

利用可见/近红外光谱测定小麦叶面积指数的改进研究

梁亮^{1,2}, 杨敏华¹, 臧卓²

(1. 中南大学信息物理工程学院, 湖南长沙 410083; 2. 中南林业科技大学林业遥感信息工程研究中心, 湖南长沙 410004)

摘要: 改进了小麦叶面积指数的可见/近红外光谱测定模型。以不同方法实现了小麦冠层反射光谱的预处理, 并采用偏最小二乘回归算法(PLS)建立小麦叶面积指数估测模型对其进行比较分析, 发现小波除噪结合一阶导数能最有效地消除原始光谱的噪声与背景信息, 此时 PLS 模型校正集与预测集 R^2 分别为 0.849 与 0.835。为进一步优化模型, 对经一阶导数结合小波除噪后的光谱采用主成分分析法(PCA)降维, 以前 4 个主成分(含原始光谱 84.867% 特征信息)为输入变量, 采用小二乘支撑向量机回归算法(LS-SVR)建立了小麦叶面积指数估测模型, 其校正集与预测集 R^2 分别达 0.905 与 0.883, 具有比 PLS 算法更高的精度。结果表明: 以小波除噪结合一阶导数去除小麦冠层反射光谱中的土壤背景信息以提高模型精度是可行的, 且 LS-SVR 是建模的优选方法。

关键词: 可见/近红外光谱; 叶面积指数; 小波除噪; 导数; 支撑向量回归; 小麦
中图分类号: TP79; S123 **文献标识码:** A

Improvement research of wheat leaf area index measurement using visible/near-infrared spectra

LIANG Liang^{1,2}, YANG Min-hua¹, ZANG Zhuo²

(1. School of Info-Physics and Geomatics Engineering, Central South University, Changsha 4100831, China;
2. Research Center of Forestry Remote Sensing and Information Engineering, Central South University of Forestry and Technology, Changsha 410004, China)

Abstract: The model to measure the leaf area index (LAI) of wheat with visible/near-infrared reflectance spectra is improved. The wheat canopy reflectance spectra is pretreated by different methods, and then the LAI estimation models are established by partial least square (PLS) algorithm to comparative analysis different pretreatments. It is found that the pretreatment method of wavelet denoising combined with first derivative can eliminate the noise and background information of the original spectra most effectively, with the calibration R-square 0.849 and prediction R-square 0.835, respectively. For optimizing the model, the pretreated spectra are analyzed using principal component analysis (PCA), and the anterior 4 principal components, which accounted for 84.867% variation of the original spectral information, are used as the input variables to built the LAI estimation model by least square support vector regression (LS-SVR) algorithm. The calibration R-square and prediction R-square of LS-SVR model are 0.905 and 0.883, respectively, higher than that of PLS model, which indicates that the LS-SVR model is more accurate. The results suggest that it is feasible to improve the accuracy of the LAI estimation model by eliminating the soil background information of original spectra with the pretreatment method of wavelet denoising combined with first derivative, and the LS-SVR algorithm is a preferred method for model building.

Key words: visible/near-infrared spectra; leaf area index (LAI); wavelet denoising; differential coefficient; support vector regression (SVR); wheat

基金项目: 国家自然科学基金项目(No. 30570279); 国家发展与计划委员会项目(No. A00300100584); 中南林业科技大学林业遥感信息工程研究中心开放性研究基金项目(No. RS2008k03); 中南大学拔尖博士研究生学位论文创新项目(No. 1960-71131100007); 中南大学优秀博士学位论文扶持项目(No. 2008yb024)资助。

