

基于经验模态分解与回归分析的空间外差光谱目标提取

叶松^{1,2}, 李源壮¹, 孙永丰¹, 高凤艳¹, 王新强^{1,2}, 汪杰君^{1,2}, 张文涛^{1,2}, 王方原^{1,2*}

- (1. 桂林电子科技大学 电子工程与自动化学院, 广西 桂林 541004;
2. 广西光电信息处理重点实验室, 广西 桂林 541004)

摘要: 为识别空间外差光谱仪探测目标干涉信号的特征信息, 提出一种基于经验模态分解与回归分析的空间干涉谱目标提取方法。首先对预处理后的光谱进行经验模态自适应分解, 得到各阶次固有模态分量并分别计算它们与原始光谱信号的 Pearson 相关系数, 根据相关系数分选准则判定背景与目标信息重构的分界点。然后计算重构背景与实测背景间的 Pearson 相关系数来判定经验模态分解结果。对信号主导的固有模态分量利用小波软阈值进行消噪, 重构较纯净的目标特征信息; 利用目标特征信息与原始干涉光谱信息进行多元线性回归分析获得最佳的近似滤波系数, 构造滤波器并应用到目标信号, 提取目标。最后通过差谱信号与提取的目标光谱的 Pearson 相关系数来判别提取的目标信号。实验结果表明: 经验模态分解可将背景与目标近似分离; 在未知背景信号情况下, 利用经验模态分解与回归分析可实现钾共振双线特征光谱的提取。

关键词: 空间外差光谱仪; 光谱图; 经验模态分解; 多元线性回归分析; Pearson 相关系数
中图分类号: O433 **文献标志码:** A **DOI:** 10.3788/IRLA201847.1223001

Extraction of spatial heterodyne spectroscopy target based on empirical mode decomposition and regression analysis

Ye Song^{1,2}, Li Yuanzhuang¹, Sun Yongfeng¹, Gao Fengyan¹, Wang Xinqiang^{1,2}, Wang Jiejun^{1,2},
Zhang Wentao^{1,2}, Wang Fangyuan^{1,2*}

- (1. School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin 541004, China;
2. Guangxi Key Laboratory of Optoelectronic Information Processing, Guilin 541004, China)

Abstract: The algorithm was proposed based on the empirical mode decomposition and regression analysis to extract and identify the characteristic information of spatial heterodyne spectroscopy. The spectrum which was obtained by pre-processing the original probe data was decomposed into several intrinsic mode function components by empirical mode decomposition and the each order IMF's Pearson correlation coefficient was calculated with the original spectral signal. According to the correlation coefficient classification criteria, the demarcation point of the background and target information reconstruction will be determined. Then the Pearson correlation coefficient between the reconstructed background and the measured background was calculated to determine the empirical mode decomposition

收稿日期: 2018-07-10; 修订日期: 2018-08-28

基金项目: 国家自然科学基金(41561079, 41201342); 广西自动检测技术与仪器重点实验室基金(YQ17108, YQ15111, YQ16105); 桂林电子科技大学创新团队项目; 广西研究生教育创新计划(YCSW2017145)

作者简介: 叶松(1979-), 男, 研究员, 博士, 主要从事光学遥感方面的研究。Email: yesongmail@sina.com

通讯作者: 王方原(1988-), 男, 讲师, 博士, 主要从事光电信息处理方面的研究。Email: wangfy@guet.edu.cn

results. At the same time, the signal-dominated components were de-noised respectively by the wavelet soft threshold and then the pure target characteristic signal was reconstructed. By using multiple linear regression analysis to process the target characteristic information and the original interference spectral information, the optimal coefficients of time-domain filtering will be obtained. The filter will be constructed to extract the target. Finally, the signal of extracted target will be identified by Pearson correlation coefficients. The experimental results show that the background and the target can be separated by the empirical mode decomposition. In the case of unknown background signal, the empirical mode decomposition and regression analysis can be used to extract the characteristic spectrum of potassium resonance.

