制叠加次数,以不同的叠加次数获得相同信噪比的光谱,如图 10—图 12 所示。

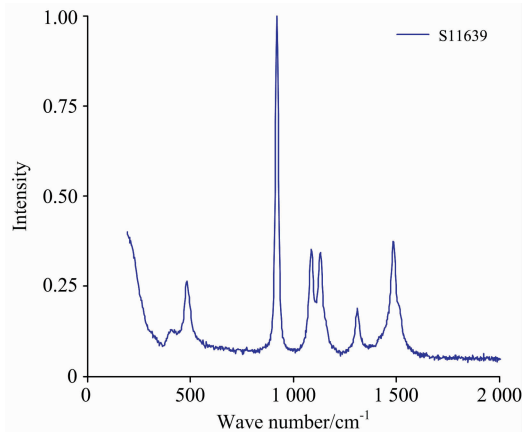


图 11 S11639 在信噪比评估为 200 时返回的光谱
Fig. 11 Spectrum returned at a signal-to-noise ratio of 200 using S11639

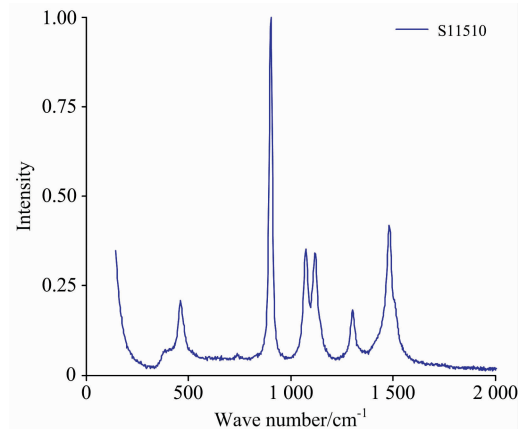


图 12 S11510 在信噪比评估为 200 时返回的光谱
Fig. 12 Spectrum returned at a signal-to-noise ratio of 200 using S11510

图 10—图 12 三条谱线表现出了良好的一致性。实验表明,本文所提出的噪声评估方法是可行的。

3 基于同步叠加算法拟合信噪比曲线

对于弱拉曼信号,受探测器饱和的限制,积分时间不可

能过大。同步叠加平均算法成为继续提升拉曼光谱信噪比的手段之一。叠加平均算法原理是基于无规则噪声累加应按照统计平均相加,而有用信号则是按照算术加法相加。根据该原理可得叠加 n 次光谱后信噪比应提高 \sqrt{n} 倍^[6]。

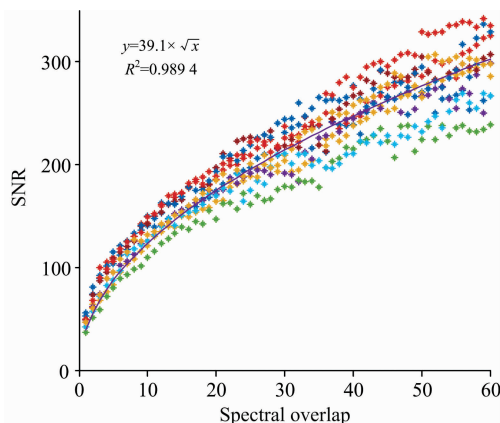


图 13 使用 ILX511B 光谱叠加过程中信噪比统计绘图

Fig. 13 SNRs statistics during the process of spectra overlapping using ILX511B

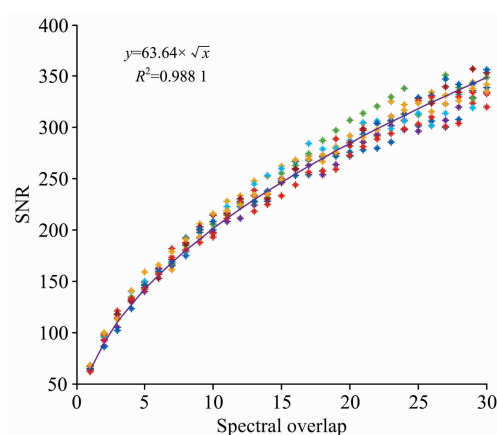


图 14 使用 S11639 光谱叠加过程中信噪比统计绘图

Fig. 14 SNRs statistics during the process of spectra overlapping using S11639

使用同步叠加平均算法对三台光谱仪进行实验,设置相同条件,使用酒精样品的 883 cm^{-1} 的特征峰,根据式(1)计算拉曼光谱信噪比。通过叠加提升所测光谱的信噪比,重复实验 10 次的数据统计如图 13—图 15。

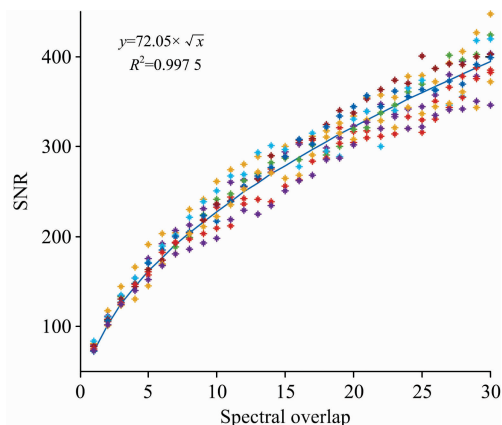


图 15 使用 S11510 光谱叠加过程中信噪比统计绘图

Fig. 15 SNRs statistics during the process of spectra overlapping using S11510

图中给出了对数据的幂指数拟合,三台光谱仪的曲线拟合优度都在 0.98 以上,与理论分析相符。图中拟合参数实为不同光谱仪单次检测的信噪比,据此参数可知,相同条件下三种不同探测器的拉曼光谱系统弱信号检测能力比为 39.1 : 63.6 : 72.1。