作者简介: 梁亮(1981-), 男, 博士生, 主要从事高光谱遥感研究。E-mail: liangliang198119@163.com

收稿日期: 2010-07-02; **修订日期:** 2010-08-27

1 引言

叶面积指数与植物的光合能力密切相关,是反映植物群体大小的良好指标,系农学、生态学与气象学重要的研究对象之一^[1-2]。叶面积指数的传统测定方法(如干重法)存在耗时长、耗费大且需损伤性采样测试等缺陷。作为一种非损伤性手段,以可见/近红外光谱技术获取农作物的叶面积指数等理化参量近年来倍受关注^[3-8]。但在野外进行光谱采集的过程中,所采集的光谱信号中往往包含有土壤等背景信息,成为影响建模精度的干扰因子。虽已有研究表明导数处理可有效去除植被光谱中的土壤背景值,使估计出的生理生化参量更为可靠^[9-10],但这一方法对噪声很敏感,在去除土壤背景信息的同时会加强光谱的噪声信息而导致信噪比降低^[11]。因此,如何在消除土壤等背景信号的同时抑制噪声,从而提高模型的精度,是一有待解决的问题。

小波分析是进行信号处理的有力工具^[12]。自Donohod等提出小波阈值去噪法以来,小波分析在信号降噪方面获得了广泛的应用^[13-15]。如果能引入小波分析技术,将小波降噪与导数处理相结合以进行小麦田间光谱数据的预处理,将可实现在去除土壤背景的同时抑制随机噪声,提高光谱信号的信噪比,使小麦叶面积的可见/近红外光谱测定获得更理想的结果。而在建模方法上,基于结构风险最小化原理的支撑向量机算法(support vector machine, SVM)表现出许多特有的优势,已在高分辨率光谱的分析中获得了成功的应用^[16-17]。故本文拟结合小波消噪与导数处理实现对小麦田间光谱的噪声消除与背景去除,并以最小二乘支撑向量机回归算法(least squares support vector machine regression, LS-SVR)建立小麦叶面积指数测定模型,从而实现小麦叶面积指数的快速、无损与精确测定。

2 材料与方法

2.1 田间胁迫试验

实验区位于北纬 $40^{\circ}10'31'' \sim 40^{\circ}11'18''$,东经 $116^{\circ}26'10'' \sim 116^{\circ}27'05''$ 。为保证小麦叶面积指数有较大的变化范围以便于建立估测模型,对小麦进行氮肥胁迫实验,将实验田12个小区(面积为 $60\text{ m} \times 60\text{ m}$)内的小麦分为6个梯度(每个梯度设置两个重复小区)进行施肥,其氮肥施用量分别为 0 kg/hm^2 (N1), 75 kg/hm^2 (N2), 150 kg/hm^2 (N3), 225 kg/hm^2 (N4), 300 kg/hm^2 (N5)与 375 kg/hm^2 (N6)。

2.2 数据采集

以ASD公司出产的FieldSpec Pro FR地物光谱仪采集拔节期小麦的冠层光谱(仪器光谱范围 $350 \sim 2500\text{ nm}$,在 $350 \sim 1000\text{ nm}$ 波段光谱分辨率 3 nm ,采样间隔为 1.4 nm ;在 $1000 \sim 2500\text{ nm}$ 波段光谱分辨率为 10 nm ,采样间隔为 2 nm)。光谱采集时天气晴朗,风力小于3级,无卷云与浓积云。时间范围规定为地方时(此处即为北京时间) $10:00 \sim 15:00$,以保证有较高的太阳高度角。测量过程中,光谱采集人员着深色服装,面向太阳立于目标区后方。传感器探头(25° 视场角)垂直向下,高度保持在冠层上方 1.3 m 附近,每一样本重复测量10次取均值,且每半小时用参考板对仪器进行一次校正,以消除环境变化所带来的影响。

光谱采集时进行同步采样,将叶片带回实验室进行叶面积指数测定,作为建模的参考数据。叶面积指数的测定采用干重法,即对同一处理取 $50 \sim 100$ 片叶子进行面积测量后,烘干称重,再根据被测对象的干重反推出叶面积,并以叶面积仪(C1-203型)进行校正。试验共采集样本174份,随机抽取142份作为训练集,其余32份作为预测集。光谱数据在ASD View Spec Pro中以ASCII码形式导出,后续处理分别在Unscramble 9.8与Matlab 7.0中进行。

