

基于统计特征和显著度的光谱信号提取算法

吴江波, 贾云伟*, 姚城斌, 郝晨翔, 王 坤

天津理工大学天津市先进机电系统设计与智能控制重点实验室, 天津 300384

摘要 各类光谱信号都会受到噪声和基线畸变的影响, 在提取光谱信号过程中若不考虑基线畸变和噪声的影响, 将会严重影响信号提取的精度和准确性, 所以需要在信号提取前消除噪声和基线畸变的影响。大多数信号提取算法的步骤是先提取整体基线, 再提取信号, 这样难以保证基线的提取精度。为了降低信号提取过程中背景噪声、基线畸变等不利因素的影响, 根据信号的存在总是会导致该区域的统计特征不同于背景的特点, 提出了一种基于显著度和统计特征的光谱信号检测与提取算法(SSD算法)。首先, 在待测数据的不同尺度空间中计算出信号在各尺度下的显著度, 将检测出的显著信号点作为候选信号点; 其次, 利用信号特征去除候选信号点中的伪信号点; 最后, 对候选信号点所在区域采用二次多项式进行基线拟合以剔除伪信号区域并实现最终的信号提取。为验证 SSD 算法的综合性能, 首先, 通过仿真的方法对高斯信号和矩形信号在不同基线类型、不同信噪比下进行实验; 然后将该算法与 AirPLS 算法、Wavelet 算法以及 DoG 算法对两种信号在不同信噪比, 不同基线类型下的提取结果进行比较。仿真实验结果表明: 与其他算法相比, SSD 算法信号提取结果基本不受信号类型和基线畸变类型的影响, 且当信噪比大于 40 时基本不受信噪比的影响; 在不同基线畸变类型下, SSD 算法对两种信号提取结果的准确度、稳定性、离散度均较好, 其他算法则只适用于某种基线畸变类型。从总体提取结果上看, SSD 算法提取结果的绝对误差的均值仅为 AirPLS 算法的 8.71%、Wavelet 算法的 3.52%、DoG 算法的 2.01%; 绝对误差的均方根也仅为 AirPLS 算法的 13.08%、Wavelet 算法的 5.45%、DoG 算法的 3.11%。因此, 所提出的 SSD 算法在提取信号时具有良好的综合性能, 能够在不同的信噪比及基线畸变情况下准确的提取出信号。

关键词 光谱信号; 信号提取; 信号检测; 基线畸变; 显著度; 统计特征

中图分类号: TN911.7 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2021)07-2294-07

引言

无论是用拉曼光谱进行分子光谱检测, 还是用红外光谱鉴别化学物种, 在提取信号的过程中, 不可避免的会受到噪声和基线畸变的影响^[1]。假设检测数据中不存在基线漂移、扭曲等干扰, 则非零区域不是信号就是噪声, 只要采用合适的去噪算法消除噪声干扰即可检测出信号, 如阈值算法, 然而实际的检测数据中基线漂移、扭曲等干扰却不可避免的存在, 这些干扰使得信号区域和非信号区域的取值均可能非零, 增大了信号区域的检测难度, 为此很多算法在检测信号前都需要先提取基线以减少基线影响, 例如激光诱导击穿光谱(LIBS)定量分析技术^[2-3], Eugene Druker 提出的伽马射线光峰提取算法^[4]。然而对基线进行整体提取其精度很难得到

保证, 这是因为基线干扰因素复杂多变, 很难用模型对其进行整体描述, 即使采用迭代算法, 例如自适应迭代加权惩罚最小二乘法(AirPLS)^[5-6], 由于平滑性参数 λ 和忠于原始数据参数 s 之间的相互制约, 其基线的提取精度也并不理想, 从而导致信号检测与提取受到影响。也有人采用小波分解的方法消除基线影响, 或者直接用小波分解提取信号^[7-8], 但是小波分解精度常常受小波基和分解次数的影响, 而且小波基和分解次数通常根据个人经验来进行选取。Lowe^[9]提出了 Difference of Gaussian(DoG)算法用来处理图像, 在信号提取方面也可以使用 DoG 算法^[10]。虽然, DoG 算法可以自动提取同一频谱中不同宽度的信号, 并且提取信号的精度受基线和噪声的影响也比较小, 但是应用 DoG 算法实时性较差。为此, 本文提出了一种基于统计特征和显著度的信号检测算法。

收稿日期: 2020-07-02, 修订日期: 2020-11-22

基金项目: 国家自然科学基金项目(61201081, 61873188), 天津市自然科学基金项目(18JCYBJC19300)资助

作者简介: 吴江波, 1996 年生, 天津理工大学机械工程学院硕士研究生 e-mail: wujiangbo_vip@163.com

* 通讯作者 e-mail: yunweijia@tjut.edu.cn

所谓显著度是指某点或某区域与其周围的差异性程度^[11]，显著度检测则是指检测不同于周围的某点或某区域^[12]。即使背景中存在噪声、基线畸变等干扰，信号的存在总是会导致该区域的统计特征不同于背景。基于统计特征和显著度的信号检测算法正是检测统计特征不同于其周围的某点或某区域并将其视为信号。