Key words: spatial heterodyne spectrometer; spectrum; empirical mode decomposition; multivariate linear regression analysis; Pearson correlation coefficient

0 引言

空间外差光谱技术 (Spatial Heterodyne Spectroscopy, SHS) 是一种新发展起来的具有超高分辨率的遥感技术。它综合空间调制技术和衍射光栅技术于一体,可在窄带的光谱范围内实现超高分辨光谱探测,使其对大气气体的精细谱线探测以及复杂背景下的微量气体目标信号定量反演具有显著优势^[1-2]。国内外包括美国威斯康辛大学、加拿大宇航局、中国科学院安徽光学精密机械研究所(下称中科院安光所)等研究机构分别将该技术用于大气 OH 自由基、水汽、二氧化碳等目标遥感探测研究,为遥感信号的解析与识别提供必要手段^[3-4]。

早在 1997 年,国内研究者利用傅里叶变换红外遥感技术对复杂大气背景下的目标信号采用差谱法来进行探测与识别。该方法需要事先获取纯净的背景光谱,对于动态目标以及大气背景无法获取的情况,利用差谱法难以实现对目标信号的精确探测与识别。美国的 Small G.W. 和中科院安徽光机所均开展过基于傅里叶变换干涉图直接分析的目标识别处理方法^[5-7]。该方法需要事先预知目标信号的特征峰,同时利用设计的高斯带通滤波器对干涉图滤波消除背景以及噪声干扰,它仅适用于单一吸收特征光谱,不适用多个连续吸收光谱。但就 SHS 而言,在获得精细谱线信息的同时需要研究一种对两个及两个以上的特征光谱识别的计算方法。

文中将经验模态分解 (Empirical mode decomposition, EMD) 与多元线性回归的分析方法相

结合对空间外差光谱仪探测的目标信号进行提取与识别。以钾共振双线光谱(766.5、769.9 nm)目标探测为例,利用 EMD 算法对钾共振双线特征光谱信号进行分解,获取各个阶次的固有模态分量(Intrinsic mode function, IMF),计算不同阶次的固有模态分量与原光谱信号间的 Pearson 相关系数,根据相关系数分选准则选取背景与目标光谱信号的分界点,之后进行信号的重构,利用重构背景光谱信号与实测背景光谱信号的 Pearson 相关系数来判定 EMD 方法的可行性;然后利用小波软阈值消噪技术对重构的近似目标信号消噪以便获取纯净的目标信号;最后对获取的纯净目标信号和原始信号进行多元线性回归分析,求解近似系数,构造滤波器,对目标进行滤波处理,通过计算无背景的目标信号光谱与滤波提取的目标光谱间 Pearson 相关系数,判别在未知背景信号情况下,EMD 与回归分析算法是否能实现光谱信号的有效提取与识别。

1 基本理论

1.1 SHS 探测原理

空间外差光谱仪采用两倾斜衍射光栅代替传统迈克尔逊干涉仪中两平面反射镜,两束相干光以空间调制的方式形成频域外差干涉,其结构原理如图 1 所示,光栅与光轴成 Littrow 角 θ 放置,当轴向光入射至光栅上,某一波数的光将以 θ 角原方向衍射回来,此波数称为 Littrow 波数 σ_0 ,其两出射波面与光轴垂直;而非 Littrow 波数的光经衍射返回后,其出射波面与光轴成一夹角 $\pm\gamma$,形成一定频率的空间干