2.3 光谱预处理方法比较分析

田间采集的原始小麦反射光谱包括三方面的信息。其一是目标物(小麦)的光谱信号,系目标物在一定光谱区域内所表现出的特征信息;其二是背景信息,主要由土壤等下垫面所引起;其三是随机噪声,主要由仪器的暗电流以及环境中无法控制的变化所引起。在这三方面的信息中,目标物的光谱信号系需要提取的对象,后两者则为需要消除的对象。本文首先采用小波降噪对光谱进行处理(小波各参数选择方案经正交试验确定为:小波类型取haar,分解层数为5,阈值方案选择Fixed form threshold,噪声结构定为Unscaled white noise),并与传统的平滑与滤波方法进行比较以选取较优的噪声去除方法。然后将降噪后的光谱进行导数处理,并与基线校正、标准化、归一化以及多元散射校正等方法进行比较以确定较优的背景去除方法。

2.4 建模方法比较分析

SVM是一种基于结构风险最小化原理的机器学习算法,可在保证训练精度的同时降低学习机器

的复杂度以获得较好的泛化能力,在解决小样本、非线性等问题方面具有突出的优势^[18]。而利用可见/近红外光谱测定农作物理化参量的研究,其最终目的正在于获得一个具有良好预测能力的模型。因此,本文尝试利用 SVM 中的 LS-SVR 算法建立小麦叶面积指数测定模型,并与传统的偏最小二乘算法(PLS)进行比较分析,以实现建模方法的优化选择。

3 结果与分析

3.1 小麦的可见/近红外反射光谱特征

图 1 为小麦的可见/近红外反射光谱曲线。其中 1770 ~ 1940 nm 与 2350 ~ 2500 nm 波段信噪比过低,已去除。由图所示,由于叶绿素对蓝光的强吸收作用,小麦冠层光谱在 350 ~ 500 nm 反射率较低。此后,因叶绿素对绿光的吸收相对较低,反射率逐渐上升,并在 560 nm 附近达到峰值(即绿峰)。随后,因叶绿素对红光的强吸收作用,反射率又开始下降,在 670 nm 附近形成波谷。而从 690 nm 后,由于小麦叶片中海绵状叶肉细胞对近红外光的强反射功能^[19],反射率急剧上升,形成绿色植物特有的红边与近红外高台阶,直到 1280 nm 后才因叶片内水份的吸收而下降。总体而言,小麦的光谱曲线表现出典型绿色植物的光谱特征。

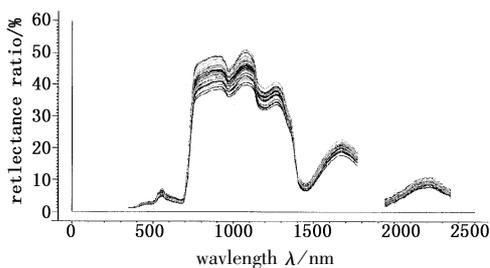


图 1 小麦冠层的可见/近红外反射光谱

Fig. 1 the visible/near-infrared reflective spectra of wheat canopy

3.2 光谱预处理方法的评价与优化选择

本文所采用的各种光谱预处理方法如表 1 所示。为获得最佳的光谱预处理方法,利用校正集的 142 份样本,以 PLS 算法对经各种方法处理后的光谱数据分别建模,并以校正决定系数(calibration r-square, 记为 $C - R^2$)、校正均方根误差(root mean square error of calibration, RMSEC)、检验决定系数(validation r-square, 记为 $V - R^2$)以及检验均方根误差(root mean square error of validation, RMSEV)为指标进行比较分析,从而筛选出最佳的光谱预处理方法。

表 1 不同预处理模型的 R^2 与 RMSE
Tab. 1 R^2 and RMSE of model with different pretreatment methods

