

## 土壤导数光谱小波去噪与有机质吸收特征提取

刘 炜<sup>1</sup>, 常庆瑞<sup>1\*</sup>, 郭 曼<sup>1</sup>, 邢东兴<sup>1,2</sup>, 员永生<sup>1</sup>

1. 西北农林科技大学资源环境学院, 陕西 杨凌 712100
2. 咸阳师范学院资源环境系, 陕西 咸阳 712000

**摘 要** 使用高光谱仪 ASD Field Spec 在波长范围 400~1 000 nm 内, 采集有机质含量水平不同的土壤反射光谱数据, 之后求取一阶导数(一阶导数光谱)并进行小波去噪; 通过数值积分计算、相关分析, 从去噪后的一阶导数光谱中提取吸收面积表征有机质含量变化。结果显示: (1) 土壤一阶导数光谱中含有大量噪声, 致使光谱曲线轮廓和有机质的吸收特征难以识别。(2) 当小波分解层数  $J=3$  时, 小波去噪在平滑土壤一阶导数光谱曲线的同时, 较好的保持了曲线上多个敏感波段的光谱响应特征。(3) 从去噪后的一阶导数光谱中提取的吸收面积  $S(538, 586)$  与有机质含量的相关系数为  $-0.8763$ , 较好的反映了土壤有机质含量的变化情况, 可以用于土壤有机质含量测算。

**关键词** 可见/近红外光; 土壤有机质; 数值积分; 特征提取

**中图分类号:** S15 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2011)01-0100-05

### 引 言

土壤有机质是土壤中有机的总称, 它不仅为植物和微生物的生命活动提供养分和能量, 而且能够促进土壤结构形成, 改善土壤物理性状, 提高土壤保肥能力。传统的土壤有机质化学测定方法虽然测定精度较高, 但耗时、费力、成本高, 并且存在着有害、污染等问题, 因而难以满足现代农业生产需求。土壤高光谱数据具有高分辨率特性, 其在经过一阶求导运算后, 可以增强光谱曲线上局部位置的细微差别, 并能够提供精细的光谱形态变化信息。因此, 可以从土壤一阶导数光谱中寻找有机质的吸收特征, 进而提取特征参数用于有机质含量的快速测算<sup>[1-5]</sup>。

近年来, 国内外关于土壤反射光谱的大量研究表明, 可见/近红外光谱段(400~1 000 nm)是土壤有机质最主要的光谱响应区域。Kishnan 等<sup>[6]</sup>从四种类型的土壤反射光谱中寻找有机质的吸收特征, 结果表明可见光谱段 564 和 623 nm 是探测有机质含量的敏感波段; 近红外光谱段没有有机质吸收引起的反射峰、吸收谷。Sudduth 等<sup>[7]</sup>结合差光谱与比光谱的算法, 从可见/近红外光谱段土壤反射光谱中, 选取了 8 个波段的吸光度, 建立了对土壤有机质含量的检测模型。

Ben-Dor<sup>[8]</sup>、李民赞<sup>[1]</sup>、刘伟东<sup>[9]</sup>、谢伯承<sup>[10]</sup>、刘焕军<sup>[11]</sup>等则利用相关分析和回归分析, 从土壤一阶导数光谱中提取有机质含量、全氮含量以及水分含量的信息。

大量研究表明, 高光谱数据经求导运算后会引入大量噪声, 导致光谱信噪比下降, 光谱曲线轮廓和关于目标物质的吸收特征难以识别<sup>[1-4]</sup>。目前, 光谱去噪多采用移动平均法(moving average), 然而, 该方法不能有效区分光谱信号中的高频噪声与细节特征, 致使在平滑曲线的同时, 往往会损失掉光谱信号中锐变尖峰成分所携带的重要信息。因此, 选择适当的去噪方法是改善一阶导数光谱分析精度, 提升一阶导数光谱分析能力的一个关键突破口。此外, 土壤有机质的构成成分十分复杂, 功能团多样, 其光谱特性并不适宜采用几个单波段的光谱反射率加以解释<sup>[5, 12]</sup>。