涉条纹,不同波长的光出射波面与光轴的夹角 γ 取决于光栅方程:

$$\sigma(\sin\theta + \sin(\theta - \gamma)) = m/d \quad (1)$$

式中: σ 为入射光波数; m 为光栅的衍射级(常取 $m=1$); $1/d$ 为光栅刻线密度。若任意波数 σ 的光束与 Littrow 波数 σ_0 的光束出射波面成一夹角 γ , 则其两出射波面相差角度为 2γ , 波数为 σ 的两束光形成的干涉空间频率为:

$$f_x = 2\sigma \sin\gamma \approx 4(\sigma - \sigma_0) \tan\theta \quad (2)$$

当入射光为 $S(\sigma)$, 所得的干涉图为^[8-9]:

$$I(x) = \int_0^\infty S(\sigma) (\cos[8\pi(\sigma - \sigma_0)x \tan\theta]) d\sigma \quad (3)$$

干涉图是光程差 u 的函数, 其中 $u = 4x \tan\theta$, 令 $\sigma' = \sigma - \sigma_0$, 则公式(3)可表示为:

$$I(u) = \int_0^\infty S(\sigma') (\cos[2\pi\sigma' u]) d\sigma' = FT[S(\sigma')] \quad (4)$$

通过对干涉图进行傅里叶变换可恢复入射光光谱 $S(\sigma)$ 。

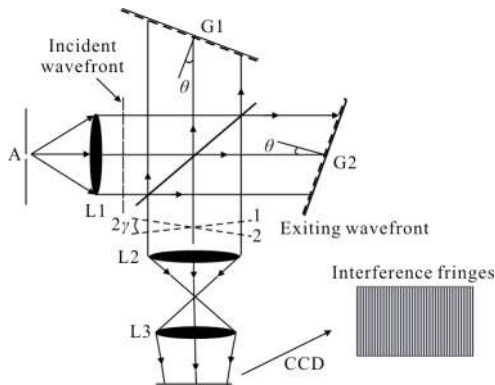


图 1 空间外差光谱仪结构原理图

Fig.1 Structure diagram of spatial heterodyne spectroscopy

根据辐射传输以及朗伯比尔定律, 光谱仪接收到的大气辐射信号既包含目标信号, 还包含大气背景以及噪声信号。在不考虑仪器偏置情况下, 光谱仪最终探测到的光谱可表示为^[10]:

$$S = R \cdot (L_B + L_O + L_S) \quad (5)$$

式中: R 为探测器的仪器响应函数; L_B 为背景辐射信号; L_O 为目标辐射信号; L_S 为噪声辐射信号。将公式(5)代入公式(4)可得:

$$I(u) = FT(RL_B + RL_O + RL_S) \quad (6)$$

由公式(6)可知: 干涉图表示背景信号、目标信号

以及噪声信号的辐射谱的傅里叶变换后叠加的形式。为了获取目标信号的特征信息必须将背景信号以及噪声信号去除掉。在实际探测情况下, 常常无法预先获取纯背景干涉信号, 因此需要设计一种方法从采集到的原始信号中分离出目标信号和背景信号。

1.2 经验模态分解

经验模态分解利用信号内部时间尺度的变化作能量与频率的解析, 其不同于小波变换等算法, 它是数据驱动算法, 可以将非线性、非平稳态信号分解为多个线性、稳态的固有模态分量。文中利用其分解特性, 将钾共振双线特征光谱从背景噪声中提取出来, 提高数据精度及有效性, 为后期钾共振双线特征光谱数据的应用打下基础。

经验模态分解是由 Huang 等人在 1998 年提出的自适应信号分解方法, 它根据信号本身的时间尺度特征将信号分解为一组固有模式分量 c_i 和一个残余分量 r_n ^[11-12], 其分解过程为:

- (1) 找出原始信号的极大值与极小值点;
- (2) 采用三次样条插值函数分别将极大值与极小值点拟合为上下包络线 W_{UP} 和 W_{LOW} , 并计算出其均值包络线 M :

$$M = \frac{W_{UP} + W_{LOW}}{2} \quad (7)$$

- (3) 利用原始信号与均值信号做差获得一个新数据, 重复以上操作直到均值包络线 M 趋于零, 此时获得第一阶固有模态分量 c_1 ;