预处理方法 Pretreating methods	评价指标 (Evaluating indexes)			
	校正集 $R^2(C - R^2)$	RMSEC	检验集 $R^2(V - R^2)$	RMSEV
不处理	0.762	0.605	0.749	0.623
均值平滑	0.780	0.571	0.764	0.596
S. Golay 平滑	0.782	0.570	0.766	0.597
小波除噪	0.789	0.567	0.771	0.591
一阶导数	0.792	0.554	0.760	0.606
小波除噪 + 一阶导数	0.849	0.451	0.817	0.498
小波除噪 + 二阶导数	0.885	0.393	0.799	0.524
小波除噪 + 标准归一化	0.834	0.473	0.815	0.502
小波除噪 + 多元散射校正	0.829	0.479	0.809	0.510
小波除噪 + 标准化	0.823	0.488	0.804	0.517
小波除噪 + 基线校正	0.830	0.477	0.806	0.514

经比较发现,原始光谱(光谱不经任何预处理)所建立的模型效果最差, $C - R^2$ 与 $V - R^2$ 仅为 0.762 与 0.749, RMSEC 与 RMSEV 高达 0.605 与 0.623。对光谱进行除噪后,模型精度有一定程度的改善。在移动平均平滑、S. Golay 平滑以及小波除噪三种方法中,以小波除噪的效果最好,其模型 $C - R^2$ 与 $V - R^2$ 在分别达 0.789 与 0.771, RMSEC 与 RMSEV 分别为 0.567 与 0.591, 优于 S. Golay 平滑等常用的除噪方法。而不经除噪步骤,直接将原始光谱进行一阶导数处理后建立模型,其 $C - R^2$ 与 $V - R^2$ 仅为 0.792 与 0.760, RMSEC 与 RMSEV 分别达 0.554 与 0.606, 相对原始光谱模型虽有所改善,但结果不理想,检验模型的精度尤低,尚达不到经小波或平滑除噪后的精度。表明一阶导数虽有去除光谱背景信息的功能,但同时加强了光谱中的噪声信息,单独使用无法取得理想的结果。

而光谱经小波除噪后,再经导数、标准归一化等处理方法处理均可有效消除背景值的影响,明显提高模型的精度。相比其他背景去除方法,导数处理尤其具优势。光谱经一阶导数与二阶导数处理后,所

建立的模型 $C-R^2$ 分别达 0.885 与 0.849, RMSEC 仅为 0.393 与 0.451, 对校正集样本均有良好的拟合结果。但对模型进行检验时, 发现经二阶导数处理的模型 $V-R^2$ 下降到 0.799, RMSEV 上升至 0.524, 表明光谱经二阶导处理所建立的模型预测能力较差, 不稳健, 须剔除; 而经一阶导处理的模型 $V-R^2$ 达 0.817, RMSEV 在 0.498 以下, 既具有精确的拟合结果, 也具有好的预测能力, 是一种优选方法。说明在光谱信息的预处理中, 小波除噪结合一阶导数的方法可在去除土壤等背景信息的同时有效抑制噪声, 为最佳预处理方法。

3.3 建模方法的优化选择

为择取较好的建模方法, 以偏最小二乘支撑向量机回归算法 (LS-SVR) 建立模型, 与上文中最优 PLS 模型 (即经小波除噪结合一阶导数预处理后所建立的模型) 进行比较分析后实现模型的选优。

与 PLS 算法不一样, 支撑向量机算法不包括降维的步骤。而由于小麦反射光谱维数较高, 如果将其全部引入模型, 不但会增加建模难度, 还会引入噪声而降低预测精度, 为避免这一问题, 在建模前需要采用主成分分析法 (PCA) 对光谱数据进行降维处理。数据经 PCA 降维后, 前 4 个主成分的累积贡献

率为 84.867%, 已包含了原始光谱绝大部分的特征信息。将其作为输入变量, 以校正集的 142 份样本在 LIBSVM 中建立最小二乘支撑向量机回归模型 (LS-SVR)。LS-SVR 模型选用径向基核函数 (RBF), 对结果影响较大的两个参数 (惩罚系数 C 与 RBF 核函数系数 γ) 通过交叉验证确定, 其余参数采用默认值。交叉验证采用格网搜索法 (grid-search) 进行。运行程序后发现, 当 C 为 50.028 时, γ 为 8.016 时模型可取得最佳结果。