1 信号提取算法

1.1 显著度计算

由于显著度是指某点或某区域不同于其周围的程度，因此原始数据在不同尺度下的显著度 $H(x, r)$ 计算如式(1)

$$H(x, r) = \begin{cases} I(x) - \frac{[I(x-r) + I(x+r)]}{2} & \text{信号为正信号} \\ \frac{[I(x-r) + I(x+r)]}{2} + I(x) & \text{信号为负信号} \end{cases} \quad (1)$$

式(1)中, $I(x)$ 为原始检测数据, r 为 $I(x)$ 的平移量, 同时也代表显著度尺度, x 为原始数据的横坐标, 不同的 r 产生不

同尺度的显著度信号, 所有尺度的显著度信号构成 $I(x)$ 的显著度空间。其中正信号是指信号幅值大于所对应基线的信号, 例如 Raman 信号、NMR 信号; 负信号是指信号幅值小于所对应基线的信号, 例如气体吸收信号。

为方便理解, 也可以令

$$B_l(x, r) = \frac{[I(x-r) + I(x+r)]}{2} \quad (2)$$

此时 $B_l(x, r)$ 可看作 $I(x)$ 的尺度为 r 的背景信号, 所有尺度的背景信号构成 $I(x)$ 的背景空间。

1.2 理想信号和基线含线性畸变时信号的显著度特点分析

图 1 各子图中的黑色实线为理论光谱, 其余不同颜色的实线及黑色虚线为在不同尺度下的显著度。当信号为理想信号时(基线无非线性畸变, 信号中无噪声), 由图 1 可以看出, 信号区域上的各点在显著度空间中, 一定有大于 0 的点与之对应, 并且绝不会出现小于 0 的情况; 而非信号区域上的各点在显著度空间中, 只能等于 0 或小于 0。因此, 检测显著度空间中大于 0 的点, 并对其进行合并, 即可检测到信号区域。

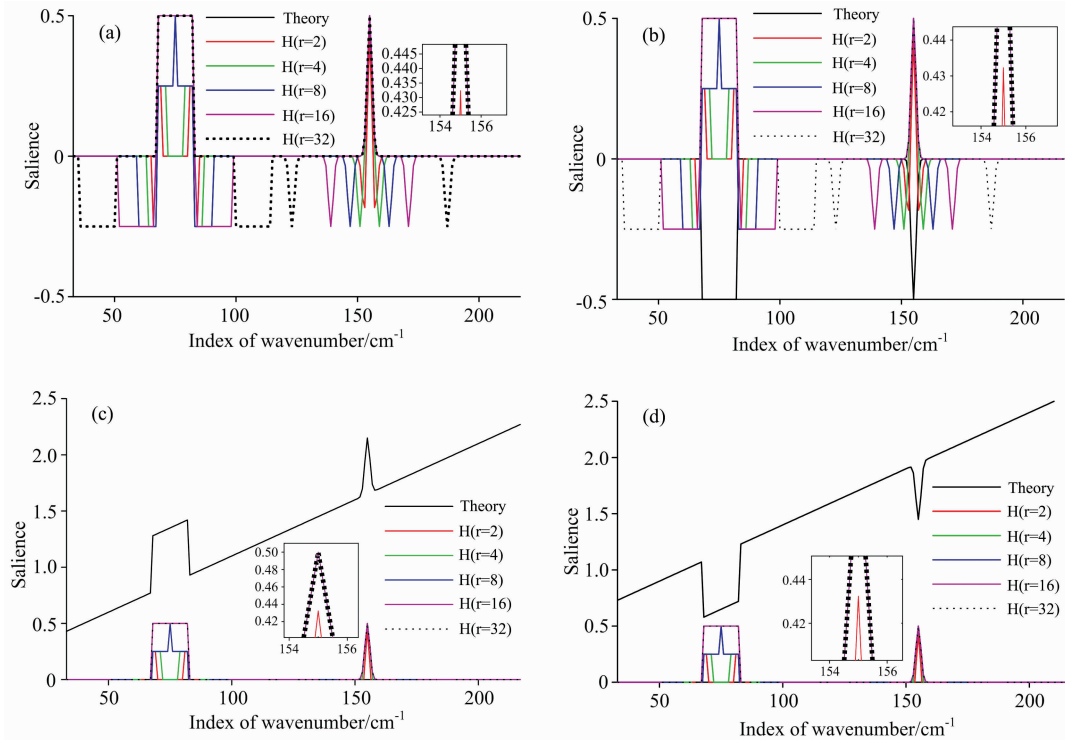


图 1 理想信号及基线存在线性畸变时的信号及其显著度空间示意

(a): 理想正信号; (b): 理想负信号; (c): 基线存在线性畸变的正信号; (d): 基线存在线性畸变的负信号

Fig. 1 The signal in ideal or the baseline having linear distortion and spatial representation of saliency

