

文章编号: 1001-9014(2009)02-0107-04

# 基于小波变换的噪声及背景同时去除方法 在血糖近红外无创检测中的应用

张广军, 李丽娜, 李庆波, 徐玉坡

(北京航空航天大学 精密光机电一体化技术教育部重点实验室, 北京 100191)

**摘要:** 对于血糖近红外无创检测, 光谱信号中的各种噪声以及水分等物质的强吸收产生的背景信号, 影响了光谱定量校正模型的预测精度. 利用小波变换, 可将光谱信号分解为多尺度的近似成分与细节成分, 根据无用信息变量消除判据可判定代表背景信息的高尺度近似成分及代表噪声的低尺度细节系数, 去除相应的成分即可同时去除光谱信号中的背景与噪声. 本文将这种小波变换与无用信息变量消除判据相结合的预处理方法应用于人体血糖无创检测研究中, 并分析了该方法对不确定因素较多的复杂光谱模型的适用性问题. 实验结果表明应用小波变换结合无用信息变量消除判据的方法可以有效地同时去除血糖无创检测近红外光谱信号中的背景信息和噪声, 提高光谱定量校正模型的预测精度, 对于人体血糖无创检测具有重要应用价值.

**关键词:** 血糖无创检测; 小波变换; 背景; 噪声; 近红外光谱

中图分类号: O657.33 文献标识码: A

## APPLICATION OF DENOISING AND BACKGROUND ELIMINATION BASED ON WAVELET TRANSFORM TO BLOOD GLUCOSE NONINVASIVE MEASUREMENT OF NEAR INFRARED SPECTROSCOPY

ZHANG Guang-Jun, LI Li-Na, LI Qing-Bo, XU Yu-Po

(Precision Opto-mechatronics Technology, Key Laboratory of Education Ministry  
Beihang University, Beijing 100191, China)

**Abstract:** Noises and background produced by the absorption of water in spectra signal are the main interference for prediction accuracy of quantitative calibration in blood glucose noninvasive measurement. By utilizing wavelet transform, the spectra signal can be split into multi-scale approximation and detail component. According to uninformative variable elimination (UVE) criterion, it is possible to distinguish the high-scale approximation component which presents background and low-scale detail coefficients which present noise from original spectra, and then to eliminate the corresponding component, so the background and noise can be eliminated simultaneously. In this study, the pretreatment method of wavelet transform combined with UVE criterion was applied to the research of blood glucose noninvasive measurement based on near infrared spectroscopy, and the adaptability of this pretreatment method to the uncertainty complex spectra model was also discussed. The results indicate that this pretreatment method is effective to eliminate background and noise in near infrared spectra of blood glucose noninvasive measurement. It is helpful to improve the prediction accuracy of quantitative calibration for blood glucose noninvasive measurement.

**Key words:** blood glucose noninvasive measurement; wavelet transform; background; noise; near infrared spectroscopy

### 引言

近红外光谱结合多变量分析技术, 是目前认为

具有发展前景的血糖无创检测技术之一<sup>[1-3]</sup>. 由于人体血液及组织中葡萄糖的正常比例只有水的 0.1%, 生物组织中水的吸收产生的背景信息和仪器测量中

收稿日期: 2008-05-26, 修回日期: 2008-09-25

基金项目: 国家自然科学基金(60708026)资助项目

作者简介: 张广军(1965-), 男, 天津人, 教授, 主要从事精密光机电系统和生物光谱检测技术方面的研究.

Received date: 2008-05-26, revised date: 2008-09-25

不可避免的各种噪声是血糖无创检测中近红外光谱的主要干扰,如何从复杂、重叠的光谱信号中提取微弱的血糖信息,以提高多变量校正模型的预测精度是近红外无创检测血糖的技术难点之一。

采用有效的预处理方法对获得可靠的分析结果具有重要作用<sup>[4,5]</sup>。基于小波变换的背景与噪声同时去除的预处理方法<sup>[6,7]</sup>,是对光谱信号进行小波分解后根据小波变换的线性原理重建各尺度的近似成分和细节成分,根据无用信息变量消除判据(Uninformative Variable Elimination, UVE)<sup>[8,9]</sup>,判断分别代表背景近似成分和代表噪声的细节系数的尺度,去除其相应的成分后进行信号重构,即可实现同时去除光谱信号中的背景与噪声。