建模完成后, 为比较 LS-SVR 模型与 PLS 模型的精度, 将两种模型对校正集的拟合结果与标准值进行回归分析, 并以模型复相关系数 (R^2)、误差均方根 (RMSE)、斜率 (Slope) 与偏移 (Offset) 四个指标对模型进行评价。两类模型对校正集的拟合结果如图 2 所示。两类模型均具有较高的精度, 但 LS-SVR 模型的样本点在回归线附近更为集中, 其校正集决定系数 ($R^2 = 0.905$) 比 PLS 模型 ($R^2 = 0.849$) 高, 校正集误差均方根 (RMSEC = 0.335) 比 PLS 模型 (RMSEC = 0.451) 低, 而斜率 (Slope = 0.892) 则比 PLS 模型 (Slope = 0.849) 更接近于 1, 偏移 (Offset = 0.246) 也比 PLS 模型 (Offset = 0.344) 更小。说明各指标均表明 LS-SVR 模型对校正集样本具有更好的拟合结果。

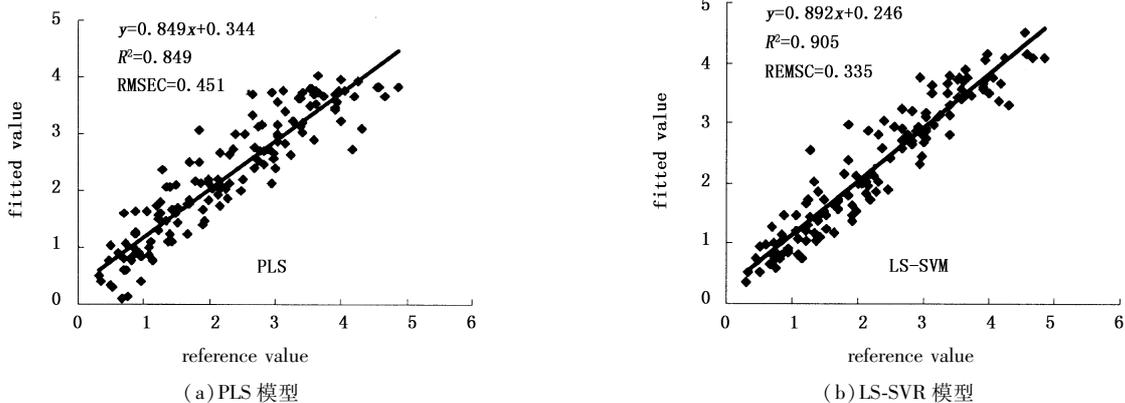
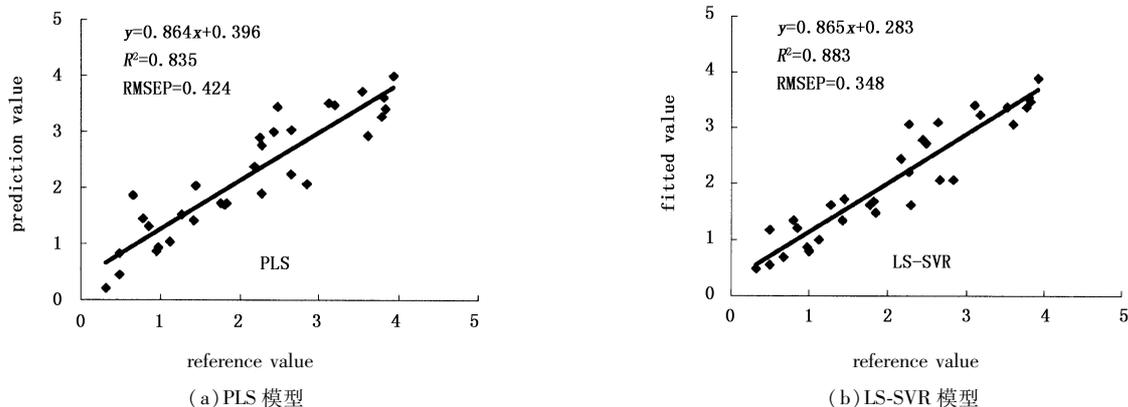