(a): Ideal positive signal; (b): Ideal negative signal; (c): Positive signal with linear distortion in the baseline; (d): Negative signal with linear distortion in the baseline

1.3 基线存在非线性畸变时信号的显著度特点分析

当基线存在非线性畸变时, 显著度特性会有一些的变化, 如图 2(a)所示, 当显著度尺度较大时, 不仅信号区域会出现大于 0 的点, 在非信号区域也会出现大于 0 的点。但是

其特点不同, 主要体现在以下三点: (1)信号区域大小随着尺度增大而基本收敛, 其最大值为 2 倍的尺度大小; (2)信号区域显著度值随着尺度增大而基本收敛。而基线存在非线性畸变引起的大于 0 的区域不一定随尺度减小而收敛且显著度

值也不一定随尺度增大而收敛,这是由于受基线非线性形状的影响。(3)若误将非线性畸变中的一段曲线当作信号(伪信号),则拟合基线与伪信号将非常相似,而真正的信号区域

与基线并不一致,信号区域的拟合基线与信号相差较大,其差约为显著度。

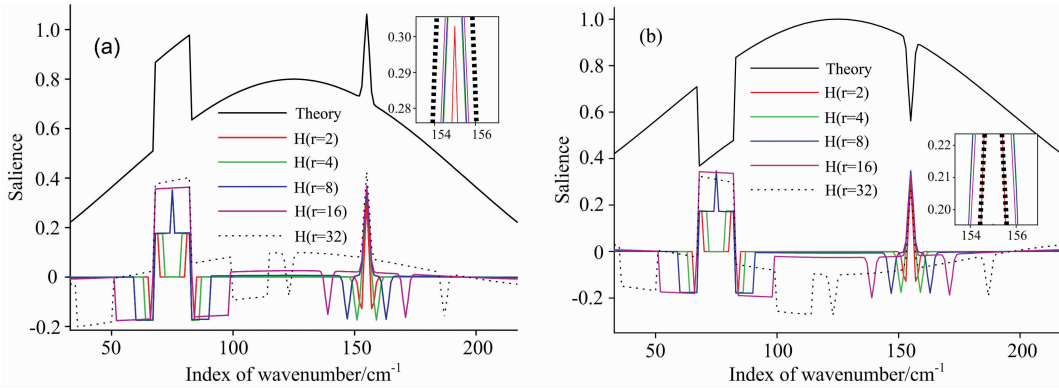


图 2 基线存在非线性畸变时的信号及其显著度空间示意

(a): 正信号及其显著度空间; (b): 负信号及其显著度空间

Fig. 2 The signal of baseline having linear distortion and spatial representation of saliency

(a): Positive signal and its saliency space; (b): Negative signal and its saliency space

因此可以认为满足以下三个特征的信号为真实信号:

Feature 1

$$H(x, r) > 0 \quad (3)$$

Feature 2: 设连续大于 0 的检测区域的数据个数为 N_{α} , N_{α} 需满足式(4)

$$N_{\alpha} \leq 2 \times r \quad (4)$$

Feature 3: 设原信号区域为 $D_s(x, r)$, 用二次多项式拟合得到的对应区域的基线为 $B_s(x, r)$, 则应满足式(5)~式(7)。

$$N_{\alpha} > \frac{N_{\alpha}}{2} \quad (5)$$

其中, N_{α} 为所判断区域可能为信号点的数据点个数, 其定义如式(6)所示

$$N_{\alpha} = \sum_{i=1}^{N_{\alpha}} \Delta \left(\frac{|D_s(x, r) - B_s(x, r)|}{H_s(x_i, r)} \right) \quad (6)$$

其中, x_i 为各个候选信号点的横坐标, i 为各个候选信号点在候选信号区域的索引值。

关于运算符 Δ 运算规则如式(7)

$$\Delta(y) = \begin{cases} 0 & \text{当 } y < \frac{2}{3} \\ 1 & \text{当 } y \geq \frac{2}{3} \end{cases} \quad (7)$$

即若某尺度下某区域同时满足式(3)~式(7), 才认为其为信号区域。由于式(5)~式(7)可适当消除尺度过大的影响, 因此该算法对尺度范围 r 和尺度间隔 Δr 两者的设置都不敏感。

图 3 显示了信号检测过程, 黑色直线代表原始光谱, 不同颜色的彩色直线和黑色虚线代表在不同尺度范围内的信号检测结果, 绿色加号代表满足 Feature 1 和 Feature 2 的区域, 红色圆圈代表满足 Feature 1, Feature 2 和 Feature 3 的区域。从图中可见, 通过式(3)~式(7)可以很好的将信号区域检测出来。

