

文章编号: 1672-8785(2020)12-0030-06

一种基于小波最佳分解层数的红外 光谱基线校正算法

吕子敬 张 鹏 张志辉 闫继送 韩 强

(中国电子科技集团公司第四十一研究所, 山东 青岛 266555)

摘 要: 红外光谱分析技术是一种用于食品健康检测、生物制药和环境监测等方面的高新技术。为了去除应用过程中的基线漂移现象, 提出了一种基于小波最佳分解尺度的红外光谱基线校正算法。首先, 对原始光谱信号多次进行每一层的小波分解, 同时也进行每一层的小波重构。然后算得每一层的信噪比, 并通过信噪比比对法获得除噪后的光谱信号。接着对该信号进行每一层的小波分解, 得到每一层小波细节和小波逼近的频率。分别将两个频率做除法并求出比值。然后比较每层的比值大小, 并选择最大的比值作为分解层数的最佳值。将最佳分解层数下的小波逼近系数置零后再进行小波重构, 获得基线校正后的光谱信号。通过实验验证发现, 该算法不仅可以为小波分解的最佳层数提供依据, 而且能够在更好地保留有用信号的同时, 除去高频噪声以及低频基线干扰。基线校正比较充分, 效果良好。

关键词: 红外光谱分析技术; 基线漂移; 基线校正; 信噪比; 小波分解; 小波重构

中图分类号: TH744 文献标志码: A DOI: 10.3969/j.issn.1672-8785.2020.12.005

An Infrared Spectral Baseline Correction Algorithm Based on Wavelet Optimal Decomposition Layer Number

LV Zi-jing, ZHANG Peng, ZHANG Zhi-hui, YAN Ji-song, HAN Qiang

(The 41st Institute of CETC, Qingdao 266555, China)

Abstract: Infrared spectroscopy technology is a new and high technology used in food health inspection, biological pharmacy and environmental monitoring. In order to eliminate baseline drift during application, an infrared spectral baseline correction algorithm based on wavelet optimal decomposition scale is proposed. First, the multi-layer wavelet decomposition and reconstruction are performed on the original spectral signal, and the signal-to-noise ratio of each layer is calculated. Then, the denoised spectral signal is obtained by SNR contrast. The multi-layer wavelet decomposition is performed on the denoised signal to obtain the frequencies of the wavelet detail and wavelet approximation. The two frequencies are divided and the ratio is calculated. Then, the ratios of all the layers are compared, and the maximum one is selected as the best number of decomposition layers. Finally, the spectral signal after baseline correction can be obtained by setting the wavelet

收稿日期: 2020-08-23

作者简介: 吕子敬(1985-), 男, 山东青岛人, 硕士, 主要从事光谱分析算法方面的研究。

E-mail: 570824026@qq.com

approximation coefficient to zero at the optimal decomposition layer number and reconstructing the wavelet. After the experiment verification, it is found that this algorithm can not only provide the basis for the optimal number of layers for wavelet decomposition, but also can remove the high frequency noise and low frequency baseline interference while retaining the useful signals better. The baseline correction is quite sufficient and the effect is good.

Key words: infrared spectroscopy technology; baseline drift; baseline correction; signal-to-noise ratio; wavelet decomposition; wavelet reconstruction

