

小波阈值降噪模型在红外光谱信号处理中的应用研究

吴桂芳^{1,2}, 何勇^{1*}

1. 浙江大学生物工程与食品科学学院, 浙江 杭州 310029

2. 内蒙古农业大学机电工程学院, 内蒙古 呼和浩特 010018

摘要 针对近红外光谱经常受到噪声干扰的特点, 提出了利用小波阈值降噪方法进行光谱数据的降噪处理, 以山羊绒表面油脂的近红外光谱检测为例, 对比分析了三种小波阈值降噪模型(Penalty 阈值降噪模型、Brige-Massart 阈值降噪模型、缺省阈值降噪模型)的降噪性能。对降噪后的光谱数据采用偏最小二乘和支持向量机回归相结合建立了校正和预测模型, 通过对比校验参数 R^2 , RMSEC, RMSEP, 分析评价了三种小波阈值降噪模型的降噪效果。结果表明: 三种降噪模型都能在一定程度上降低光谱信号的噪声, 提高信噪比, 改善光谱预测模型的精度, 其中, Brige-Massart 阈值降噪模型和缺省阈值降噪模型的降噪效果明显优于 Penalty 阈值降噪模型, 与原始光谱信号建模的预测精度($R^2=0.793$, RMSEC=0.233, RMSEP=0.225)相比较, 经过 Brige-Massart 阈值降噪模型降噪后的光谱信号建模的预测精度($R^2=0.882$, RMSEC=0.144, RMSEP=0.136)和经过缺省阈值降噪模型降噪后的光谱信号建模的预测精度($R^2=0.876$, RMSEC=0.151, RMSEP=0.142)均有较大程度的改善和提高, 说明提出的小波阈值降噪方法能有效地降低原始光谱噪声作用, 使光谱数据多变量分析模型更具有代表性和稳健性, 从而可以提高模型的预测精度。

关键词 近红外光谱; 小波; 阈值; 降噪

中图分类号: TS102 **文献标识码**: A **DOI**: 10.3964/j.issn.1000-0593(2009)12-3246-04

引言

红外光谱技术具有样品处理简单, 分析速度快, 可以同时测定多种组分, 实现非破坏性和无污染性测量等优点, 在食品和农产品品质检测中得到广泛的应用, 受到分析界的广泛重视^[1,2]。红外光谱分析技术的难点在于提取信息的过程, 光谱数据在用于建模前需要预处理, 预处理方法的选取对所建立光谱模型的预测精度和稳定性具有非常重要的意义。由于光谱仪采集的光谱信号往往受到环境温度、化学成分和含量的波动以及电阻的热噪声等因素干扰, 通常会含有噪声, 如果直接使用原始光谱信号进行定量分析, 必然会影响到最终的分析结果的准确性与精度, 因此, 在使用近红外光谱进行品质检测时, 需要有效的降噪预处理^[3-5]。

小波分析是近 20 年来发展起来的一种时频分析方法, 小波分析克服了短时傅里叶变换在单分辨率上的缺陷, 具有多分辨率分析的特点, 在时域和频域都有表征信号局部信息的能力, 时间窗和频率窗都可以根据信号的具体形态动态调

整, 在小波分析中, 应用最广泛的是信号处理和图像处理, 而在这两个领域里应用最多的就是信号的降噪和压缩^[6-8]。由于在正交小波中, 正交基的选取比传统方法更接近实际信号的本身, 所以通过小波变换可以容易地分离出噪声或其他我们不需要的信息, 因此在这类应用中小波分析有着传统方法无可比拟的优势。在小波分析用于降噪的过程中, 关键在于阈值的确定, 合理的确定有效阈值可以在很大程度上提高小波分析的降噪能力^[9-12]。

本研究对比分析三种不同的小波阈值降噪模型对光谱数据进行降噪的效果, 以改善光谱数据的信噪比, 进一步提高多变量分析模型的预测能力。

1 小波阈值降噪分析的模型

1.1 小波降噪分析的原理

小波分析用于降噪过程, 可分为如下三步, 即

(1) 分解过程: 选定一种小波, 对信号进行 N 层小波(小

收稿日期: 2008-12-26, 修订日期: 2009-03-28

基金项目: 国家自然科学基金项目(30671213), 国家“十一五”科技支撑项目(2006BAD10A09)和国家高技术研究发展计划“863”计划项目(2007AA10Z210)资助