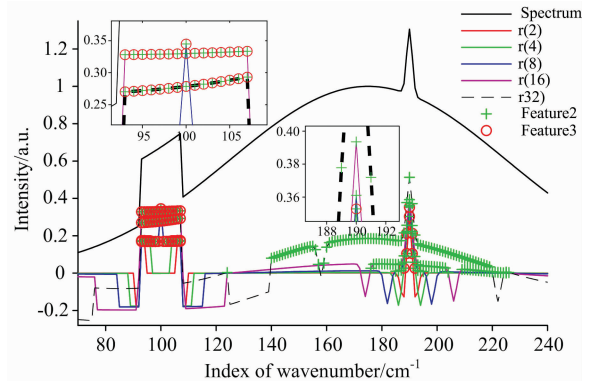


图 3 信号检测过程

Fig. 3 The process of signal detection

1.4 信号提取算法

上述分析[式(3)~式(7)]均未考虑噪声影响。而实际数据中总是不可避免的存在噪声, 式(5)~式(7)本身就是统计特征, 具有一定的抗噪声能力, 为了降低噪声的影响, 只对式(3)进行改进。首先如式(8)所示统计相邻数据点之间的显著度差值均方值, 然后将式(3)用式(9)替代。

$$\mu_r = \sqrt{\frac{1}{N_h - 1} \sum_{i=1}^{N_h} [H(x, r) - H(x_{i-1}, r)]^2} \quad (8)$$

$$H(x, r) \geq k\mu_r \quad (9)$$

其中, N_h 为数据总个数, k 为比例系数。 k 越大, 其抗干扰能力越强, 但是也容易丢失小信号。如果噪声为零均值高斯噪声, 根据 3σ 原则, 当信噪比大于 16 ($\text{SNR} = 20 \log \left(\frac{V_s}{V_n} \right)$, 其中 V_s 为信号幅值, V_n 为噪声幅值), 信号强度大于 $6\mu_r$, $k=3$ 时, 误判率和漏判率均小于 0.5%。

若想将信号提取出来, 不仅需要知道各个信号区域, 还需要知道各信号区域所对应的基线。当各信号区域已知时,

只需对其两侧一定范围内的邻域数据进行二次多项式拟合即可得到该信号区域所对应的基线 B_s 。与对整体数据进行基线拟合相比,这种局部二次多项式拟合基线不仅实时性较好,最重要的是提高了基线拟合的准确性。设各信号区域所对应的原始数据为 D_s ,则最终提取的信号 S 可由式(10)得到。

$$S = \begin{cases} D_s - B_s & \text{信号为正信号} \\ B_s - D_s & \text{信号为负信号} \end{cases} \quad (10)$$

因此最终的信号提取算法为:

- (1) 利用式(1)计算信号在各尺度的显著度;
- (2) 利用式(8)计算显著度差值均方差;
- (3) 利用式(9)和式(4)获取候选信号点;
- (4) 对候选信号点进行合并得到信号区域;
- (5) 利用二次多项式拟合得到信号区域的基线;
- (6) 利用式(5)一式(7)去除伪信号区域;
- (7) 利用式(10)提取信号。

2 实验结果与讨论

为验证 SSD 算法的有效性,本文选择 MATLAB 实验仿真平台,使用 SSD 算法对矩形信号和高斯信号在不同基线畸变类型、不同信噪比下运用不同的随机噪声分别进行了仿真实验,所有的模拟光谱信号可用式(11)来表示。基线采用比较有代表性的高斯型、指数型、二次多项式型、sin 三角函数型和线型基线畸变;在每一次的实验中我们应用 MATLAB 软件生成随机噪声,然后进行信号的提取,为使结果更具代表性,在各种情况下均进行 100 次重复性实验,即在同一理论信号、同一基线类型下共产生 100 次随机噪声得到 100 个仿真光谱,然后分析了在不同基线类型、不同信噪比下的信号提取结果;为进一步说明所提算法的有效性,还进行了 SSD 算法与不同算法的对比实验。

$$s(x) = a(x) - n(x) + b(x) \quad (11)$$

其中, $a(x)$ 为理论信号, $n(x)$ 为高斯白噪声, $b(x)$ 为理论基线, $s(x)$ 为仿真光谱。

2.1 SSD 算法信号提取性能分析

图 4 所示为含噪声且基线存在非线性畸变时某次实验的提取结果。在图 4 中,绿色实线为理论信号,黑色实线为理

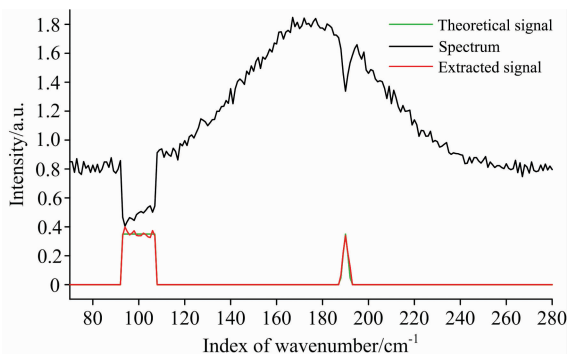


图 4 含噪声且基线存在非线性畸变时的信号提取
Fig. 4 Signal extraction with noise and the baseline in nonlinear distortion