本文将小波变换结合 UVE 判据的预处理方法应用于人体血糖无创检测研究中,对人体血浆样品检测实验和人体无创检测实验的近红外光谱数据进行分析,实现去除光谱信号中的背景及噪声干扰,并讨论了该方法对不确定因素较多的复杂光谱模型的适用性问题。

## 1 模型与算法

高频噪声、生物组织成分的光吸收产生的背景信号多分别体现在低尺度细节系数和高尺度的近似成分,因此,可以利用小波变换的多分辨分析特性识别和同时去除这些无用信息,以优化光谱质量。

小波变换结合 UVE 判据方法去除背景及噪声的原理与实现步骤如下:

首先,构造模拟数据<sup>[6,8]</sup>。具体方法是,使用高斯信号模拟包含有背景、有用信号的光谱,如式(1)所示

$$g(t) = a_1 \exp\left(-\left(\frac{t-a_2}{a_3}\right)^2\right), \quad (1)$$

其中,  $a_1$  表示峰强度,  $a_2$  表示峰位,  $a_3$  表示峰宽。根据研究对象,设置各参数使背景表现为较宽的高斯峰,人工信息为较窄的高斯峰,二者产生明显差异,微弱相关。人工噪声用人工信号的最大值 20% 的一个正态分布信号模拟。则建立的模拟数据

$$g = g_i + g_n + g_b, \quad (2)$$

其中,  $g_i$  代表人工信息,  $g_n$  表示人工噪声,  $g_b$  表示背景信息。

其次,分别将原始信号与模拟信号在适当的尺度  $j(j=0, 1, \dots, l)$  进行离散小波变换,即可获得近似成分  $f_j$ 、细节成分  $g_l$  和细节系数  $d^l$ ,并通过式(3)重构近似成分

$$f_{l-1} = g_l + f_l. \quad (3)$$

根据 UVE 判据判定背景与噪声<sup>[6]</sup>。通过式(4)计算不同变量  $b_j$  的  $|t_j|$  值的大小来判定是否为无用信息。

$$|t_j| = \left| \frac{\text{mean}(b_j)}{\text{std}(b_j)} \right|, \quad j = 1, \dots, p \quad (4)$$

本文  $b_j$  代表不同尺度的近似成分  $f_j$  或细节系数  $d^j$ 。若光谱信号低尺度  $d^j$  的  $|t_j|$  值与人工噪声  $d^j$  的  $|t_j|$  值相似,则该系数代表的成分作为噪声剔除;若光谱信号高尺度  $f_j$  的  $|t_j|$  值小于人工信息变量  $f_j$  的  $|t_j|$  值,则其对应的成分作为背景去除。去除无用信息后,重构原始信号<sup>[7]</sup>。

重构光谱信号后,建立多变量校正模型,根据预测均方根误差(Root Mean Square Error of Prediction, RMSEP)选择最优小波基。

## 2 实验部分

### 2.1 人体血浆近红外实验

仪器与试剂:采用美国 Perkin-Elmer 公司 Spectrum GX 傅里叶变换红外光谱仪,液态氮冷却的 InSb 检测器,光谱采集范围 900 ~ 3600nm,共计 8192 个波长,1mm 石英样品池,蠕动泵自动进样系统。采用日本 TBA30 全自动生化分析仪。葡萄糖由天津化学试剂三厂提供,血浆由天津医院提供。

实验步骤:在全血中加入肝素抗凝剂,在离心机中经过 1500rpm, 10min 后分离出,并在血浆中加入葡萄糖使其产生具有一定浓度变化的血糖值,然后用全自动生化分析仪和葡萄糖氧化酶法标定血糖值。

样品数据:血浆实验中共获得 36 个样品,葡萄糖浓度范围为 10.4 ~ 44.4mg/dL,随机分布,浓度标准差为 8.5mg/dL。

### 2.2 人体近红外无创实验

人体口服葡萄糖耐量试验(Oral Glucose Tolerance Test, OGTT)是一种葡萄糖负荷试验。该试验过程能在短时间内提供一个人体血糖浓度的变化范围,为研究在体无创血糖测量提供了一条独特的途径。