4 结 论

针对基于 CCD 探测器的拉曼光谱系统,提出了一种简便可行的信噪比评估方法。理论和实验表明:对于基于 CCD 探测器的拉曼光谱仪器,当确定样品和特征峰时,可以基于此方法获得信噪比,作为控制谱图质量一致性的标准。采用本文方法验证叠加平均方法拟合优度高于 98%。验证基于 ILX511B, S11639 和 S11510 探测器的三种拉曼光谱系统,相同条件下拉曼信号的实际检测能力比为 39.1 : 63.6 : 72.1。该方法对基于拉曼光谱技术的智能鉴别系统的开发具有指导意义。

References

- [1] Esmondewhite K A, Cuellar M, Uerpmann C, et al. Analytical & Bioanalytical Chemistry, 2017, 409(3): 637.
- [2] Ramasubramanyan N. Raman Spectroscopy for Bioprocess Operations, 2016.
- [3] Bartko A P. Identification of Mycoplasma Contamination in Biotechnology Production Using Raman Spectroscopy; WO, US9334520, 2016.
- [4] SUN Zhen-hua, YU Zhen-gang, HUANG Mei-zhen, et al(孙振华, 余镇岗, 黄梅珍, 等). Journal of Optoelectronics • Laser(光电子 • 激光), 2015, (6): 1132.
- [5] JIANG Cheng-zhi(姜承志). PhD Thesis, Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences(中国科学院长春光学精密机械与物理研究所博士论文), 2014.
- [6] JIANG Cheng-zhi, SUN Qiang, LIU Ying, et al(姜承志, 孙 强, 刘 英, 等). Acta Optica Sinica(光学学报), 2014, 34(6): 0630001.
- [7] Kashima N. Ieice Comex, 2014, 3(10): 300.
- [8] FAN Xian-guang, WANG Xiu-fen, WANG Xin, et al(范贤光, 王秀芬, 王 昕, 等). Spectroscopy and Spectral Analysis(光谱学与光谱

分析), 2016, 36(12): 4082.

- [9] WANG Min, LI Ting-ting, CHANG Tian-ying, et al(王 敏, 李婷婷, 常天英, 等). Semiconductor Optoelectronics(半导体光电), 2014, 35(3): 527.

Method and Application for Raman Spectra SNR Evaluation Based on Extreme Points Statistics

WANG Zi-ru¹, LIU Ming-hui², LIU En-kai¹, DONG Zuo-ren², CAI Sheng-wen¹, YIN Lei¹, LIU Feng¹

1. Nanjing S&S Instruments Co., Ltd., (Mother Company), Nanjing 210046, China

2. Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Shanghai 201800, China

Abstract In recent years, portable spectrometer technology has developed rapidly. Compared with the traditional spectrometer, CCD spectrometer in the spectral collection of the way there have been two changes: (1) The signal is superposed and integrated to generate the spectrum, and the traditional SNR estimation method cannot obtain the detector fluctuation by a single detection. (2) For the spectral noise, the detector responses to random fluctuations and scanning repetitive errors are transformed into differences in pixel response of the CCD detector, detector random noise and mode noise related to the resolution of the optical system. Therefore, it is of great practical significance to propose a more adaptive spectral quality assessment method based on the measured spectrum. According to the changes of Raman spectrometer detector, we analyze the components of the collected spectral signal, and put forward the noise model assumptions of CCD spectrometer on the basis of the analysis. According to this assumption, different signal extremum frequencies are used to separate different noise pixels and the noise frequency mode is numerically simulated. The simulation results are consistent with the assumptions. On the basis of this, we propose and experimentally validate the method for evaluating SNR of Raman spectroscopy to estimate spectral line noise through spectral line spacing. The method includes the following two steps: (1) Collecting multiple measured spectra for superposition, counting the number of spectral extreme points corresponding to different frequencies in the superposition process, and obtaining the statistical results to separate the environmental noise and dark noise in the spectrometer; (2) Applying the above separation results, the statistical average of the spectral line extreme points corresponding to the dark noise in the measured spectrum is calculated, and then the SNR is calculated according to the formula in the text. After the preliminary preparation of step (1), the method can evaluate the random noise of the CCD Raman spectrometer through a single spectrum and evaluate the spectral SNR. In this paper, three Raman spectroscopy systems with the same optical structure and different CCD detectors are used to experiment. By using this method, the spectral quality is controlled by setting the SNR threshold, and a uniform spectral curve is obtained. Based on the method proposed in this paper, the SNR is fitted to the synchronization overlapping average algorithm, and the goodness of fit is up to 98%. The method can be used to evaluate the performance of Raman spectrometer and acquire real-time quality control of Raman spectrum. Theoretical and experimental results show that for the CCD detector-based Raman spectrometer, the SNR can be obtained based on this method when determining the sample and the characteristic peak. The method can also be used to compare different configurations of Raman spectroscopy equipment and as a standard to control the quality of spectra.

Keywords Spectroscopy; SNR estimation; Noise model; Raman spectrometer; Fixed pattern noise

(Received Feb. 6, 2018; accepted Jul. 10, 2018)