0 引言

近些年来, 红外光谱分析技术在食品质量分析、制药工业以及环境监测等领域得到了广泛应用。它具有无损伤、灵敏度高和速度快等优点。随着该技术的发展, 作为基础的校正分析技术也越来越重要。在实际的红外光谱检测过程中, 由于仪器漂移、光源不稳定和外界环境振动等原因, 所测光谱数据产生了基线漂移, 测试精度也大大降低了。因此, 对光谱数据进行基线校正至关重要^[1]。目前用于光谱基线校正的算法很多, 常用的有一阶差分法、阈值拟合法和小波变换。其中, 相对较为简单的一阶差分去除基线的方法对单色谱有一定的效果, 但对连续谱的效果不是很明显, 很不容易彻底除掉基线。阈值拟合法是一种应用比较广泛的基线校正方法, 其校正结果受所选阈值大小的影响比较大: 如果所选阈值较大, 校正结果就会比较差; 如果所选阈值较小, 校正结果就会比较好。但是, 该方法耗时且效率低。小波变换是近些年较多用于红外光谱信号的一种分析工具, 能够有效地从红外光谱信号中提取出相关信息。此外, 小波变换还因具有快速准确的优点而常被用于红外光谱的基线校正。用于红外光谱基线校正的小波变换一直都是依靠经验来确定分解尺度, 所以很难达到预想的最佳处理效果。本文为此提出了一种基于小波最佳分解尺度的红外光谱基线校正算法。

1 红外光谱数据采集原理

通常采用傅里叶光谱仪来获取红外光谱数据。与色散型光谱仪相比, 傅里叶光谱仪近些年来获得了快速发展, 其性能也得到了大幅提

升。这是因为它具有工作波段较宽、分辨率较高以及测试速度较快的特性^[2-3]。鉴于此, 它在红外波段的物质定性分析方面也得到了广泛应用^[4]。由分束镜、探测器、动镜以及定镜等组成的经典迈克尔逊干涉仪, 可通过沿直线平移动镜来不断改变两臂的光程差。平面反射镜被用作动镜。在沿直线移动动镜的过程中, 其振动幅度和倾斜角度对测试数据准确度的影响较大。而在基于转镜的迈克尔逊干涉仪(见图 1)中, 为了保证测试稳定性, 将作为动镜的平面反射镜的两臂固定, 并通过扫描干涉仪中转动的转镜来获取光程差值。当两路光束与转镜之间的夹角都为 45° 时, 已被固定的两臂光程是一样的。但是其中一路的光程会随着转镜的转动而不断增加, 而另一路的光程则会不断减小。这时两路光就会产生光程差。此外, 转动的转镜和空气的折射率差异也是产生光程差的重要因素。

红外光谱数据采集的具体工作原理如下: 当从光源出射的光射入到分束镜时, 后者利用自身所具备的半反半透的性质将光分为两路。其中一路光经过反射镜 M1 和 M3 反射后按原路返回; 另外一路光经过反射镜 M2 和 M4 反射后, 也会按原路返回。这两路光经离轴抛物面反射镜反射后会发生干涉效应。此时, 探测器可以捕获到干涉光信号。在理想状态下, 该信号中的数据没有低频项。通过对这个不含低频项的干涉光信号进行快速傅里叶变换, 可以得到没有任何误差的光谱数据^[5]。但是, 现实中所采集的干涉光信号含有低频的基线信号。通过对采集的干涉信号进行快速傅里叶变换而获得的光谱信号中存在低频项。因此, 在理想

状态下获取的光谱数据与现实中采集到的光谱数据有较大的偏差。此时就需要对现实中采集到的光谱数据进行基线去除处理,以获取理想的光谱数据。

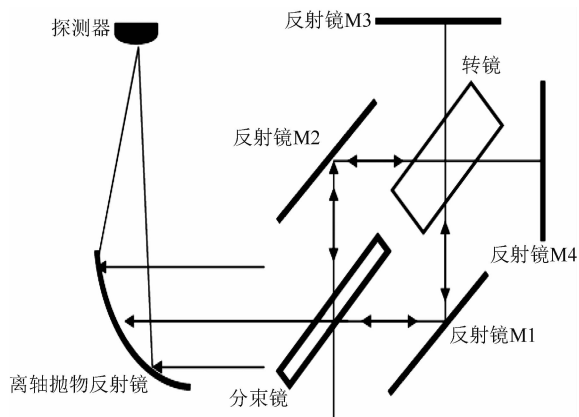


图 1 干涉信号采样的原理图

2 小波变换理论概述

小波变换是近些年发展起来的一种基于时频域的信号分析工具。它不仅具有良好的时频局部性、选基灵活性和去相关性,而且还有快速准确等优点^[7]。在某些特定条件下,函数经过伸缩或平移后产生一个函数族(即小波):

$$\varphi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \varphi\left(\frac{t-b}{a}\right) \quad (a, b \in \mathbb{R}, a \neq 0) \quad (1)$$