作者简介: 吴桂芳, 女, 1972 年生, 浙江大学生物系统工程与食品科学学院博士研究生 e-mail: wgf@sara@126.com

* 通讯联系人 e-mail: yhe@zju.edu.cn

波包)分解;

(2)作用域过程:对分解得到的各层系数选择一个阈值,并对细节系数作用软阈值处理;

(3)重建过程:降噪处理后的系数通过小波(小波包)重建恢复原始信号。

用小波降噪分析对信号降噪的模型如图 1 所示。

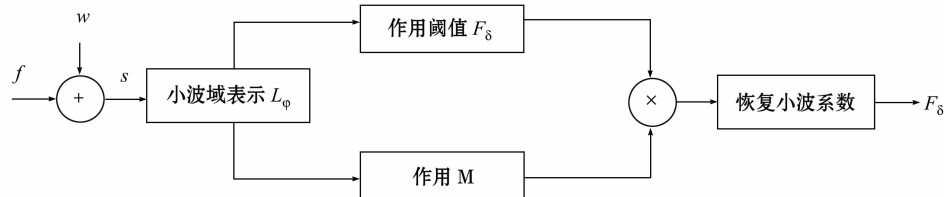


Fig. 1 De-noising model with wavelet

其中, f 为原始信号; w 为噪声信号; s 为原始信号被噪声污染后的信号; L_ϕ 为信号 s 在小波域中的表示, 即信号 s 在小波变换下的分解系数; F_δ 为阈值算子, F_δ 作用于模型的结果是模值小的系数被置为零, 只保留模值大的系数项; M 为掩码算子, M 作用于模型的结果是保留特定的系数并把其他的系数置为零。

1.2 小波阈值降噪分析的数学模型

在小波分析用于降噪的过程中, 核心步骤就是在系数上作用阈值。因为阈值的选取直接影响降噪的质量, 所以人们提出了各种理论的和经验的模型。但没有一种是通用的, 他们都有各自的适用范围, 阈值的确定主要有以下几个数学模型^[13]。

(1)小波包变换中的 Penalty 降噪模型, 阈值由下式给出。

令 t^* 为使得函数

$$\text{crit}(t) = - \sum_{k \leq t} c_k^2 + 2\sigma^2 t(\alpha + \log(n/t)) \quad (1)$$

取得最小值的 t , 其中 c_k 为小波包分解系数排序后第 k 个数。 n 为系数的总数, 那么阈值

$$\text{thr} = |c(t^*)| \quad (2)$$

式中的 σ 为信号的噪声强度, α 为经验系数, α 必须为大于 1 的实数, 随着 α 的增大, 降噪后信号的小波系数会变稀疏, 重建后的信号会变得更加光滑。 α 在采用 penalty 阈值降噪时通常取 2。

(2)Brige-Massart 降噪模型, 阈值通过如下的规则求得。

①给定一个指定的分解层数 j , 对 $j+1$ 以及更高层, 所有系数保留;

②对第 i 层 ($1 \leq i \leq j$), 保留绝对值最大的 n_i 个系数, n_i 由下式确定

$$n_i = M(j+2-i)^\alpha \quad (3)$$

式中 M 和 α 为经验系数, 缺省的情况下取 $M=L(1)$, 也就是第一层分解后系数的长度, 一般情况下, M 满足 $L(1) \leq M \leq 2L(1)$, α 的取值因用途不同, 在降噪情况下一般取 $\alpha=3$ 。