论信号叠加非线性畸变和噪声得到的仿真光谱,红色实线为用本文算法提取的信号,从图中可以看出在含噪声且基线存在非线性畸变时提取出的信号和理论信号很接近。

图 5 显示了矩形信号(信号 1)和高斯信号(信号 2)在不同基线畸变类型、不同信噪比下各进行 100 次实验时的信号提取结果。如图 5(a)所示,使用五角星、圆圈、方块、三角形以及星花分别表示在基线畸变为指数型、高斯型、二次多项式型、sin 函数型以及线型下对信号的提取结果;红色实线代表 SSD 算法对信号 1 的提取结果,蓝色实线代表 SSD 算法对信号 2 的提取结果。图 5(a)为信号 1 和信号 2 在不同基线畸变、不同信噪比下的绝对误差均值的均值,该值的大小表征了实验结果的准确度;图 5(b)为信号 1 和信号 2 在不同基线畸变、不同信噪比下的绝对误差的均方根值,该值的大小表征了实验结果的稳定性;图 5(c)为信号 1 和信号 2 在不同基线畸变、不同信噪比下的均方根误差值的均值,该值大小表征了信号提取结果的离散度;图 5(d)为信号 1 和信号 2 在不同基线畸变、不同信噪比下的均方根误差值的均方根值,该值大小则是该信号区域提取结果稳定性的另一种表达。

从图 5(a)可以看出 SSD 总体提取结果的绝对误差的均值并不大,在基线畸变为 sin 型下对信号 1 提取结果的绝对误差的均值最大,但其值仍小于 0.015,说明 SSD 算法对信号提取的准确度较高。从图 5 中也可以看出,当信噪比小于 40 时,SSD 算法的准确度、稳定性和离散度变化较为明显,在信噪比大于 40 之后,基本不再受信噪比的影响,且在不同的基线畸变类型下同样符合上述规律。同时,从图 5 中可以看出 SSD 算法受基线畸变类型的影响较小。其中,当基线畸变类型为线型和二次多项式型时,算法的稳定性和离散度更好。经计算,SSD 算法对信号 1 和信号 2 提取结果的综合绝对误差的均值的绝对值分别为 0.004 4 和 0.004 6,对信号 1 和信号 2 的综合绝对误差的均方根的绝对值都为 0.002 9,可见 SSD 算法对信号提取的稳定性和准确度基本不受信号类型的影响。在信噪比为 20 的情况下,SSD 算法的准确度和稳定性受到影响最大的主要原因是基线拟合受噪声影响较大,从而影响所提取的信号幅值。

2.2 SSD 算法与不同算法的对比分析

图 6 显示 SSD 算法、AirPLS 算法、Wavelet 算法以及 DoG 算法在不同信噪比、不同基线畸变类型下某次信号提取的实验结果,从图中可以看出 SSD 算法既可以准确的检测出信号范围也可以对光谱进行校正。DoG 算法虽然可以检测信号范围,但是该范围并不准确,这主要是因为 DoG 检测受尺度的影响,当尺度较大时,部分非线性畸变区域会被当作信号区域检测出来。而 AirPLS, Wavelet 算法只是对光谱进行校正,并没有检测信号范围,如果再用 3σ 原则或者其他算法对校正光谱进行信号检测,则可能会在检测时引起附加误差,从而降低信号提取评价标准的可信度。为尽量避免附加误差,本文在计算误差时,认为所有的信号点都可以被检测出来,从而在整个理论信号范围内用校正的光谱强度与理论强度直接相减得到各点的误差值,对这些误差值再求均值和均方根值。