小波变换(wavelet transform)可以将一维空间域信号映射到二维空间——频率域, 并能根据需要以任意尺度的空域或频域分辨率来观察信号。基于小波变换的阈值去噪方法利用目标信号与噪声的频率差异, 在含噪信号不同的小波分解尺度上通过设定阈值来消除噪声, 该方法在去除噪声的同时能够较好的保留目标信号中的细节特征。吸收面积(absorb area)是指在敏感波长范围内, 光谱反射率与对应波长范围包围的面积<sup>[1, 10]</sup>。由于覆盖了较多的敏感波段数, 包含了更多

收稿日期: 2010-03-02, 修订日期: 2010-06-06

基金项目: 国家“十一五”科技计划支撑重大项目(2006BAD09B0603), 国家自然科学基金项目(30872073)和国家(973 计划)项目(2007CB407203)资助

作者简介: 刘 炜, 1978 年生, 西北农林科技大学资源环境学院在读博士研究生 e-mail: york5588@nwsuaf.edu.cn

\* 通讯联系人 e-mail: chqr@nwsuaf.edu.cn

的光谱反射能量,因而,相对于单波段光谱反射率,吸收面积受光谱测试环境中不确定因素的影响更小,具有更强的抗干扰能力,更好的稳定性。目前,对土壤一阶导数光谱进行小波去噪,并从去噪后的一阶导数光谱中提取特征参数检测有机质的研究尚不多见。为此,试验选取基于小波变换的阈值去噪方法对土壤一阶导数光谱进行去噪处理<sup>[13]</sup>,并从去噪后的一阶导数光谱中提取具有合适波长覆盖范围的吸收面积表征有机质含量变化。研究旨在为实现土壤有机质导数光谱检测,以及提高高光谱参数准确性和实用性方面提供依据。

## 1 材料与方 法

### 1.1 样品采集与制备

供试土壤样品采自陕西省眉县,采样区土壤为褐土,质地为壤质粘土,土层深厚。实验依据土壤剖面发生层次,分层采集土壤样品,采样深度为 0~60 cm。为了从光谱数据中消除或降低土壤水分、土壤粒度等因素对土壤有机质吸收特征的影响,将土壤样品置于实验室内自然风干,之后用木棍滚压,并去除沙砾石块及植物残体,接下来研磨、过 100 目尼龙筛。每个土壤样品分成两份,一份采用重铬酸钾法测定土壤有机质含量,另一份用来测量光谱数据。

### 1.2 光谱测量

使用高光谱仪 ASD Field Spec 在波长范围 400~1 000 nm 内连续测量经预处理后的土壤样品的反射光谱数据,光谱采样间隔 1.4 nm,重采样间隔 1 nm。测量光谱前将土壤样品放置在直径 16 cm,深度 3 cm 的盛样皿上,调整盛样皿使其处于水平位置,平整土样表面使样品厚度均匀。暗室内测试,光源为能够提供平行光的 1 000 W 的镁光灯,距土壤表层中心 70 cm,光源照射方向与垂直方向的夹角为 15°。经多次实验后光纤探头的视场角选定为 7.5°,置于离土样表面 40 cm 的垂直上方接收光谱数据。测试前以白板定标,每个土壤样品采集 10 条光谱数据,然后将其算术平均值作为该土壤样品的实际光谱反射数据。对土壤光谱求取了一阶导数,有关求导运算参考文献[1]和文献[4]。

## 2 结果与分析

### 2.1 不同小波分解层数下土壤一阶导数光谱去噪处理

图 1 显示了可见/近红外光谱段波长范围 400~1 000 nm 内,有机质含量不同(25.67, 9.37, 7.31, 4.20  $\text{g} \cdot \text{kg}^{-1}$ )的土壤反射光谱曲线。从中可以看出,在沿着波长增加的方向上,土壤光谱反射率大致以线性趋势从 8% 增长至 40%,其间没有反映样品信息,突出的反射峰、吸收谷。对于不同的有机质含量水平,该波长范围内光谱反射率随有机质含量水平的提高呈下降趋势,表现出负相关性。整体而言,400~1 000 nm 内,土壤有机质含量的变化可以从土壤反射光谱曲线的分异表现中得到一定程度的反映,但有机质含量不同的各条光谱曲线之间距离较近,形态相似,没有显著的峰谷特征。