- (4) 将第一阶固有模式分量从原始信号中分离即可得到残余项, 再对残余项重复以上操作直到残余项变为单调函数, 即可终止分解, 得到一组固有模式分量 c_i 和一个残余分量 r_n , 则原始信号经由 EMD 得到的信号可表示为^[13]:

$$x(t) = \sum_{i=1}^{n-1} c_i + r_n \quad (8)$$

1.3 多元线性回归分析

多元线性回归分析是统计学的一个重要分支, 它是表征自变量与因变量间的相关关系的一种统计学方法^[14]。在文中研究中利用 EMD 获得的背景以及目标信号均为回归分析中的因变量, 原始探测信号为自变量, 利用多元线性回归分析即可获得自变

量与因变量间的回归关系。在回归分析中要对自变量 x_i 进行循环移位生成 $(n+1) \times (n+1)$ 矩阵。因此回归系数 h_k 通过下式产生:

$$h=(X^T X)^{-1} X^T Y \tag{9}$$

式中: X 为原始数据的矩阵形式; X^T 为 X 的转置; Y 为因变量; h 为回归系数。

1.4 Pearson 相关系数

Pearson 相关也称为积差相关,主要用于度量两个随机变量之间线性关系的强弱,一般情况下,在大的样本空间中,相关系数越大表明变量数据之间相关程度越高,其表达式为:

$$R(x(t),c_i(t))=\frac{\text{cov}(x(t),c_i(t))}{\sqrt{\text{cov}(x(t))} \cdot \sqrt{c_i(t)}} \tag{10}$$

式中: $\text{cov}(\cdot)$ 为协方差。通常情况下通过相关系数 R 的取值范围来判断两个变量的相关强弱,如公式(11)所示:

完全相关	$R=1$
强相关	$R \geq 0.7$
弱相关	$0 < R < 0.7$
无相关	$R=0$

(11)

文中利用相关系数 R 的大小来判断处理前后的两个数据之间的相关强弱。

2 算法描述

由 EMD 将光谱信号 $x(t)$ 按照从低阶次到高阶次的顺序进行分解,低阶次小尺度的 IMF 表示信号的高频信息,高阶次大尺度的 IMF 表示信号的低频信息。而原始光谱信号由背景辐射信号、目标辐射信号以及噪声辐射信号组成,它们有着高低不同的频率,通常情况下,背景信息一般表现为高阶次的低频 IMF 和残余项的重构信号,噪声信号与目标信号为低阶次高频 IMF,因此选择合适尺度的 IMF 组合重构即可得到近似背景信号,将其和噪声与目标信号的混合信号分离。为了获得较为纯净的目标信号,还必须对分解得到的噪声与目标的混合信号进行小波软阈值消噪,其主要步骤如下:

首先对原始信号进行经验模式分解,得到一组 IMF 和一个残余分量,接着分别计算不同阶次的 IMF 与原始光谱信号的相关系数。

然后绘制相关系数变化曲线,根据相关系数分

选准则,选取曲线中第一个局部极小值点 k ,则第 $(k+1)$ 个点为判断背景与目标信号的界定点,1 至 $(k+1)$ 阶 IMF 重构信号为噪声与目标混合信号, $(k+2)$ 至 n 阶 IMF 重构信号为近似背景信号^[15-16]。

最后计算上一步提取的背景信号与噪声目标混合信号以及实测光谱信号间的相关系数,根据相关系数性质判断算法 EMD 的有效性;对噪声和目标的混合信号进行小波软阈值消噪得到较为纯净的目标信号,再利用该消噪后的信号与原始信号进行回归分析求解近似滤波系数,然后将获取的滤波系数形成的滤波器对探测信号进行滤波得到目标信号,完成信号的提取。其小波软阈值函数为:

$$w_{j,k} = \begin{cases} \text{sng}(w_{j,k})(|w_{j,k}| - th) & |w_{j,k}| \geq th \\ 0 & |w_{j,k}| < th \end{cases} \tag{12}$$

