

文章编号: 1672-8785(2016)09-0030-07

SEM 模型与正交试验在多源光谱 分析应用中的对比研究

— 以吉林西部盐碱土为例

陈星任 韩 阳* 肖 晗 王家琪

(东北师范大学地理科学学院, 吉林长春 130024)

摘 要: 在地物多角度偏振反射光谱的观测中, 除土壤自身的理化性质外, 探测时所采用的偏振角、探测天顶角等角度因子也会对地物光谱产生明显的影响。通过引入结构方程模型 (Structural Equation Modeling, SEM), 对多因素影响下的吉林西部盐碱土进行了偏振、多角度和高光谱等多源光谱分析, 发现几何角度因素的影响强度大于土壤自身性质 (含盐量) 的影响强度。设计了三因素两水平的正交试验, 并进行了分析, 结果同样表明角度因素的影响更显著, 即在单因子影响的分析中, 偏振角、探测天顶角对所测光谱的影响要大于土壤自身含盐量的影响, 而且不同因素的交互作用也会对光谱特征造成一定的影响。对 SEM 模型与正交试验这两种分析方法进行了对比, 得出了这两种方法在光谱分析中的差异与优势。

关键词: SEM; 正交实验; 偏振; 多角度; 高光谱

中图分类号: TP79 **文献标志码:** A **DOI:** 10.3969/j.issn.1672-8785.2016.09.006

Comparative Study of SEM Model and Orthogonal Test in Application of Multi-source Spectral Analysis —A Case Study of Saline Alkali Soil in Western Jilin

CHEN Xing-ren, HAN Yang, XIAO Han, WANG Jia-qi

(College of Geography Science, Northeast Normal University,
Changchun 130024, China)

Abstract: In the observation of multi-angle polarized reflective spectra of objects, in addition to the physical and chemical properties of soil self, the angle factors such as polarization angle, detection zenith angle and so on will have significant impact on the spectra of objects too. Through the introduction of structural equation model (SEM), the multi-source spectral analysis including polarization, multi-angle and hyperspectral technologies is made for the saline alkali soil in western Jilin under the influence of multiple factors. It is found that the impact strength of the geometry factor is greater than that of the nature (salinity) of soil self. The orthogonal tests of three factors and two levels are designed and analyzed. The results also show that the impact of the angle factor is more significant. That is, in the

收稿日期: 2016-05-11

基金项目: 国家自然科学基金项目 (41301364), 吉林省科技发展计划项目 (20150520067JH)

作者简介: 陈星任 (1994-), 男, 河南新密人, 本科, 专业为地理信息科学。 E-mail: 215241857@qq.com

* 通讯作者 E-mail: hany025@nenu.edu.cn

analysis of single factor, the impact of polarization angle and detection zenith angle on the measured spectra is greater than that of the salinity of soil self. Furthermore, the interaction of different factors will have a certain impact on the spectral characteristics too. The SEM is compared with the orthogonal test. The difference and advantages of these two methods in spectral analysis are obtained.

Key words: SEM; orthogonal experiment; polarization; multi-angle; hyperspectrum

0 引言

近年来, 吉林西部盐渍化土壤的面积不断扩大, 土壤的质量退化加剧, 区域土地资源遭到了严重的破坏。随着用遥感技术对盐渍化土壤进行的研究日益深入, 在利用多波段、多时相、高光谱遥感数据提高遥感对地物识别能力的分析过程中, 角度信息在遥感图像识别和分类方面的重要性愈发引起人们的注意^[1]。因此, 对多角度反射光谱的研究逐渐成为一个热点。多角度的遥感改变了传统遥感只能从单一方向得到目标投影而缺乏足够的信息以推断地物信息的情况, 使人们在获得平面信息的基础上增加了三维信息。20 世纪 70 年代末, 美国科学家 Suits、Goel、Strahler 以及李小文院士等人对二向性反射做了双向反射分布函数 (Bidirectional Reflectance Distribution Function, BRDF) 的研究, 以期实现对目标地物空间结构的反演^[2-4]。偏振反射是二向反射的孪生姐妹, 人们在对二向反射进行研究的同时也势必要对它的偏振反射进行研究^[5-6]。

地物的反射光谱往往受多种因素的影响, 偏振角、探测天顶角、入射天顶角、方位角等角度因素以及土壤自身的一些性质如 pH、含水量、含盐量、有机质等都会对土壤的光谱产生一定的影响。因此, 在探究这些影响的过程中, 统计学方法起很重要的作用。东北师范大学的韩阳、赵云升曾对影响植物叶片偏振高光谱的因素做了交互作用的研究, 发现偏振角、入射天顶角、叶绿素以及三个因素之间的交互作用都可以对植物叶片的偏振反射产生影响。孤立地考虑单因素的影响是不全面的^[7]。除了设计正交试验外, SEM 作为一种多变量的统计学方法, 同样

适于多因素的的研究。虽然这种方法只是近几年才被引入到生态学以及环境科学研究中, 但是已经呈现出很好的发展势头, 为许多复杂问题的解决提供了独特的视角^[8]。但在地理学以及光谱分析中, 这种方法却仍未被学者引用。