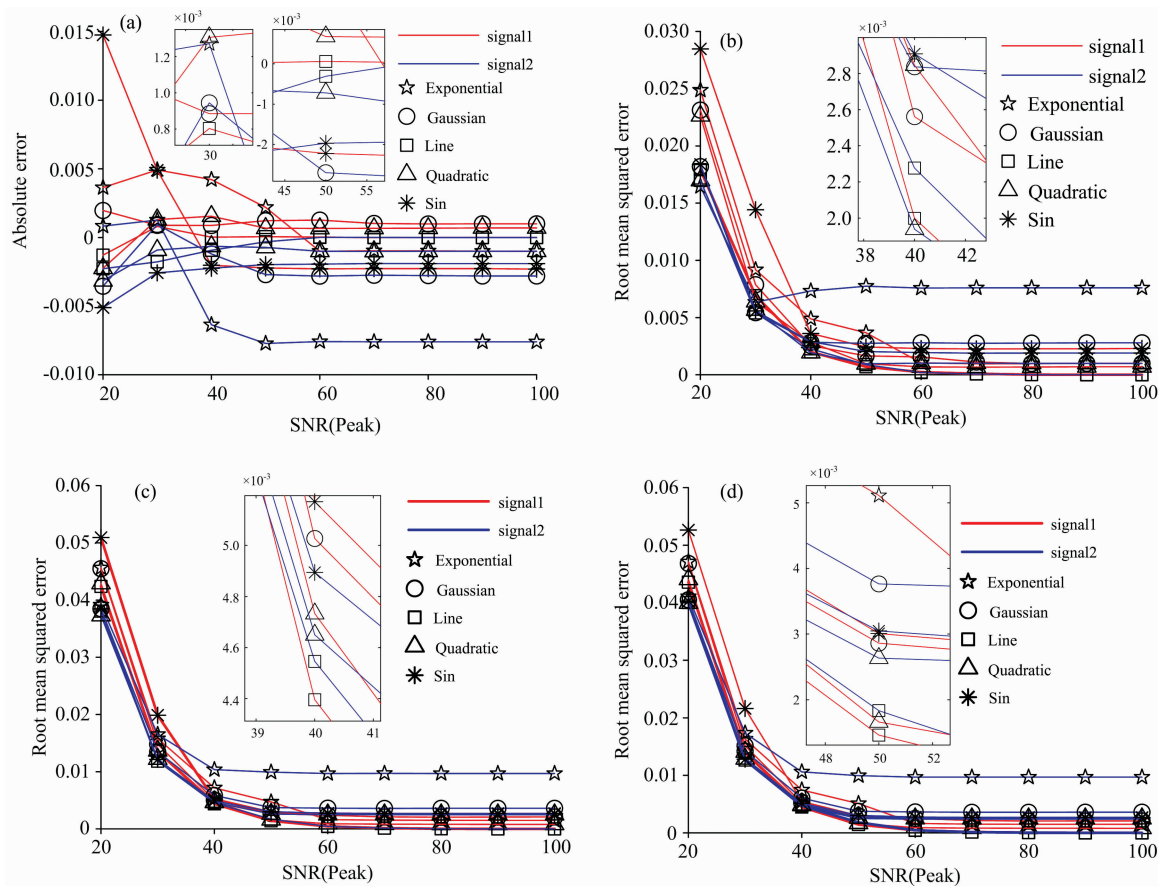


图 5 不同信噪比、不同基线、不同信号 100 次实验的信号提取结果

(a): 信号提取结果的绝对误差的均值; (b): 信号提取结果的绝对误差的均方根值; (c): 信号提取结果的均方根误差的均值; (d): 信号提取结果的均方根误差的均方根

Fig. 5 Results of signal extraction from 100 experiments with different signal-to-noise ratios, different baselines, and different signals

(a): The mean value of the absolute error of the signal extraction result; (b): The root mean square value of the absolute error of the signal extraction result; (c): The mean value of the root mean square error of the signal extraction results; (d): The mean square root of the mean square error of the signal extraction results

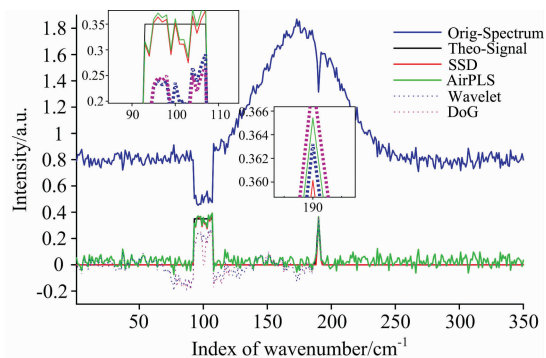


图 6 不同提取算法的提取效果比较

Fig. 6 Comparison of extraction effects of different extraction algorithms

图 7 显示了在不同信噪比、不同基线畸变类型对不同信号的提取结果统计, 图中不同颜色的直线分别指不同的提取

算法, 不同形状的点则代表提取信号时的基线类型, Exponential、Gaussian、Line、Quadratic、Sin 分别为指数型基线、高斯型基线、线型基线、二次函数型基线、Sin 型基线。图 7 (a)和(c)中各个点代表了对两种信号进行 100 次实验得到的绝对误差的均值, 该值的大小表征了算法提取结果的准确度; 图 7(b)和(d)中各点代表了对两种信号进行 100 次实验得到的绝对误差均值的均方根, 该值的大小体现了算法提取结果的稳定性。

从图 7(a)–(d)可以看出, (1)SSD 算法、AirPLS 算法以及 Wavelet 算法信号提取结果的绝对误差的均值几乎不受信噪比的影响, 说明这三种算法的准确度受信噪比的影响较小; (2)在不同的信噪比下, 与其他算法相比, SSD 算法对矩形信号和高斯信号两种信号的提取结果的绝对误差的均值都比较小, 说明 SSD 算法有较高的准确度; (3)SSD 算法的绝对误差均值的均方根也明显较小, 说明其也有较好的稳定性。