图2 PLS模型与LS-SVR模型对校正集样本的拟合结果($n=142$)

Fig.2 standard value versus fitted value of PLS model and LS-SVR model in calibration set ($n=142$)

为进一步比较两类模型对未知样本的适应性, 利用所建立的 PLS 模型与 LS-SVR 模型分别对预测集样本进行预测。将两类模型预测的结果分别与标准值进行回归分析以评价其预测精度, 发现 LS-SVR 模型预测集决定系数的取值 ($R^2 = 0.883$) 高于 PLS 模型 ($R^2 = 0.835$), 预测误差均方根的取值 (RMSEP = 0.348) 则比 PLS 模型 (RMSEP = 0.424) 低, 而斜

率 (Slope = 0.865) 比 PLS 模型 (Slope = 0.864) 更接近于 1, 偏移量 (Offset = 0.283) 也比 PLS 模型 (Offset = 3.396) 更小, 即 LS-SVR 模型具有更高的预测精度 (图 3)。结合上文的分析可知, 与 PLS 模型相比, LS-SVR 模型不但对检验集样本具有更佳的拟合结果, 也具有更佳的预测结果, 即对未知样本的适应性更好, 是建模的优选方法。

图3 PLS模型与LS-SVR模型对预测集样本的预测结果($n=32$)Fig.3 standard value versus predicted value of PLS model and LS-SVR model in prediction set($n=32$)

4 结论与讨论

叶面积指数是表征植被冠层结构最基本的参量之一,与植物蒸腾作用、光合作用、水分利用以及生产力构成等均具有密切关系,是进行农作物监测与管理的重要指标^[1-2]。利用田间可见/近红外光谱进行作物叶面积指数的测定,可实现这一指标的快速、无损获取,为形成合理的农艺决策提供科学依据。但在建模分析中,如何去除土壤等背景信息的影响是一尚需解决的问题。本文的研究表明,以小波除噪结合一阶导数去除小麦冠层反射光谱中的土壤背景信息以提高模型的精度是可行的,且利用LS-SVR算法所建立的模型具有比传统PLS算法更高的精度,是建模的优选方法,从而在光谱预处理与建模方法两方面对小麦叶面积指数的可见/近红外光谱测定模型进行了改进。

由于光谱仪的随机噪声以及土壤等背景信息对其他作物类型,或同一作物的其他理化参量具有类似的影响,因此,这一方法也可为其他作物类型以及其他理化参量测定模型的改进提供参考。

参考文献:

- [1] Moulin S, Bondeau A, Delecalle R. Combining agricultural crop models and satellite observations: from field to region scales[J]. *International Journal of Remote Sensing*, 1998, 1(6): 1021-1036.
- [2] Inoue Y. Synergy of remote sensing and modeling for estimating eco-physiological processes in plant production [J]. *Plant Production Science*, 2003, 6(1): 3-16.
- [3] Broge N H, Mortensen J V. Deriving green crop area index and canopy chlorophyll density of winter wheat from spectral reflectance data[J]. *Remote Sensing of Environment*, 2002, 81: 45-57.
- [4] Yang Minhua, Liu Liangyun, Liu Tuanjie, et al. Experiment study on high spectrum data inversion of wheat physical and chemical parameters[J]. *Acta Geodaetica et Cartographica Sinica*, 2002, 31(4): 316-321. (in Chinese) 杨敏华, 刘良云, 刘团结, 等. 小麦冠层理化参量的高光谱遥感反演试验研究[J]. *测绘学报*, 2002, 31(4): 316-321.
- [5] Wang Jihua, Huang Wenjiang, Zhao Chunjiang, et al. Research on foliage bio-chemical constituents and bud's quality with spectrum reflectivity [J]. *Journal of remote sensing*, 2003, 7(4): 277-284. (in Chinese) 王纪华, 黄文江, 赵春江, 等. 利用光谱反射率估算叶片生化组分和籽粒品质指标研究[J]. *遥感学报*, 2003, 7(4): 277-284.
- [6] Haboudanea D, Miller J R, Pattey E, et al. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture[J]. *Remote Sensing of Environment*, 2004, 90: 337-352.
- [7] Roshanak Darvishzadeh R, Skidmore A, Atzberger C, et al. Estimation of vegetation LAI from hyperspectral reflectance data: Effects of soil type and plant architecture[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2008, 10: 358-373.
- [8] Müller K, Böttcher U, Meyer-Schatz F, et al. Analysis of vegetation indices derived from hyperspectral reflection measurements for estimating crop canopy parameters of oilseed rape (*Brassica napus* L.) [J]. *Biosystems engineering*, 2008, 101: 172-182.
- [9] Demetriades-Shah T H, Steven M D, Clark J A. High resolution derivative spectra in remote sensing [J]. *Remote Sensing of Environment*, 1990, 33: 55-64.
- [10] Zhang Liangpei, Zheng Lanfen, Tong Qingxi. Estimation of biological variants with hyperspectra [J]. *Journal of re-*