接下来,求取土壤反射光谱的一阶导数(一阶导数光谱),如图 2 所示。从中可以看出,波长范围 400~1 000 nm 内,一阶导数光谱存在大量噪声,光谱曲线的轮廓及其对有机质含量变化的响应特征都难以识别。为此,采用基于小波变换的阈值去噪方法(以“sym 8”作为小波母函数,选择“Heursure”阈值选取规则和“sln”阈值调整方法<sup>[13]</sup>),对一阶导数光谱进行去噪处理,图 3(a)~(e)分别显示了在波长范围 450~650 nm 内,当小波分解层数  $J=1\sim 5$  时的去噪结果。

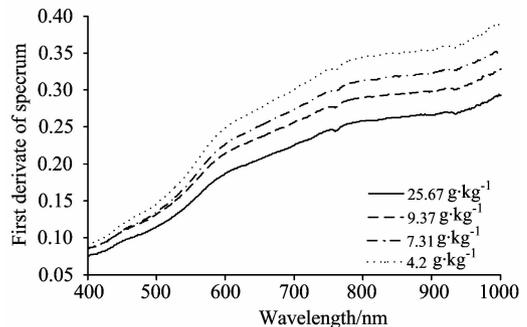


Fig. 1 Soil spectra under different organic content levels

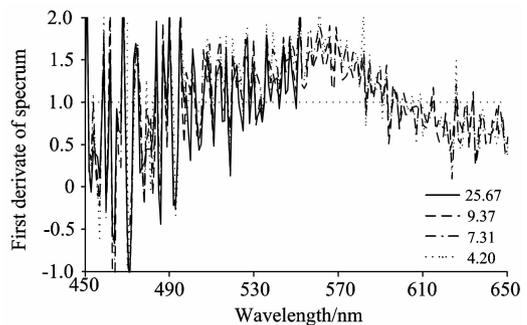


Fig. 2 First of derivative of soil spectra under different organic content levels

从图 3 中可以看出,  $J=1$  时,一阶导数光谱的去噪程度较低,光谱曲线上仍保留有大量噪声,其轮廓走向以及关于有机质含量变化的响应特征仍难以识别;  $J=2$  时,光谱噪声得到了一定程度的去除,可以大致识别出光谱曲线的轮廓。但波长范围 450~650 nm 内,光谱曲线振荡的频率较高、幅度较大,不利于特征参数提取;  $J=3$  时,一阶导数光谱中大部分噪声被有效去除,去噪后的光谱曲线呈开口向下,较为平滑的喇叭口状。同时还可以看到,去噪后的光谱曲线上还保留有多个反射峰、吸收谷,其特征凸显,细致可辨,所处的波长位置也能够被准确定位。尤其是在 562 nm 处,各条光谱曲线都一致达到了最大值,并且在该波长位置,对不同的有机质含量水平,一阶导数光谱值表现出了突出的响应特征。此外,还有 529, 547, 553, 571, 577 nm 等多个位置处的波峰、波谷或拐点,其一阶导数光谱值也具有显著的响应差异。  $J=4$  或 5 时,去噪后的一阶导数光谱曲线表现的十分光滑,显然在去除噪声的同时,光谱曲线上大部分的极值点也被平滑掉了。