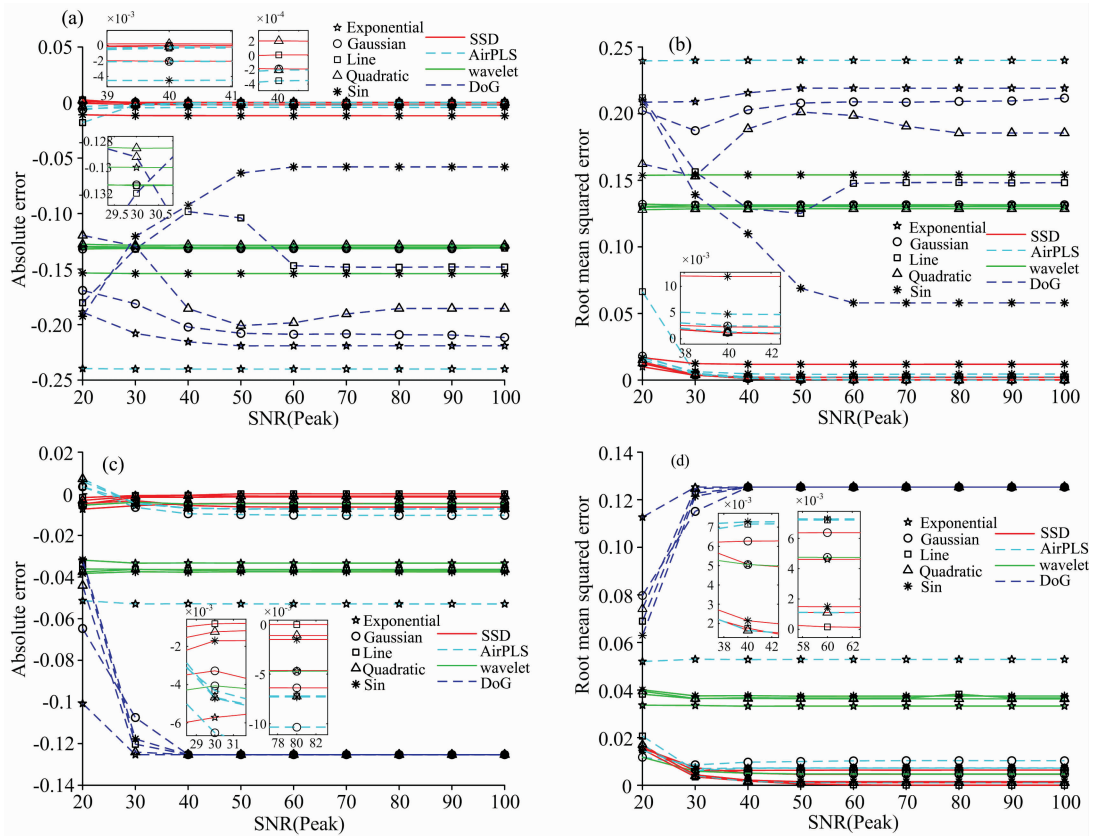


图 7 不同算法在不同信噪比、不同基线畸变类型下对不同信号的提取结果统计

- (a): 信号 1 提取结果的绝对误差的均值;
- (b): 信号 1 提取结果的绝对误差的均方根值;
- (c): 信号 2 提取结果的绝对误差的均值;
- (d): 信号 2 提取结果的绝对误差的均方根值

Fig. 7 The extraction results of different algorithms to different signals under different signal-to-noise ratios and different baseline types

- (a): The mean value of absolute error of signal 1 extraction result;
- (b): The root mean square value of the absolute error of the signal 1 extraction result;
- (c): The mean value of absolute error of signal 2 extraction result;
- (d): The root mean square value of the absolute error of the signal 2 extraction result

与其它算法相比, SSD 算法提取信号的准确度和稳定性均较好, 且基本不受基线畸变类型和信号类型的影响; Air-PLS 算法在基线畸变类型为高斯型、二次多项式型以及 sin 型时对两种信号的提取结果和 SSD 算法提取结果最相近, Wavelet 算法只有在基线畸变为高斯型时对信号 2 的提取结果的准确度以及稳定性和 SSD 算法相近。

表 1 综合了所有情况下四种算法的提取结果。可见, SSD 算法提取结果的绝对误差的均值仅为 AirPLS 算法的

8.71%、Wavelet 算法的 3.52%、DoG 算法的 2.01%; 绝对误差的均方根也仅为 AirPLS 算法的 13.08%、Wavelet 算法的 5.45%、DoG 算法的 3.11%。无论是图 7 还是表 1 都表明, SSD 算法在提取信号时具有良好的性能。