以吉林省西部典型的盐渍化土壤为例, 我们进行了多角度偏振高光谱的多源光谱探测, 并进行了土壤的盐碱度分析。在多种因素作用下, 利用 SEM 分析了各个因素对多源光谱的影响程度, 并设计了科学的正交试验, 进行了交互作用分析, 对两种分析方法所得出的结果进行了对比, 得出了两种方法在光谱分析中的差异与优势。

1 实验数据采集与分析方法

1.1 数据采集

研究区选择在吉林省白城市, 白城市位于吉林省西北部, 嫩江平原的西部, 科尔沁草原东部, 是吉林西部盐碱地分布最广的区域之一, 土壤的盐渍化程度很高。土壤样本的采集时间为 2015 年 8 月, 研究区内共有三块样本汇集土地, 在每块土地上设置五个取样点, 并利用五点取样的方式取土。构建土壤剖面, 采集土壤表层 0~20 cm 的土壤, 并进行 GPS 定位, 记录采样地点。将样品分为两组, 一组用于理化性质的测定, 一组用于在偏振光谱实验室进行光谱信息测量。

在东北师范大学“土壤地理实验室”内预处理野外采集的盐渍化土壤样本。具体步骤是, 将植物残体等杂质剔除、研磨、过 1 mm 孔径筛、烘箱烘干, 用电导仪测定土壤样品的电导率和土壤的 pH 值, 建立可溶性盐与电导率的线性关系, 计算土壤的含盐量。在原土的基础上添加碳酸钠配置样品梯度。

用美国 ASD 公司生产的野外便携式高光谱仪采集光谱数据, 光谱仪的波段范围为 350~2500 nm, 分辨率在 350~1000 nm 波段为 3 nm, 在 1000~2500 nm 波段为 10nm。数据获取的平台为东北师范大学空间偏振高光谱实验室中的地物光谱多角度测试平台(自行设计、制造), 可自行调节各种角度的位置(入射角、探测天顶角、偏振角等), 在探测光谱数据时可以获得多角度偏振高光谱信息。

1.2 SEM 简介和模型设计

SEM 是基于变量的协方差矩阵来分析变量之间关系的一种统计方法, 所以也称为协方差结构分析, 探索系统内变量间的“因果”关系强弱是其核心内容。相比于传统的统计分析方法, 其优势在于能够同时处理多个因变量。

1.2.1 SEM 变量类型

SEM 包含以下变量类型:

(1) 显变量: 显变量又称观察变量, 指在试验过程中及观测过程中能够直接测量出来的变量数据类型, 比如土壤的含盐量、pH 值和有机质含量等。潜变量则可通过观察变量来间接测得。

(2) 潜变量: 分为内生潜变量和外源潜变量。潜变量指无法直接测量的一种变量类型, 通常用来表示一种较抽象的影响因素, 但可以通过显变量间接测得, 如在土壤评价中对于“土壤理化性质”的评价, 便可以通过对土壤的含水量、土壤的含盐量和土壤的 pH 值等理化指标来进行测量。

(3) 残差项: 又称为误差变量, SEM 容许自变量和因变量含测量误差, 残差项指在分析过程中预测变量无法解释的那部分效应。

1.2.2 SEM 的统计学基础

SEM 主要包括测量模型和结构模型, 其中测量模型反映潜变量与显变量之间的关系, 结构模型反映潜变量之间的结构关系。SEM 通过三个矩阵方程来代表:

$$\eta = B\eta + \Gamma\xi + \zeta \quad (1)$$

$$Y = \Lambda_y\eta + \varepsilon \quad (2)$$

$$X = \Lambda_x\xi + \sigma \quad (3)$$

式(1)表示结构模型, 其中 η 和 ξ 分别为内生潜变量和外源潜变量, η 通过 B 和 Γ 系数矩阵和误差向量 ζ 把内生潜变量和外源潜变量联系起来。式(2)和式(3)表示测量模型, X 和 Y 分别为外源潜变量与内生潜变量的可测变量, Λ_y 、 Λ_x 为外源潜变量与内生潜变量同其各自可测变量的关联系数矩阵, 可测变量可以通过测量模型来反映潜变量^[9]。 ε 和 σ 为相应的残差项。

1.2.3 初始模型的设计

方向反射研究中涉及到的一些变量如角度因素、土壤自身理化性质等不能被准确地直接测量, 在 SEM 分析研究中可以把这些变量作为外源潜变量。用含盐量、含水量、pH 和有机质含量等作为土壤理化性质的外显指标; 用偏振角、探测天顶角等角度信息作为角度影响因子的外显指标。这两种潜变量都会对土壤的光谱特征产生影响。将研究区土壤偏振反射高光谱的特征波段作为内生潜变量, 分别选取 443 nm、670 nm 和 865 nm 三个波段的反射率作为反应光谱特征的指标。选取这三个波段是因为 POLDER 卫星在监测土壤时主要通过这三个波段进行监控, 选取这三个波段更有说服力。变量与变量之间的连线称为路径, 其能代表量量之间的相互影响关系。进行路径分析后可得出不同变量影响强度的大小。图 1 是在 lisrel 软件中建立的一个简单的结构方程模型。