式中, a 代表的是控制伸缩的尺度参数, b 代表的是控制位置的平移参数, $\varphi(t)$ 代表的是小波基。

$W_f(m, n)$ 表示信号函数 $f(t)$ 的离散小波变换^[8]:

$$\begin{aligned} W_f(m, n) &= \langle f(t), \varphi_{m,n}(t) \rangle \\ &= a_0^{-m/2} \int_{-\infty}^{+\infty} f(t) \varphi(a_0^{-m} t - nb_0) dt \end{aligned} \quad (2)$$

式中, $a_0 > 1$, $m \in \mathbb{Z}$, $b_0 \in \mathbb{R}$, $n \in \mathbb{Z}$ 。 m 和 n 为离散值。

小波变换原理的本质是信号函数 $f(t)$ 被映射到小波空间内。小波系数 $W_f(m, n)$ 表示一种函数分量,即信号函数 $f(t)$ 可用部分的小波函数 $\varphi_{m,n}(t)$ 所代表的一种分量。另一种对小波变换本质的描述方式是小波函数的线性组合可由信号或任意函数表示^[9]。

对于信号来说,小波变换之所以能够在工程实践中得到广泛应用,是因为其具有较强的自适应性。在信号的时频局部化分析过程中,它的频率窗和时间窗都能够发生改变^[10]。图 2 所示为信号 X 的整个小波分解过程。

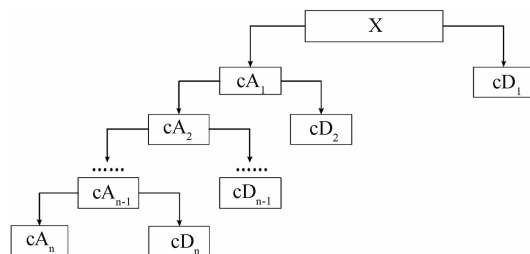


图 2 小波分解过程

图 2 中的离散值即为信号 X 。当信号 X 经过小波变换之后,分解到第 1 层时,获得用 cA_1 表示的小波逼近值和用 cD_1 表示的小波细节值;然后对 cA_1 进行分解,即分解到第 2 层,可获得用 cA_2 表示的小波逼近值和用 cD_2 表示的小波细节值。由此可以依次推出,多次分解小波逼近可以获取更多层的小波细节值和小波逼近值^[11]。信号 X 的高频与低频部分分别用小波细节值与小波逼近值来表示。分解的层数越多,信号频率越低,小波逼近值就越小。因为所有离散信号的离散点都是有限的,所以最大分解层数 N 须满足信号离散点个数大于或等于 $2N$ 的要求。由于光谱的有用信号主要集中在离散信号的高频区间,而对真实光谱信息产生干扰的基线通常分布于低频部分,因此通过给所有低频部分的小波系数都赋予零值就能去除光谱信号中的基线信息^[7]。

影响光谱基线校正效果的另一个重要因素是光谱信号分解层数的确定。当信号的分解尺度过于彻底时,基线信号去除不彻底,仍会对真实的光谱信号造成干扰;当信号的分解尺度不够彻底时,基线信号中有用的光谱信号就会被去除掉,导致光谱信号的失真现象,进而无法获取真实的光谱信息。根据经验确定分解尺度很难达到预想的最佳处理效果,为此本文提出了一种基于小波最佳分解尺度的红外光谱基线校正算法。