仪器与试剂:采用海洋光学 NIR256-2.5 光栅光谱仪、HL-2000 光源、R400-7-VIS-NIR Y 型反射光纤探头,光谱扫描范围 870 ~ 2565nm,扫描分辨率 6.6nm,积分时间 13ms。采用强生稳豪 II 代血糖仪测量血糖参考值。口服葡萄糖由北京燕京药业生产。

实验步骤:受试者为 1 位健康志愿者,实验前禁食 8h,实验时 5min 内饮入 100ml 含 75g 葡萄糖的水溶液,每 10min 采集手指指腹的近红外光谱,实验时保持检测位置、压力、受试者心理状态等尽可能恒定。光纤探头与指腹接触压力约压入 0.5mm,接触

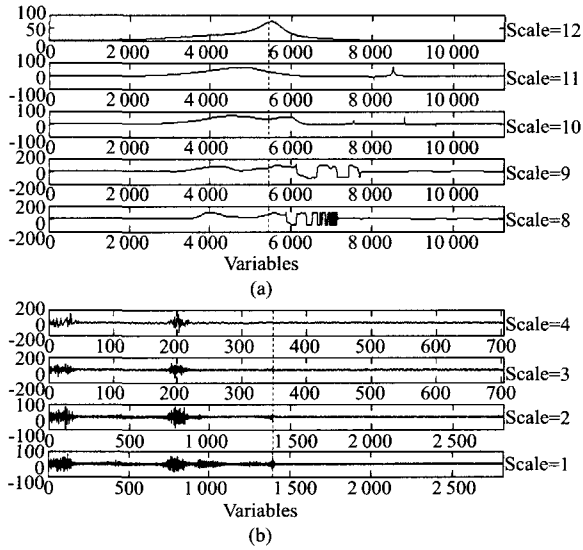


图 1 (a) 血浆近红外实验数据近似成分 t 值 (b) 血浆近红外实验数据细节系数 t 值  
 Fig. 1 (a) t value of approximation components of plasma NIR data (b) t value of detail coefficient of plasma NIR data

30s 后开始采集光谱. 同时采用强生便携式血糖仪测血糖作为建模参考值. 受试者饮入葡萄糖水溶液 1/2h 后血糖值达到最高 158. 4mg/dL, 以后慢慢回落至正常 99mg/dL 时实验结束.

样本数据: 该次实验共获得 11 个样本, 每个样本 255 个数据点. 血糖值范围 77. 4 ~ 158. 4mg/dL, 标准差为 27. 5mg/dL.

### 3 结果与讨论

#### 3.1 实验数据处理

首先, 应用小波变换结合 UVE 判据的方法对血浆近红外实验数据进行预处理. 选择 900 ~ 1900nm 波长范围的血浆近红外实验数据进行预处理. 初始选取小波基函数 db2, 并通过 Mallat 分解, 分解层数为 12. 计算重构的近似成分及细节系数的 t 值如图 1 所示, 图 1(a) 为高尺度近似成分 t 值, 根据 UVE 判据, 在尺度 11, 光谱信号的近似成分(t)值小于人工信息, 而在尺度 10 时大于人工信息, 则尺度 11 对应的成分作为背景去除; 图 1(b) 表示的是光谱信号低尺度细节系数的 t 值, 在尺度 1、2、3 时, 血浆近红外实验数据细节系数的 |t| 值频率较模拟数据高, 故尺度 1、2、3 的细节系数代表的成分可作为噪声消除.

同理, 可对 OGTT 近红外实验数据进行处理. 根据小波变换原理进行信号延拓, 初始选取小波基 db2, 分解尺度为 8, 根据 UVE 判据去除了尺度 1、7

所分别对应的包含噪声和背景的成分.