式中: $\text{sng}(\cdot)$ 为符号函数; th 为阈值。其算法流程如图 2 所示。

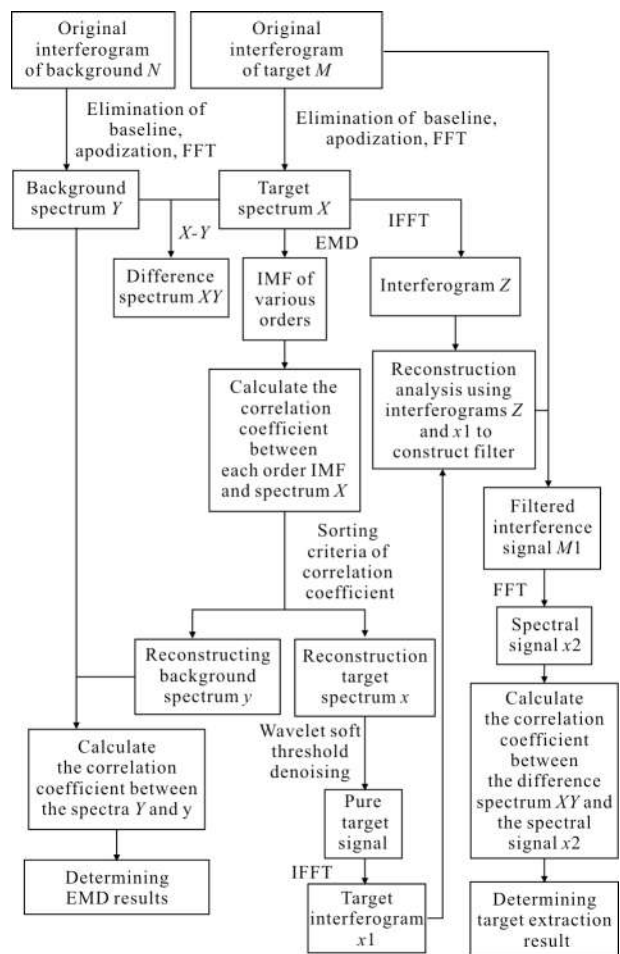


图 2 算法流程图

Fig.2 Algorithm chart

文中提出的目标信号提取算法与参考文献[17]相比存在差异,方法核心为利用高斯频域滤波器来获取有用的目标信号,而高斯频域滤波器需利用已知目标光谱的特征峰与谱线宽度来设置滤波器参数,对于未知特征峰,就无法设置该滤波器参数,则其参数的设置影响着有用目标信号及滤波系数的获取,限制了该方法的使用。而文中提到的经验模态分解与重构方法没有这些参数的限制,可自适应选取分解尺度,将目标信号与背景信号分离,提取目标信号进行回归分析获得滤波系数,弥补了其缺陷,避免了自适应差等问题。

3 实验结果与分析

为了验证上述的经验模态分解与回归分析算法的有效性,利用空间外差光谱仪探测得到的钾共振双线特征光谱信号进行验证,实现背景信号与噪声和目标的混合信号的有效分离,对噪声信号和目标的混合信号进行小波阈值消噪,以便提取目标特征信息。

测试数据由自行研制的窄波段空间外差光谱仪采集,探测目标为钾线灯。系统采用 1 024×1 024 面阵 CCD 探测器作为干涉图采集系统,主要参数为:分辨率优于 0.03 nm,光栅 Littrow 波长为 757 nm,光栅刻线密度为 600 lp/mm,视场角 θ 约为 4°。其探测的示意图如图 3 所示。