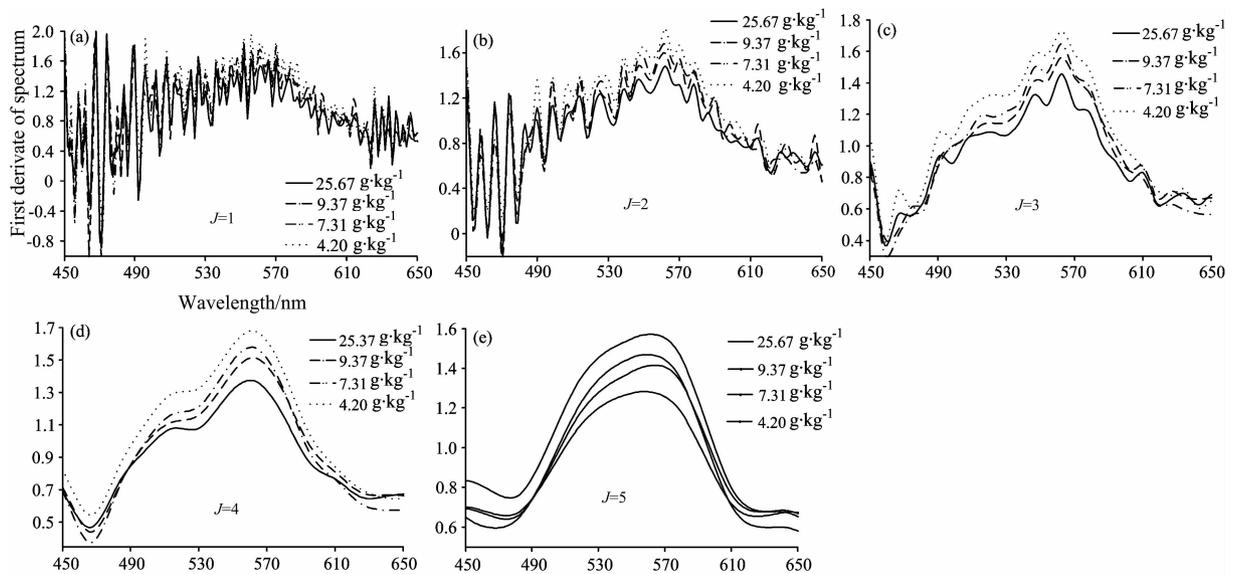


Fig. 3 Denoised first derivative of soil spectra under different soil matter organic content levels and different wavelet decomposition scales

## 2.2 不同小波分解层数及不同积分区间下土壤一阶导数光谱吸收面积的分析与计算

吸收面积是指在敏感波长范围内, 光谱反射率与对应波长范围包围的面积, 通过数值积分运算可以求得该面积值。计算吸收面积时, 需要将指定的波长范围设为积分区间, 这是求得吸收面积的一个关键因素。因为积分区间覆盖的波长范围过窄, 会使得吸收面积包含的敏感波段数量不足, 对光谱信息利用不充分; 而积分区间覆盖的波长范围过宽, 则会导致吸收面积包含数量较多的非敏感波段, 与有机质含量的相关性下降。

从图 3(c)中可以看出, 波长范围 450~650 nm 内, 去噪后的一阶导数光谱曲线呈开口向下的喇叭口状, 并且曲线上分布着多个极值点。这些极值点对应的一阶导数光谱值对不同的有机质含量水平, 都具有显著的特征差异, 特别是在喇叭口中间区域 562 nm 处左右, 光谱响应差异最为突出。然而, 当波长位置从喇叭口的中间区域, 分别向短波方向或长波方向移动时, 一阶导数光谱值会快速下降, 同时因有机质含量不同引起的光谱响应差异也在逐渐缩小, 至喇叭口短波边缘区域 450~460 nm 以及长波边缘 635~650 nm 内, 光谱响应差异基本消失。有鉴于此, 实验将波长范围  $L_1$  (561, 563),  $L_2$  (560, 564),  $L_3$  (559, 565)⋯,  $L_{72}$  (490, 634) 排成一个序列, 然后将该序列的每个波长范围设为积分区间, 求取对应的吸收面积

$$S_{ij}(562-m, 562+m) = \int_{L_m} D_{ij}(\lambda) \cdot d\lambda \quad (1)$$

$$i = 1, 2, \dots, k; j = 1, 2, 3, 4, 5; m = 1, 2, \dots, 72$$

式中,  $D_{ij}(\lambda)$  为小波去噪后的一阶导数光谱值;  $i = 1 \sim k$  分别代表  $k$  个有机质含量水平;  $j = 1 \sim 5$  分别代表去噪过程中小波分解的层数;  $m = 1, 2, \dots, 72$  对应于积分区间  $L_1 \sim L_{72}$ 。