消除噪声与背景后, 分别重构血浆近红外光谱信号和 OGTT 近红外光谱信号, 再建立偏最小二乘 (Partial Least Squares, PLS) 校正模型, 采用交互验证法评价模型的预测能力, 根据预测值均方根误差可选用最佳小波滤波器. 本文对 db2 ~ db10 进行了分析, 应用不同小波基函数得到的 RMSEP 如图 2 所示. 结果表明, 对于重构血浆近红外光谱信号和 OGTT 近红外光谱信号, 均选用 db3 时的 RMSEP 最小, 认为几种小波基中 db3 为最佳. 图 3 和图 4 所示分别为血浆近红外实验数据和 OGTT 实验数据经预处理的结果, 图中 A 表示吸光度,  $\lambda$  表示波长. 显然, 去除的信息中包含了水的主要吸收峰 1440nm 附近的区域, 认为血糖无创检测中的主要背景干扰来自水的影响.

#### 3.2 复杂光谱模型的适用性分析

与离体血浆实验相比, 人体 OGTT 实验为活体测量, 其获得的光谱数据中包含了更多的不确定因素, 如测量位置、时间、压力以及脉搏、体温等人体生理因素的影响等, 人体无创近红外实验获得的光谱模型更为复杂. 本节主要讨论小波变换结合 UVE 判据的方法对复杂模型的适用性问题. 本文应用小波变换结合 UVE 的方法分别对人体血浆近红外实验数据及人体近红外无创实验数据进行预处理, 数据预处理前后的 PLS 交互验证预测结果如表 1 所列. 通过分析得出如下结论:

(1) 对于两种实验数据, 本文应用的预处理方法与未经预处理的光谱模型分别建立 PLS 回归模型, 预测均方根误差分别减少了 16% 和 52%. 结论表明本文应用的方法能够提高光谱定量校正模型的预测精度.

(2) 对于离体血浆近红外实验数据, 使用预处理

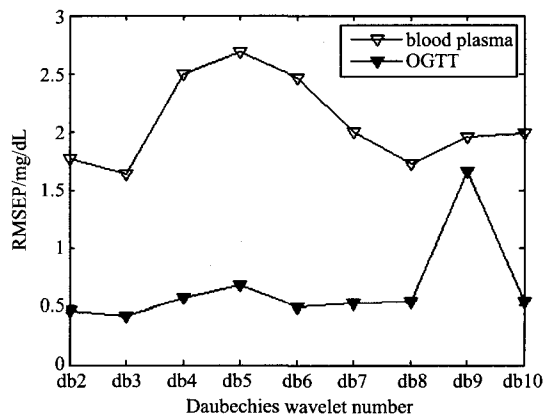


图 2 不同小波基的预测精度分析

Fig. 2 Prediction precision analysis of different wavelet filter

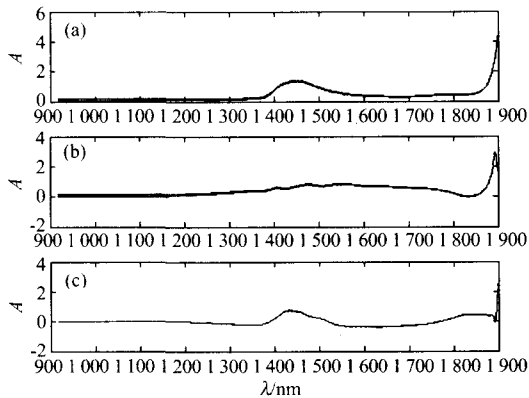


图3 血浆近红外光谱(a)原始光谱 (b)重构信号 (c)消除的无用信息

Fig. 3 Near infrared spectra of blood plasma (a) original spectra (b)reconstructed spectra (c)eliminated information

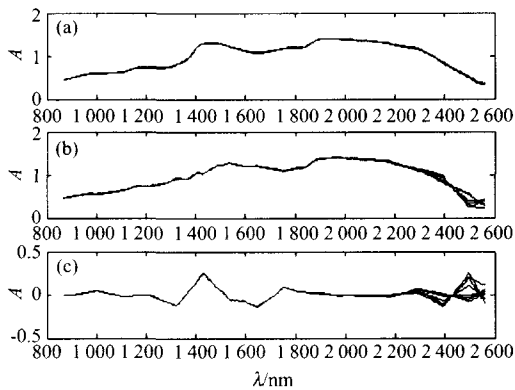


图4 OGTT 近红外光谱(a)原始光谱 (b)重构信号 (c)消除的无用信息

Fig. 4 Near infrared spectra of OGTT(a)original spectra (b)reconstructed spectra (c)eliminated information