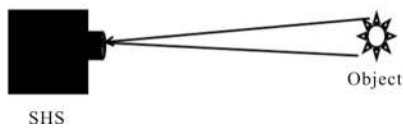


图 3 SHS 探测示意图

Fig.3 Diagram of SHS detection

3.1 干涉图处理及 EMD 分解

空间外差光谱仪探测得到的原始数据为二维干涉数据,其干涉图如图 4、图 5 所示,它们分别为钾共振双线干涉图以及大气背景干涉图。待处理光谱由多维干涉数据差分去基线、切趾、快速傅里叶变换等处理后平均得到^[18]。在已知背景信号的情况下,探测目标信号与背景信号做差,除去背景,其光谱如图 6 所示,分别为背景光谱、钾共振双线特征光谱和差

谱,从图中可知,差谱中背景因素被消除,突出了钾线双峰。在未知背景信号情况下,利用文中提到的 EMD 与回归分析算法提取钾共振双线特征光谱,消除背景影响。

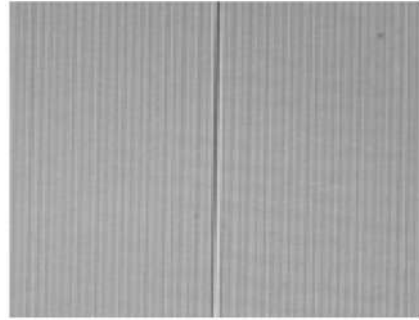


图 4 钾共振双线原始干涉图

Fig.4 Raw interferogram of potassium double resonance

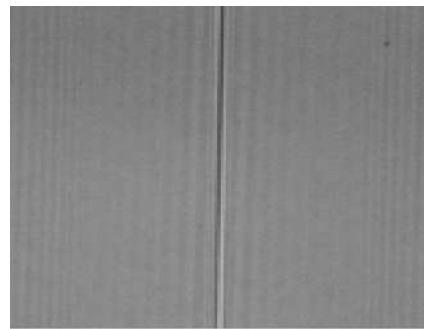


图 5 背景原始干涉图

Fig.5 Raw interferogram of background

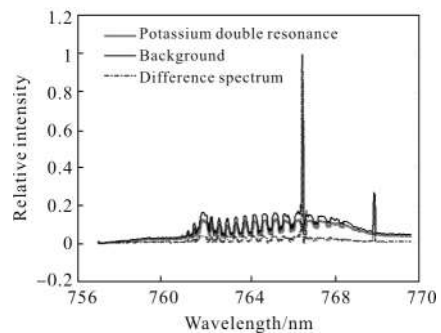


图 6 钾共振双线、背景及差谱光谱图

Fig.6 Potassium double resonance and background spectrum and difference spectrum

通过 EMD 算法,钾共振双线特征光谱被分解为不同阶次的 IMF,计算各阶次与原始钾共振双线特征光谱的相关系数,如图 7 所示,相关系数变化曲线中,第一个局部极小值出现在第 4 个点,即 k 为 4,选取第 5 点作为背景与目标信号的分界点,则 1~5

阶 IMF 为低阶高频,表现为近似目标与噪声的混合信号,之后的高阶低频 IMF 为近似背景信号;重构近似背景光谱,其光谱如图 8 所示,由于算法 EMD 并不能将背景与目标完全分离,故重构的背景信号中存在部分目标信号分量,因此会出现一个较高的峰值,其为信号的分量,且重构背景信号与实测背景信号的 Pearson 相关系数大于 0.75,为较强相关,而重构的背景信号与重构的目标噪声混合信号的 Pearson 相关系数为-0.048 4 以及重构的目标噪声混合信号与实际探测的背景信号间的 Pearson 相关系数为 0.099,为极弱相关或无相关。因此利用 EMD 算法可以获取近似背景信号,将背景信号与目标信号近似分离。