3 算法实施步骤和实验结果

如图 3 所示, 这种基于小波最佳分解尺度的红外光谱基线校正算法包括以下几个步骤。

步骤 1: 对已知的光谱信号 $S(i)$ 进行一层分解, 然后重构分解后的小波系数, 并求出信噪比 SNR1。

步骤 2: 对已知的光谱信号 $S(i)$ 进行两层分解, 然后重构分解后的小波系数, 并求出信噪比 SNR2。接着比较 SNR1 和 SNR2 的大小。如果 $\text{SNR1} > \text{SNR2}$, 那么转到步骤 3, 否则就进行三层分解, 经重构求出信噪比 SNR3, 再比较 SNR3 与 SNR2 的大小。如果 $\text{SNR2} > \text{SNR3}$, 那么转到步骤 3, 否则就进行四层分解, 依次迭代, 最终转到步骤 3。

步骤 3: 假如分解到 k 次时停止, 所获取的最优分解层数就是 k 。采用该方法进行除噪后的光谱信号为 $T(i)$ 。

步骤 4: 当光谱信号 $T(i)$ 的离散点个数为 m 时, 最大分解层数为 N 。它们之间的关系式为 $m=2^N$ 。

步骤 5: 将光谱信号 $T(i)$ 彻底分解到最后 一层, 然后计算每一层小波细节 cD_n ($n=1, 2, \dots, N$) 和小波逼近 cA_n ($n=1, 2, \dots, N$) 的能量值。其中 n 代表分解尺度。求得两个频率之和 Fd_n , 并将其作为小波细节的频率:

$$Fd_n = \sum_{i=1}^{ld_n} cD_n(i)^3 + \sum_{i=1}^{la_n} cA_n(i)^3 \quad (3)$$

式中, ld_n 表示 cD_n 的离散值个数, la_n 表示 cA_n 的离散值个数^[12]。

步骤 6: 根据步骤 5 获得的小波细节和小波逼近能量值, 计算出两个能量差值的绝对值 Fa_n , 并将其作为小波逼近的频率:

$$Fa_n = \left| \sum_{i=1}^{ld_n} cD_n(i)^3 + \sum_{i=1}^{la_n} cA_n(i)^3 \right| \quad (4)$$

步骤 7: 采用频率系数 K_n 来确定最佳分解层数, 其计算公式为 $K_n = Fa_n / Fd_n$ 。通过比较所有的 K 值, 获得最大值, 并将其作为分解层数的最佳值。

步骤 8: 根据步骤 7 获得的最佳分解层数 n , 进行光谱信号 $T(i)$ 的小波分解。当分解到

第 n 层时, 将低频小波系数置零, 然后进行信号的重构, 接着就可得到一个基线校正后的光谱信号 $M(i)$ 。

实验验证采用酒精的光谱曲线(见图 4 中的谱线 1)作为参照物。在实验过程中, 选取 db9 作为小波基, 对除噪后的光谱信号进行 8 层分解, 然后将小波逼近系数置零并重构信号, 最终获取了去除基线后的光谱(见图 4 中的谱线 2)。可以看出, 光谱基线去除前后的对比效果非常明显。该算法不但提高了基线校正的效果, 而且具有简单快速的优点, 同时还大大增强了效率。它不仅可以为小波分解的最佳尺度提供依据, 而且能够在更好地保留有用信号的同时, 除去高频噪声以及低频基线干扰。基线校正比较充分, 效果良好。为了进一步验证该算法去除基线的可靠性, 将经目前比较常用的光谱基线校正算法——阈值拟合法处理后的酒精光谱曲线作为对象进行了对比(见图 5)。可以看出, 两种校正算法获得的曲线相似度非常高。计算结果表明, 经两种算法校正后, 光谱的相关系数值达到 0.987, 相关度非常高。本文算法与阈值拟合法对光谱数据的校正处理效果非常一致, 再次证明了该算法基线校正的可靠性。