表1 人体血浆及人体无创近红外实验数据应用预处理方法前后预测结果比较

Table 1 Prediction results comparison of NIR data of plasma and OGTT before and after using this pre-treatment method

	血浆实验		人体 OGTT 实验	
	RMSEP (mg/dL)	相关系数	RMSEP (mg/dL)	相关系数
未用预处理	1.7	0.97	15.3	0.77
应用预处理	1.5	0.98	7.4	0.95

理方法前后的相关系数分别为 0.97 和 0.98, 差异不明显; 对于光谱模型相对复杂的人体 OGTT 近红外实验数据, 使用预处理方法前后的效果差别明显. 结论表明本文应用的预处理方法对于存在不确定因素较多的在体检测数据建立的 PLS 模型具有较强的适应性.

#### 4 结语

本文通过小波变换结合无用信息变量消除判据的预处理方法, 对人体血浆近红外实验数据和人体无创检测近红外实验数据进行处理, 实验结果表明, 该法能同时去除近红外光谱信号中水的吸收产生的背景信号及噪声, 提高了近红外方法无创测量血糖的定量校正模型精度, 且对于存在不确定因素较多的复杂光谱模型具有较好的适应性, 对血糖无创近红外检测的研究具有重要意义.

#### REFERENCES

- [1] LI Qing-Bo, XU Ke-Xin, WANG Yan. Primary discussion on prerequisites to noninvasive blood glucose [J]. *Journal of Tianjin University* (李庆波, 徐可欣, 汪曦. 人体血糖浓度无创检测的必要测量条件. *天津大学学报*), 2003, **36**(2): 139—142.
- [2] LI Gang, WANG Yan, LI Qiu-Xia, et al. Theoretic study on improving noninvasive measurement accuracy of blood component by dynamic spectrum method [J]. *J. Infrared Millim. Waves* (李刚, 王焱, 李秋霞, 等. 动态光谱法对提高近红外无创血液成分检测精度的理论分析. *红外与毫米波学报*), 2006, **25**(5): 345—348.
- [3] LI Qing-Bo, ZHANG Guang-Jun, XU Ke-Xin, et al. Application of digital fourier filtering pretreatment method to improving robustness of multivariate calibration model in near infrared spectroscopy [J]. *Spectroscopy and Spectral Analysis* (李庆波, 张广军, 徐可欣, 等. 应用数字傅立叶滤波方法提高近红外光谱多元校正模型稳健性的研究. *光谱学与光谱分析*), 2007, **27**(8): 1484—1488.
- [4] LIU Yan-De, LUO Ji, CHEN Xing-Miao. Analysis of soluble solid content in nanfeng mandarin fruit with visible near infrared spectroscopy [J]. *J. Infrared Millim. Waves* (刘燕德, 罗吉, 陈兴苗. 可见/近红外光谱的南丰蜜桔可溶性固形物含量定量分析. *红外与毫米波学报*), 2008, **27**(2): 119—122.
- [5] ZHU Shi-Ping, WANG Gang, YANG Fei, et al. Rapid detection method of the spicy components in zanthoxylum bungeagum maxim by near infrared spectroscopy [J]. *J. Infrared Millim. Waves* (祝诗平, 王刚, 杨飞, 等. 基于近红外光谱的花椒麻味物质快速检测方法. *红外与毫米波学报*), 2008, **27**(2): 129—132.
- [6] Chen Da, Shao Xueguang, Hu Bin, et al. A Background noise elimination method for quantitative calibration of near infrared spectra [J]. *Analytica Chimica Acta*, 2004, **511**(1): 37—45.
- [7] Tan Hu-Wei, Brown S D. Wavelet analysis applied to removing non-constant varying spectroscopic background in multivariate calibration [J]. *Journal of Chemometrics*, 2002, **16**(5): 228—240.
- [8] Centner V, Massart D L, De Noord O E, et al. Elimination of uninformative variables for multivariate calibration [J]. *Analytical chemistry*, 1996, **68**(21): 3851—3858.
- [9] Jouan-Rimbaud D, Walczak B, Poppi R J, et al. Application of wavelet transform to extract the relevant component from spectral data for multivariate calibration [J]. *Analytical chemistry*, 1997, **69**(21): 4317—4323.