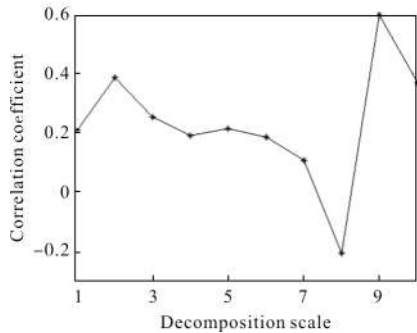


图 7 相关系数变化曲线

Fig.7 Correlation coefficient curve

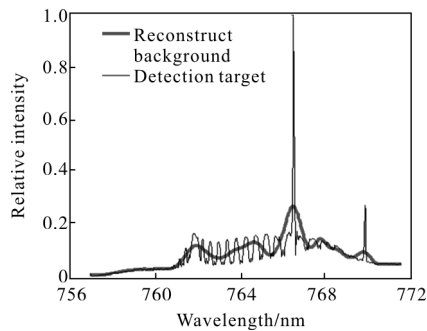


图 8 探测目标和重构背景光谱图

Fig.8 Detection target and reconstruct background spectrum

3.2 目标的提取与识别

由算法 EMD 得到的 1~5 阶 IMF 表现为近似目标与噪声的混合信号,重构目标信号与噪声信号的混合信号,对其进行小波软阈值去噪,保留了原始目标信号的细节信号,之后利用该信号与原始信号进行多元线性回归分析求解近似滤波系数,然后将获

得的滤波系数形成的 FIR 滤波器应用到原始信号,得到滤波后的目标光谱,其光谱如图 9 所示,分别为探测目标与背景的差谱和探测目标经滤除背景得到的目标光谱,它们间的 Pearson 相关系数为 0.937,则相关性较强,相似度较高,二者基本一致;与图 6 相比,钾共振双线特征波峰被保留,其他干扰波峰被大幅度削弱,突显了目标峰值,因此在未知背景信号的情况下,利用 EMD 与回归分析算法可以实现钾共振双线特征光谱目标的提取。

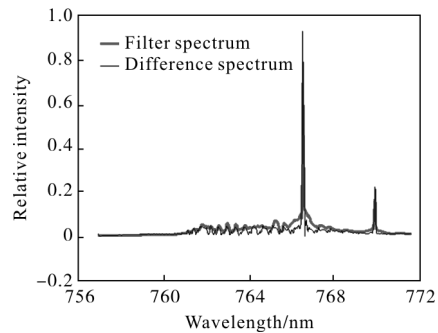


图 9 差谱与目标滤波光谱图

Fig.9 Difference spectrum and the spectrum of filtering target

4 结论

文中提出的基于 EMD 与回归分析的目标提取算法可自适应地选取分解尺度,实现算法,通过实测背景与重构背景的比较,验证了 EMD 的有效性,但是重构背景无法显示其细节,只能获得背景大概轮廓。经验模态分解与回归分析的结合,较好地将目标信号从混合信号中提取出来,保留了钾共振双线特征波峰,降低了旁峰的干扰,提高了数据的实用性,便于后期的数据处理及钾共振双线特征光谱的相关应用;但是还存在不足,提取出的目标信号强度削弱严重,如何提高目标信号强度,这将是后期的研究工作。

参考文献:

[1] Harris W M, Roesler F L, Harlander J, et al. Applications of reflective spatial heterodyne spectroscopy to UV exploration in the solar system[C]// SPIE, 2004, 5488: 886-897.