③缺省阈值的降噪模型, 阈值由如下公式给出

$$\text{thr} = \sqrt{2 \log(n)} \sigma \quad (4)$$

其中 n 为信号的长度, σ 为信号的强度。

2 小波阈值降噪法在光谱数据处理上的应用

2.1 材料、设备与方法

山羊绒采用内蒙古自治区的鄂尔多斯市阿尔巴斯白山羊绒(Albas), 随机采取山羊绒样本 120 份(每份 5 g), 全部样本随机分成建模集和预测集, 建模集有 80 个样本, 预测集有 40 个样本。光谱测量仪器是日本 JASCO Model FTIR-4000 傅里叶红外光谱仪)。采用光谱漫反射模式, 每个样本采集 40 次, 并对采集值取平均。由于要分析光谱信号的噪声, 为保证检测范围内存在一定的噪声, 采用了光谱仪器的全波段光谱信号 ($350 \sim 7\,800 \text{ cm}^{-1}$), 山羊绒表面油脂含量采用瑞典 FOSS2050 自动索氏脂肪分析仪测量, 对每个样本的脂肪含量严格按照 GB/T 5413.3—1997 检测。油脂含量测定值为山羊绒样本中的油脂的质量百分数(%)。试验中所测得的山羊绒表面油脂含量在 $0.32\% \sim 1.27\%$ 范围内, 标准差(S)为 0.153 ^[14,15]。

2.2 山羊绒样本的近红外漫反射光谱

实验所得近红外光谱曲线如图 2 所示。图 2 中纵坐标为光谱反射率, 横坐标为波数, 范围是 $350 \sim 7\,800 \text{ cm}^{-1}$, 包括近红外波段 ($350 \sim 4\,000 \text{ cm}^{-1}$) 和中红外波段 ($4\,000 \sim 7\,800 \text{ cm}^{-1}$)。从图中可以看出试验获得的原始数据受到噪声的干扰非常大, 为了降低光谱信号的噪声, 提高信噪比, 改善采用多变量分析方法对光谱数据的建模精度, 我们分别采用了三种小波降噪模型(小波包变换中的 Penalty 阈值模型, Brige-Massart 阈值模型, 缺省阈值的确定模型)对试验所得的光谱信号进行降噪处理, 三种小波阈值降噪数学分析模型用 Matlab 软件编程实现, 并将降噪结果显示出来, 分析结果如图 3 所示。降噪后的光谱在进行分析之前需要进行光谱预处理。本文采用采用 Savitzky-Golay 平滑法, 平滑点数为 9, 然后采用多元散射校正(MSC)对光谱数据进行处理。所有的预处理过程在 Unscrambler V9.7 软件中进行。经过降噪以及预处理之后的光谱曲线如图 4 所示, 取 ($500 \sim 7\,000 \text{ cm}^{-1}$)