- mote sensing, 1997, 1(2): 111 - 114. (in Chinese)
张良培, 郑兰芬, 童庆禧. 利用高光谱对生物变量进行估计[J]. 遥感学报, 1997, 1(2): 111 - 114.
- [11] Yang Minhua. Hyperspectral remote sensing information collection related to precision agriculture [D]. Beijing: Chinese agriculture university, 2002, 86 - 90. (in Chinese)
杨敏华. 面向精准农业的高光谱遥感作物信息获取[D]. 北京: 中国农业大学, 2002: 86 - 90.
- [12] Liu Mingcai. Wavelet analysis and application [M]. Beijing: Qinghua university press, 2005: 125. (in Chinese)
刘明才. 小波分析及其应用[M]. 北京: 清华大学出版社, 2005: 125.
- [13] To A C, Moore J R, Glaser S D. Wavelet denoising techniques with applications to experimental geophysical data [J]. Signal Processing, 2009, 89(2): 144 - 160.
- [14] Huang Fei-jiang, Zhu Shou-ye. Infrared image denoising based on wavelet transform and improved SVD [J]. Laser & Infrared, 2009, 39(3): 335 - 338. (in Chinese)
黄飞江, 朱守业. 基于小波变换和改进 SVD 的红外图像去噪[J]. 激光与红外, 2009, 39(3): 335 - 338.
- [15] Ji Hu, Liu Qin-lei, Zheng Yong-huang, et al. De-noising of infrared images based on stationary wavelet transform and Bayes estimation [J]. Laser & Infrared, 2009, 39(6): 677 - 680. (in Chinese)
季虎, 刘钦磊, 郑永煌, 等. 基于离散平稳小波变换和 Bayes 估计的红外图像去噪方法[J]. 激光与红外, 2009, 39(6): 677 - 680.
- [16] Wu Di, He Yong, Feng Shuijuan, et al. Application of infrared spectra technique based on LS-support vector machines to the non-destructive measurement of fat content in milk powder [J]. Journal of Infrared and Millimeter Waves, 2008, 27(3): 180 - 184. (in Chinese)
吴迪, 何勇, 冯水娟, 等. 基于 LS-SVM 的红外光谱技术在奶粉脂肪含量无损检测中的应用[J]. 红外与毫米波学报, 2008, 27(3): 180 - 184.
- [17] Sun D Y, Li Y M, Wang Q. A unified model for remotely estimating chlorophyll a in lake taihu, china, Based on SVM and in situ hyperspectral data [J]. IEEE Transaction on Geoscience and Remote sensing, 2009, 47(8): 2957 - 2965.
- [18] Vapnik V N. Statistical Learning Theory. New York: Wiley, 1998.
- [19] Jensen J R. Introductory digital image processing a remote sensing perspective [M]. Cheng Xiaoling, Transl. Beijing: China Machine Press, 2007, 288 - 296. (in Chinese)
延森. 遥感数据图像处理导论[M]. 陈晓玲, 译. 北京: 机械工业出版社, 2007: 288 - 296.