通过式(1)可以看出, 一阶导数光谱在采取不同的小波

分解层数去噪处理后, 每个有机质含量水平对应 72 个吸收面积。这些吸收面积的积分区间以  $L_1$  (561, 563) 为中心, 每次同时分别在短波方向上递减 1 nm 和长波方向上增加 1 nm, 逐步扩展至  $L_{72}$  (490, 634), (也就是说, 上述序列中, 积分区间的覆盖宽度以 2 nm 为起点, 2 nm 为步长, 经 71 次累加增至 144 nm)。图 4(a)~(e) 显示了不同有机质含量水平下, 当积分区间逐步扩展时, 吸收面积值的变化趋势。从中可以看出, 小波分解层数  $J=1 \sim 5$  时, 吸收面积值随积分区间的逐步扩展, 都大致呈现出增长趋势。其中, 当小波分解层数  $J=1$  或 2 时, 随着积分区间向喇叭口两侧边缘区域逐步扩展, 吸收面积曲线出现了幅度越来越大的振荡, 这与在这个小波分解层数下, 位于喇叭口的两侧边缘区域较高强度的噪声有关。而当小波分解层数  $J=3 \sim 5$  时, 吸收面积值的增长趋势相对平稳。

实验计算了当积分区间逐步扩展时吸收面积与有机质含量的相关系数, 结果如图 5(a)~(e) 所示。从中可以看出, 小波分解层数  $J=1$  或 2 时, 相关系数曲线波动频繁, 起伏强烈, 表现不够稳定。小波分解层数  $J=4$  或 5 时, 相关系数曲线呈开口向上的抛物线状, 并且随积分区间的逐步扩展, 相关系数的绝对值从 0.880 连续下降至 0.100。而当小波分解层数  $J=3$ , 积分区间取  $L_1$  (561, 563)~ $L_{44}$  (518, 606) 时, (即积分区间的覆盖宽度从 2 nm 增至 88 nm 时), 相关系数曲线在 -0.770~-0.880 之间小幅波动, 其中, 当积分区间取波长范围 518~606 nm 的中间区域  $L_{15}$  (547, 577)~ $L_{30}$  (532, 592) 时, (即积分区间的覆盖宽度从 30 nm 增至 60 nm 时), 相关系数的绝对值稳定在 0.860~0.880 之间, 尤其是当积分区间取  $L_{24}$  (538, 586) 时, (也就是积分区间的覆盖宽度为 48 nm 时), 相关系数达到了负的最小值 -0.876 3。鉴于在该小波分解层数下, 较宽波长范围 538~586 nm 内, 相关系数值在较高的水平上, 具有十分稳定的表现, 作者认为, 当小波分解层数  $J=3$ , 积分区间取  $L_{24}$  (538, 586) 时, 对