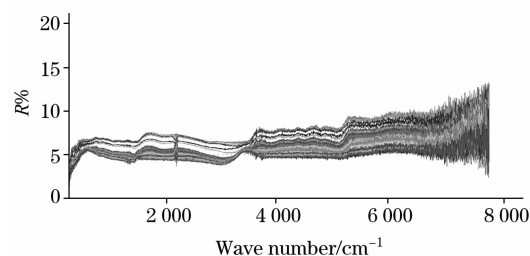


Fig. 2 Near infrared reflectance spectra of cashmere

的光谱数据用于分析建模^[16-20]。

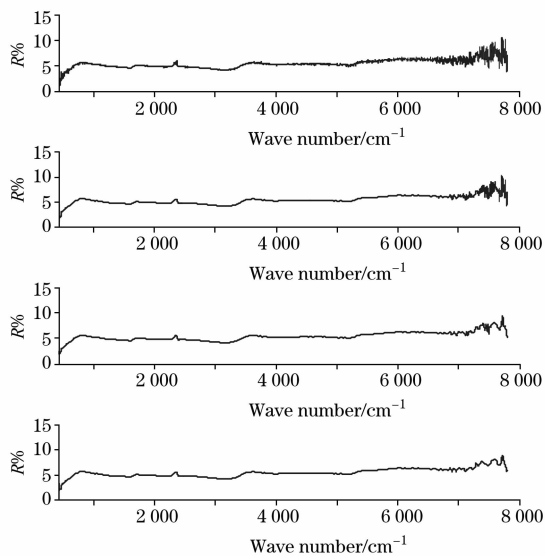


Fig. 3 De-noising results of near infrared reflectance spectra

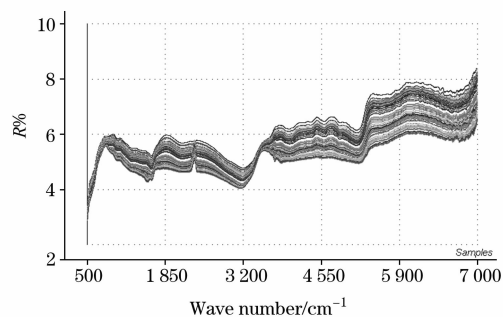


Fig. 4 De-noising and preprocessed results of near infrared reflectance spectra

2.3 小波阈值降噪模型的评价

为评价三种小波阈值降噪模型对光谱数据降噪效果,本研究分别对原始光谱和经过上述三种小波阈值模型降噪后的光谱数据分别应用多变量分析方法(偏最小二乘与支持向量机回归相结合)和山羊绒表面油脂含量建立校正模型和预测模型。采用决定系数(R^2)、校正模型标准差(RMSEC)、预测

模型标准差(RMSEP)来验证模型的预测精度,由表 1 可知,三种降噪模型都能在一定程度上降低光谱信号的噪声,提高信噪比,改善光谱预测模型的精度,在三种小波阈值降噪模型中,Brige-Massart 阈值降噪模型和缺省阈值的降噪模型的降噪效果明显的优于 Penalty 阈值降噪模型的降噪效果,同原始光谱信号建模的预测精度($R^2 = 0.793$, RMSEC = 0.233, RMSEP = 0.225)相比较,经过 Brige-Massart 阈值降噪模型降噪后光谱信号建模的预测精度($R^2 = 0.882$, RMSEC = 0.144, RMSEP = 0.136)和经过缺省阈值降噪模型降噪后光谱信号建模的预测精度($R^2 = 0.876$, RMSEC = 0.151, RMSEP = 0.142)有较大程度的改善和提高,说明小波阈值降噪方法能有效地降低原始光谱的噪声作用。

Table 1 Precision comparison of calibration set and prediction set between noisy spectral and de-noised spectral

De-noising type	R^2	RMSEC	RMSEP
None(Original spectral)	0.793	0.233	0.225
Penalty	0.829	0.194	0.189
Brige-Massart	0.882	0.144	0.136
Default threshold	0.876	0.151	0.142

3 结 论

小波变换近年来在分析化学中得到了广泛的应用,小波变换可以只对某频率段的局部信号进行特殊方式的频谱处理,可以有效地滤除信号中的高频或倍频噪声,在小波变换用于降噪的过程中。阈值的选取决定降噪的质量,本研究以山羊绒表面油脂的近红外光谱检测为例,选取了三种小波阈值降噪模型(Penalty 阈值降噪模型、Brige-Massart 阈值降噪模型、缺省阈值降噪模型)分别对光谱数据降噪,结果表明,三种降噪模型都能在一定程度上降低光谱信号的噪声,改善光谱信号的质量,三种方法中,Brige-Massart 阈值降噪模型和缺省阈值降噪模型的降噪效果明显的优于 Penalty 阈值降噪模型,可以很大程度的提高信噪比,改善近红外光谱对山羊绒表面油脂含量的检测精度,可以进一步应用于近红外光谱测量的定性、定量分析研究之中。

参 考 文 献

- [1] Weyer L G. Applied Spectroscopy Reviews, 1985, 21(1, 2): 1.
- [2] Stark E, Luchter K. Applied Spectroscopy Reviews, 1986, 22(4): 335.
- [3] Poornachandra S. Digital Signal Processing, 2008, 18: 49.
- [4] Eiceman G A, Wang M, Prasad S, et al. Analytica Chimica Acta, 2006, 579: 1.
- [5] Pasti L, Walczak B, Massart D L, et al. Chemometrics and Intelligent Laboratory Systems, 1999, 48: 21.
- [6] Cinquemani E, Pillonetto G. Automatica, 2008, 44: 2288.
- [7] Azzalini A, Farge M, Schneider K. Applied and Computational Harmonic Analysis, 2005, 18: 177.
- [8] Kubota H, Tanki N, Nanjo T. International Congress Series, 2004, 1268: 1238.
- [9] XIA Guo-rong, XU Zhi-sheng, MA Zhen-yan(夏国荣, 徐志胜, 马振燕). Journal of Test and Measurement Technology(测试技术学报), 2007, 21: 319.
- [10] WANG Xin, ZHU Gao-zhong(王 新, 朱高中). High Voltage Engineering(高压电技术), 2008, 34: 342.