[2] Li Shuang, Qiu Zhenwei, Wang Xiangjing. Stray light simulation and analysis of space-borne spatial heterodyne spectrometer for monitoring greenhouse gases[J]. *Infrared and Laser Engineering*, 2015, 44(2):

- 616–619. (in Chinese)
- [3] Wei Qiuye, Wang Xianhua, Ye Hanhan, et al. Research of spatial heterodyne spectroscopy for atmospheric CO₂ remote sensing with high precision [J]. *Acta Optica Sinica*, 2014, 34(8): 0801006. (in Chinese)
- [4] Wang Xinqiang, Ye Song, Xiong Wei, et al. Calibration of fundamental frequency wavelength of spatial heterodyne spectrometer [J]. *High Power Laser and Particle Beams*, 2013, 25(4): 827–831. (in Chinese)
- [5] Tarumi T, Small G W, Combs R J, et al. Digital filtering implementations for the detection of broad spectral features by direct analysis of passive Fourier transform infrared interferograms [J]. *Applied Spectroscopy*, 2004, 58(4): 432–441.
- [6] Tarumi T, Small G W, Combs R J, et al. Infinite impulse response filters for direct analysis of interferogram data from airborne passive Fourier transform infrared spectrometry [J]. *Vibrational Spectroscopy*, 2005, 37(1): 39–52.
- [7] Jiao Yang, Xu Liang, Gao Minguang, et al. Real-time data processing of remote measurement of air pollution by infrared passive scanning imaging system [J]. *Acta Physica Sinica*, 2013, 62(14): 124–130. (in Chinese)
- [8] Ye Song, Fang Yonghua, Hong Jin, et al. System design of spatial heterodyne spectrometer [J]. *Optics and Precision Engineering*, 2006, 14(6): 959–964. (in Chinese)
- [9] Ye Song, Gan Yongying, Xiong Wei, et al. Baseline correction of spatial heterodyne spectrometer using wavelet transform [J]. *Infrared and Laser Engineering*, 2016, 45(11): 1117009. (in Chinese)
- [10] Yu Lijuan, Xiong Wei, Xiong Huimin. Weak signal feature extraction algorithm for passive infrared spectra [J]. *Laser & Infrared*, 2008, 38(3): 289–292. (in Chinese)
- [11] Bogdan Mijović, Ivan Gligorijević, Maarten De Vos, et al. Source separation from single-channel recordings by combining empirical-mode decomposition and independent component analysis [J]. *IEEE Transactions on Biomedical Engineering*, 2010, 57(9): 2188–2196.
- [12] Yu Chuanjin, Li Yongle, Zhang Mingjin. Comparative study on three new hybrid models using Elman Neural Network and Empirical Mode Decomposition based technologies improved by Singular Spectrum Analysis for hour-ahead wind speed forecasting [J]. *Energy Conversion and Management*, 2017, 147: 75–85.
- [13] Akhand Rai, Upadhy S H. Bearing performance degradation assessment based on a combination of empirical mode decomposition and *k*-medoids clustering [J]. *Mechanical Systems & Signal Processing*, 2017, 93: 16–29.
- [14] Wu Di. Research on gas detection using independent component analysis [D]. Harbin: Harbin Institute of Technology, 2008. (in Chinese)
吴迪. 基于独立成分分析的气体识别方法研究 [D]. 哈尔滨: 哈尔滨工业大学, 2008.
- [15] Zhang Chunmin, Ren Wenyi, Mu Tingkui, et al. Empirical mode decomposition based background removal and de-noising in polarization interference imaging spectrometer [J]. *Optics Express*, 2013, 21(8): 2592–2605.
- [16] Liu Xiaoguang, Hao Shasha, Wang Guanglei, et al. Micro mechanical gyro denoising method based on EMD autocorrelation [J]. *Journal of Chinese Inertial Technology*, 2016, 24(4): 537–541. (in Chinese)
- [17] Zhang Wentao, Gao Fengyan, Ye Song, et al. Target recognition on the finite impulse response and regression analysis from spatial heterodyne spectrometer [J]. *Opto-Electronic Engineering*, 2016, 43(7): 1–8. (in Chinese)
- [18] Yuan Zongheng, Li Shu, Ye Song, et al. Feasibility analysis of rocket plume detection based on spatial heterodyne spectroscopy [J]. *Infrared and Laser Engineering*, 2015, 44(10): 2867–2873. (in Chinese)