4 结束语

具有高精度、高分辨率和宽工作波段等优点的傅里叶光谱仪已经广泛应用于食品质量检验、生物工程和环境监测等领域。随着科技的不断发展, 它也将会在上述领域发挥更加重要的作用。对傅里叶光谱仪所测出的光谱数据进行预处理是得到真实光谱数据的重要环节, 而光谱的基线校正更是其中的重中之重。用于红外光谱基线校正的小波变换一直都依靠经验来确定分解尺度, 很难达到预想的最佳处理效果, 为此本文提出了一种基于小波最佳分解尺度的红外光谱基线校正算法, 并通过实验验证了该算法去除基线的可靠性。同时还将其与阈值拟合法进行了比对, 证明了两者的校正结果的一致性。可以看出, 本文算法在去除基线方面

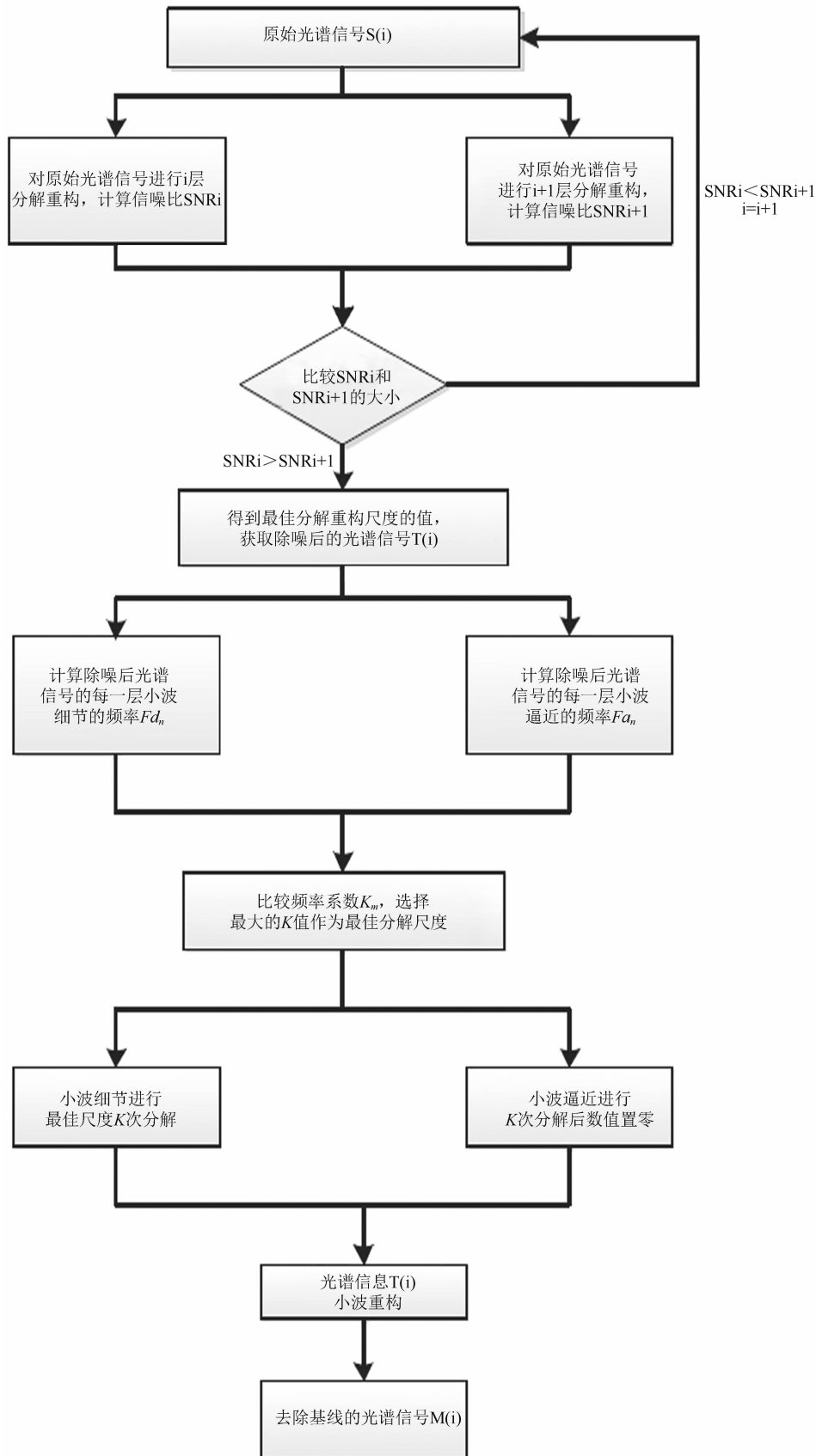


图3 算法流程图

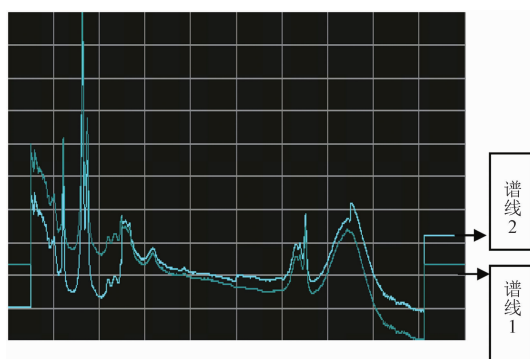


图 4 基于酒精光谱曲线的基线校正前后对比

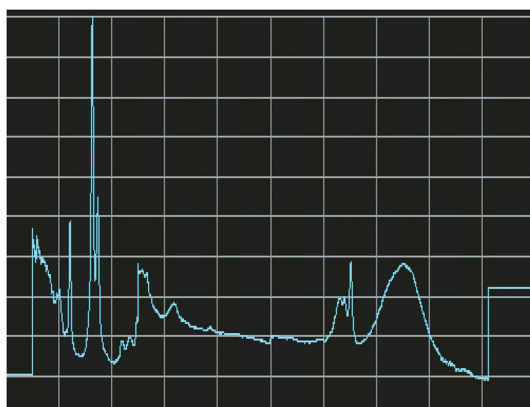


图 5 采用阈值拟合法校正后的酒精光谱曲线

具有一定的实用意义。

参考文献

- [1] 兰天鸽, 方勇华. 红外光谱信号预处理的新方法 [J]. *红外与激光工程*, 2007, **36**(2): 257–260.
- [2] 代作晓. 迈克尔逊型傅里叶光谱仪动镜直线运动 [J]. *红外*, 2002, **23**(2): 1–9.

- [3] 殷纯永. *现代干涉测量技术* [M]. 天津: 天津大学出版社, 1999.
- [4] Henry B, Luc R, Francois C. Evolution of FTIR Technology as Applied to Chemical Detection and Quantification [C]. *SPIE*, 2004, **5269**: 132–142.