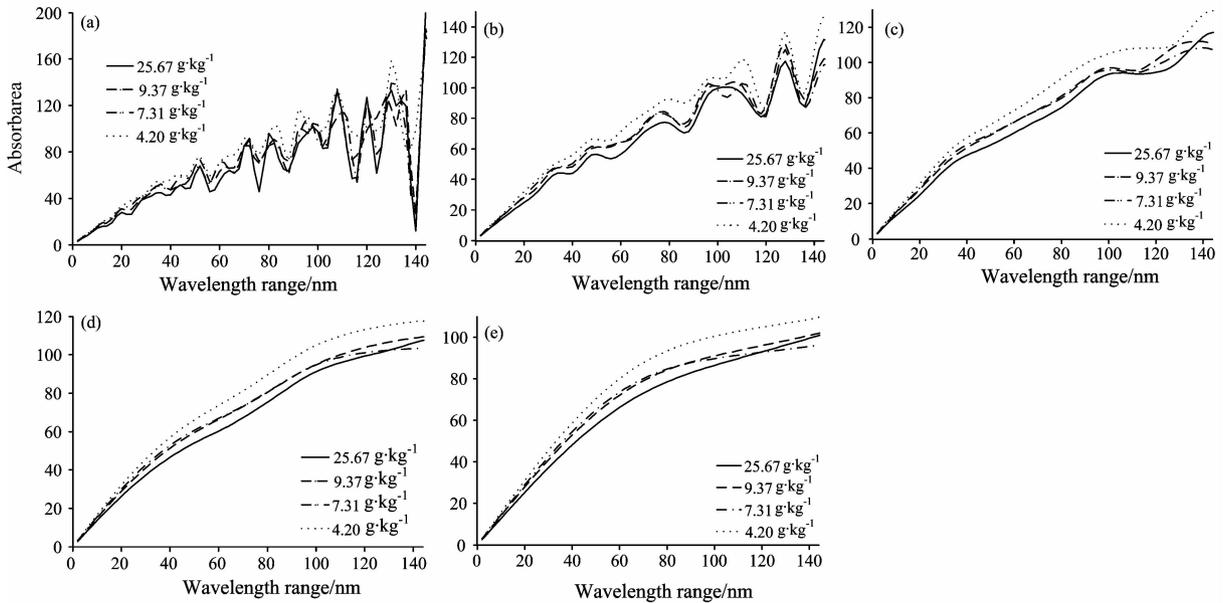


Fig. 4 Absorption area under different wavelength ranges and at different wavelet decomposition scales

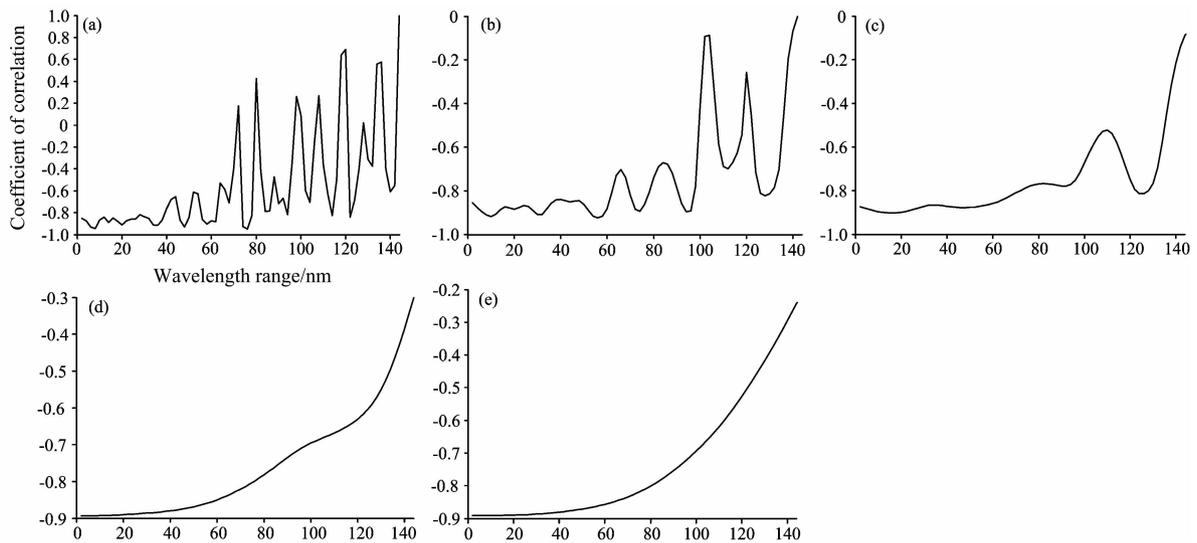


Fig. 5 Correlation coefficients between absorption areas and soil organic matter contents under different wavelength ranges and different wavelet decomposition scales

应的吸收面积  $S(538, 586)$  最适合用作光谱指数测算土壤有机质含量。

### 3 结论与讨论

(1) 土壤高光谱数据经一阶求导运算后含有大量噪声, 致使光谱曲线轮廓和有机质的吸收特征难以识别。

(2) 当小波分解层数  $J=3$  时, 基于小波变换的阈值去噪方法在平滑土壤一阶导数光谱曲线的同时, 较好的保持了曲

线上的多个敏感波段的光谱响应特征。

(3) 从去噪后的一阶导数光谱中提取的吸收面积  $S(538, 586)$  与有机质含量的相关系数  $-0.8763$ , 较好地反映了土壤有机质含量的变化情况。

高光谱曲线上的峰谷特征是探测目标物质各项组成成分的关键区域<sup>[1-4]</sup>, 除一阶求导运算外, 还可以通过更高阶次的导数运算, 从土壤高光谱数据中挖掘有机质的吸收特质, 进而提供具有更高可靠性和准确性的特征参数, 这将是本试验下一步研究工作的重点。