- [11] ZHU Hua, WU Chuan-sheng, WANG Xiao-mei(朱 华, 吴传生, 汪小梅). Computer Applications(计算机应用), 2007, 27: 2605.
- [12] ZHAO Jun-long, TAN Cheng-qian, JIAO Ji-tian, et al(赵军龙, 谭成仟, 焦积田, 等). Journal of Xi'an University of Science and Technology(西安科技大学学报), 2007, 27(2): 263.
- [13] DONG Chang-hong, GAO Zhi, YU Hai-xiao(董长虹, 高 志, 余海啸). Principle and Application of Wavelet Analysis Toolbox(小波分析工具箱原理与应用). Beijing: National Defence Industry Press(北京: 国防工业出版社), 2004. 1.
- [14] SHAO Yong-ni, HE Yong(邵咏妮, 何 勇). Journal of Infrared and Millimeter Waves(红外与毫米波学报), 2006, 25(6): 478.
- [15] CHU Xiao-li, XU Yu-peng, LU Wan-zhen(褚小立, 许育鹏, 陆婉珍). Chinese Journal of Analytical Chemistry(分析化学), 2008, 36(5): 702.
- [16] Liu Fei, He Yong, Wang Li. Analytica Chimica Acta, 2008, 615(1): 10.
- [17] Bona M T, Andrés J M. Analytica Chimica Acta, 2008, 624(1): 68.
- [18] Ripoll G, Albertí P, Panea B, et al. Meat Science, 2008, 80: 697.
- [19] Wu Di, He Yong, Feng Shui-juan. Analytica Chimica Acta, 2008, 610(2): 232.
- [20] Cleve E, Bach E, Schollmeyer E. Analytica Chimica Acta, 2000, 420: 163.

Application of Wavelet Threshold Denoising Model to Infrared Spectral Signal Processing

WU Gui-fang^{1, 2}, HE Yong^{1*}

1. College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310029, China

2. College of Mechanical and Electrical Engineering, Inner Mongolia Agricultural University, Huhhot 010018, China

Abstract Aimed at noise interference of infrared spectra, an example of using infrared spectra to detect fat content value on the surface of cashmere was applied to evaluate the effect of wavelet threshold denoising. The denoising capabilities of three wavelet threshold denoising models (penalty threshold denoising model, Brige-Massart threshold denoising model and default threshold denoising model) were compared and analyzed. Denoised spectra and measured cashmere fat content values were used for calibration and validation with multivariate analysis (partial least squares combined with support vector machine). The authors analyzed and evaluated denoising effects of these three wavelet threshold denoising models by comparing parameters (R^2 , RMSEC and RMSEP) obtained through calibration and validation of denoised spectra with these three wavelet threshold denoising models respectively. The results show that the three wavelet threshold denoising models all can denoise the infrared spectral signal, increase signal to noise ratio and improve precision of prediction model to some extent; Among these three wavelet threshold denoising models, the denoising effect of Brige-Massart threshold denoising model and default threshold denoising model were significantly better than that of default threshold denoising model; Compared with the prediction precision ($R^2 = 0.793$, RMSEC = 0.233, RMSEP = 0.225) of multivariate analysis model established with original spectra, the prediction precision ($R^2 = 0.882$, RMSEC = 0.144, RMSEP = 0.136) of multivariate analysis model established with spectra denoised by Brige-Massart threshold denoising model and the prediction precision ($R^2 = 0.876$, RMSEC = 0.151, RMSEP = 0.142) both had much more improvements. All the above illustrates that wavelet threshold denoising models can denoise infrared spectral signal effectively, make multivariate analysis model of spectral data and measured cashmere fat values more representative and robust, and so it can improve detection precision of infrared spectral technique.

Keywords Infrared spectroscopy; Wavelet; Threshold; De-noising

(Received Dec. 26, 2008; accepted Mar. 28, 2009)

* Corresponding author