- [5] 叶松, 熊伟, 乔延利, 等. 空间外差光谱仪干涉图数据处理 [J]. *光谱学与光谱分析*, 2009, **29**(3): 848–852.
- [6] Yu L, Liu H. Efficient Feature Selection via Analysis of Relevance and Redundancy [J]. *Journal of Machine Learning Research*, 2004, **25**(12): 1205–1224.
- [7] 方勇华, 孔超, 兰天鸽, 等. 应用小波变换实现光谱的噪声去除和基线校正 [J]. *光学精密工程*, 2006, **14**(6): 1088–1092.
- [8] 孙延奎. *小波分析与应用* [M]. 北京: 机械工业出版社, 2005.
- [9] 谢杰成, 张大力, 徐文立. 一种小波去噪方法的几点改进 [J]. *清华大学学报(自然科学版)*, 2002, **42**(9): 1269–1272.
- [10] 陈刚, 温志渝, 杨桂荣, 等. 小波变换在微型光谱仪光谱信号处理中的应用 [J]. *光学精密工程*, 2002, **10**(6): 552–558.
- [11] 田高友, 袁洪福, 刘慧颖, 等. 小波变换用于近红外光谱性质分析 [J]. *分析化学*, 2004, **32**: 1125–1130.
- [12] 李庆波, 徐玉波, 李响, 等. 一种改进的小波基线校正方法的研究 [J]. *光电子·激光*, 2009, **20**(5): 698–701.