## References

- [1] LI Min-zan, HAN Dong-hai, WANG Xiu(李民赞, 韩东海, 王 秀). Spectral Analysis Technology and Its Application(光谱分析技术及其应用). Beijing: Science Press(北京: 科学出版社), 2006. 115.
- [2] LU Wan-zhen, YUAN Hong-fu, XU Guang-tong, et al(陆婉珍, 袁洪福, 徐广通, 等). Modern Near Infrared Spectroscopy Analytical Technology(现代近红外光谱分析技术). Beijing: China Petrochemical Press(北京: 中国石化出版社), 2007. 35.
- [3] CHU Xiao-li, YUAN Hong-fu, LU Wan-zhen(褚小立, 袁洪福, 陆婉珍). Progress in Chemistry(化学进展), 2004, 16(4): 528.
- [4] WANG Ji-hua, ZHAO Chun-jiang, HUANG Wen-jiang(王纪华, 赵春江, 黄文江). Quantitative Remote Sensing Technique and Application in Agriculture(农业定量遥感基础与应用). Beijing: Science Press(北京: 科学出版社), 2008. 32.
- [5] SHEN Run-ping, DING Guo-xiang, WEI Guo-shuan, et al(沈润平, 丁国香, 魏国栓, 等). Acta Pedologica Sinica(土壤学报), 2009, 46(3): 391.
- [6] Krishnan P, Alexander J D, Butler B J, et al. Soil Science Society of American Journal, 1980, 44: 1282.
- [7] Sudduth K A, Hummel J W. Transaction of the ASAE, 1991, 34(4): 1900.
- [8] Ben-Dor E, Banin A. Remote Sensing of Environment, 1994, 48: 261.
- [9] LIU Wei-dong, Baret F, ZHANG Bing, et al(刘伟东, Baret F, 张 兵, 等). Acta Pedologica Sinica(土壤学报), 2004, 41(5): 700.
- [10] XIE Bo-cheng, XUE Xu-zhang, LIU Wei-dong, et al(谢伯承, 薛绪掌, 刘伟东, 等). Acta Pedologica Sinica(土壤学报), 2005, 42(1): 171.
- [11] LIU Huan-jun, ZHANG Bai, ZHAO Jun(刘焕军, 张 柏, 赵 军, 等). Acta Pedologica Sinica(土壤学报), 2007, 44(1): 27.
- [12] PENG Jie, ZHANG Yang-zhu, ZHOU Qing, et al(彭 杰, 张杨珠, 周 清, 等). Soils(土壤), 2006, 38(4): 453.
- [13] WU Gui-fang, HE Yong(吴桂芳, 何 勇). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2009, 29(12): 3246.

## Extraction of First Derivative Spectrum Features of Soil Organic Matter via Wavelet De-Noiseing

LIU Wei<sup>1</sup>, CHANG Qing-rui<sup>1\*</sup>, GUO Man<sup>1</sup>, XING Dong-xing<sup>1, 2</sup>, YUAN Yong-sheng<sup>1</sup>

1. College of Resources and Environment, Northwest A&F University, Yangling 712100, China

2. Department of Resources Environment, Xianyang Normal College, Xianyang 712000, China

**Abstract** The hyperspectral reflectance of soil was measured by a ASD FieldSpec within 400~1 000 nm. Next, its first derivative of spectra were acquired and de-noised by the threshold de-noising method based on wavelet transform. From the de-noised derivative spectra, absorption areas used as indicators for soil organic matter content were acquired by numerical integration. Results show that: (1) Because of much noise, it is difficult to identify spectra contour and features in the first derivative of soil spectra resulting from different organic content levels. (2) When the scale of wavelet decomposition was 3, the threshold de-noising method based on wavelet transform can keep the balance between smoothing curve and holding spectra features. (3) Absorption area  $S(538, 586)$  is extracted from de-noised first derivative of soil spectra, and the coefficient of correlation between it and organic matter content is 0.896 3.

**Keywords** Vis/NIR spectrum; Soil organic matter; Numerical integration; Feature extraction

(Received Mar. 2, 2010; accepted Jun. 6, 2010)

\* Corresponding author