3 结 论

提出了一种基于统计特征和显著度的信号提取算法, 实验结果表明该算法具有较高的抗噪声能力, 并且极大的降低了基线畸变的影响, 即使是在含有较高噪声和非线性畸变的情况下, 仍可以准确的提取出不同形状的信号。与 AirPLS 算法、Wavelet 算法以及 DoG 算法相比, SSD 信号提取结果的准确度、稳定性均较好, 其基本不受噪声与基线畸变类型的影响。由于本文算法不需要迭代或者卷积, 所以具有较高的实时性。因此本文提出了一种鲁棒性较好的信号提取算法。

表 1 不同算法的综合提取效果

Table 1 Comprehensive extraction effect of different algorithms

	SSD	AirPLS	Wavelet	DoG
绝对误差的均值	0.002 9	0.033 3	0.082 4	0.139 7
绝对误差的均方根	0.004 5	0.034 4	0.082 6	0.144 7

References

- [1] ZHOU Feng-bo, LI Chang-geng, ZHU Hong-qiu, et al(周风波, 李长庚, 朱红求, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2018, 38(2): 506.
- [2] Yi C, Lv Y, Xiao H, et al. Spectrochimica Acta Part B: Atomic Spectroscopy, 2017, 138: 72.
- [3] Sun Y, Hao X, Ren L. 10th International Conference on Information Optics and Photonics, 2018.
- [4] Druker E. Journal of Environmental Radioactivity, 2018, 187: 22.
- [5] Cai Y, Yang C, Xu D, et al. Analytical Methods, 2018, 10(28): 3525.
- [6] Xu D, Liu S, Cai Y, et al. Applied Optics, 2019, 58(14): 3913.
- [7] Hu H, Zhang L, Yan H, et al. IEEE Access, 2019, 7: 59913.
- [8] Xi Y, Li Y, Duan Z, et al. Applied Spectroscopy, 2018, 72(12): 1752.
- [9] Lowe D G. International Journal of Computer Vision, 2004, 60(2): 91.
- [10] Jia Y W, Liu T G, Liu K, et al. J. Lightwave Technol., 2013, 31(22): 3582.
- [11] Jia Y W, Sun S Y, Yang L, et al. Analyst, 2018, 143(11): 2656.
- [12] Cao S, Zhang W. Journal of Systems Engineering and Electronics, 2020, 31(1): 37.

Spectrum Signal Extraction Algorithm and Application Based on Saliency and Statistics

WU Jiang-bo, JIA Yun-wei*, YAO Cheng-bin, HAO Chen-xiang, WANG Kun

Key Laboratory of Advanced Mechatronics System Design and Intelligent Control of Tianjin, Tianjin University of Science and Technology, Tianjin 300384, China

Abstract Signal extraction will be affected by noise and baseline distortion in most kinds of the spectrum. If the influence of noise and baseline distortion is not considered in spectrum signal extraction, the accuracy of signal extraction will be seriously decreased. Therefore, it is necessary to eliminate the influence of noise and baseline distortion before signal extraction. However, most signal extraction algorithms' procedure is to extract the whole baseline first and then extract the signal, which makes it difficult to guarantee the extraction accuracy of the baseline. A spectrum signal detection and extraction algorithm (SSD algorithm) based on saliency and statistical characteristics was proposed because the presence of signals always causes the statistical characteristics of the signal region to be different from the background. Firstly, the signal's saliency at different scales is calculated, and the detected significant signal points are taken as candidate signal points. Secondly, the pseudo-signal points in the candidate signal points are removed based on the signal characteristic that the signal should satisfy. Finally, the quadratic polynomial is used to fit the candidate signal region's baseline to remove the false signal areas and realize the final signal extraction. Many experiments were run to verify the performance of the SSD algorithm. Firstly, gaussian signal and rectangular signal were simulated under different baseline types and signal-to-noise ratio (SNR). Then different algorithms were compared, such as the AirPLS algorithm, Wavelet algorithm and DoG algorithm, on the extraction results. Simulation experiment results show that: SSD algorithm was better than compared algorithms. The signal extraction results of the SSD algorithm were not affected by the signal type and baseline distortion type and were not affected by SNR when SNR is greater than 40. Its accuracy, stability, and dispersion were good, while the other algorithms are only applicable to a certain type of baseline distortion. From the overall extraction results, the mean value of the absolute error of the SSD algorithm is only 8.71% of the AirPLS algorithm, 3.52% of the Wavelet algorithm, and 2.01% of the DoG algorithm; the root means square of the absolute error is also only 13.08% of the AirPLS algorithm, 5.45% of Wavelet algorithm, 3.11% of DoG algorithm. Therefore, the SSD algorithm proposed in this paper has good comprehensive performance in extracting signals and can accurately extract signals under different SNR and baseline distortion.

Keywords Spectrum signal; Signal extraction; Signal detection; Baseline distortion; Saliency; Statistical features

* Corresponding author

(Received Jul. 2, 2020; accepted Nov. 22, 2020)