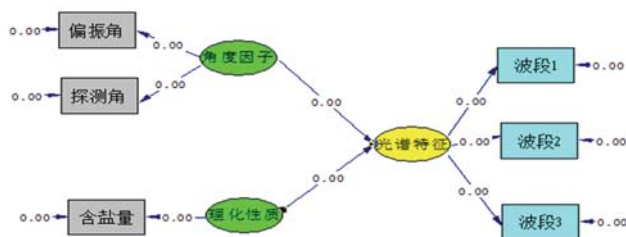


图 1 初始模型

1.3 正交实验设计

本文设置了三因素二水平的正交实验。三个因素分别为含盐量 A 、探测天顶角 B 和偏振

角 C 。东北师范大学赵云升曾探究过方位角对碱土偏振高光谱的影响, 发现在不同的方位角情况下, 其光谱特征没有特别大的差异, 具有朗伯体的性质^[10]。故本次实验未把方位角作为影响因素之一。因素 A 有两个水平, 分别为含盐量 $A_1=0.09\%$ 和 $A_2=0.11\%$; 因素 B 有两个水平, 分别为 $B_1=20^\circ$ 和 $B_2=30^\circ$; 因素 C 有两个水平, 分别为 $C_1=0^\circ$ 以及 $C_2=90^\circ$ 。实验设计方案见表 1。

表 1 因素水平表

水平	含盐量 A	探测天顶角 B	偏振角 C
1	0.09%	20°	0°
2	0.11%	30°	90°

要进行正交试验, 需进行科学的正交试验设计。由于实验设计为三因素二水平, 故通过正交表 $L_8(2^7)$ 得到表头设计方案, 见表 2。

表 2 正交试验表头设计

列号	1	2	3	4	5	6	7
因素	A	B	A*B	C	A*C	B*C	

2 结果与讨论

2.1 原始偏振高光谱的特征分析

2.1.1 不同偏振角状态下的吉林西部盐碱土偏振高光谱分析

土壤在不同偏振状态下显示出不同的光谱特征。为了更直观地展现其影响, 选取 0° 偏振及 90° 偏振状态下的盐碱土偏振反射高光谱。图 2 表明, 在探测天顶角和土壤含盐量相同的情况下, 0° 偏振时的光谱反射率要高于 90° 偏振时的光谱反射率, 且影响较明显。不同状态下的盐渍化土壤光谱曲线与传统的盐渍化土壤高光谱反射率图有相似的特征, 整体上随着波长的增长而平缓上升。在 1400 nm 和 1900 nm 附近有两个明显的吸收谷, 其形成原因是土壤中水的吸收作用。在 2100 nm 附近有一个反射峰。

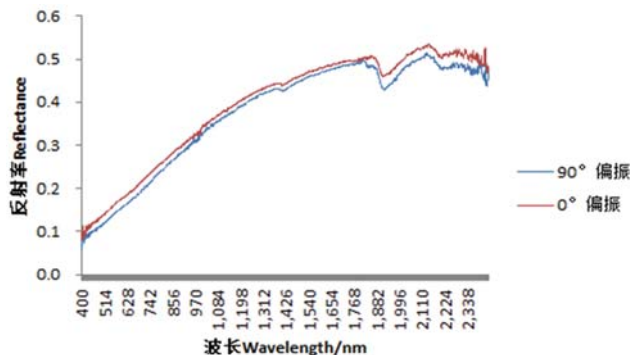


图 2 不同偏振角下的盐渍化土壤光谱曲线

2.1.2 不同探测天顶角状态下的吉林西部盐碱土偏振高光谱分析

由图 3 可以看出, 30° 探测天顶角状态下的偏振反射率要高于 20° 时的偏振反射率, 反映探测天顶角对土壤光谱反射率存在一定的正影响, 但相比于偏振角, 其影响不那么明显。

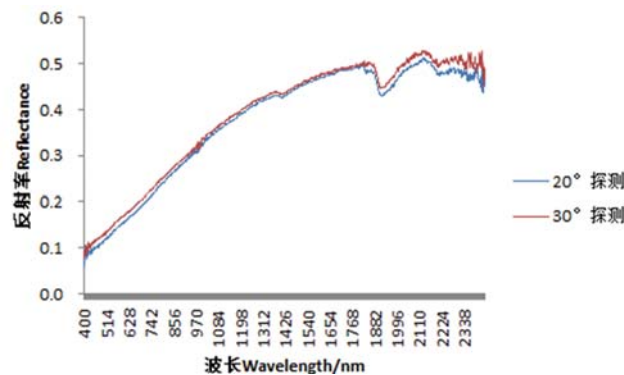


图 3 不同探测天顶角下的盐渍化土壤光谱曲线

2.1.3 不同含盐量下的盐渍化土壤偏振高光谱分析

选取含盐量 0.09% 和 0.11% 的盐渍化土壤光谱曲线进行分析, 这主要因为白城地区盐渍化土壤的含盐量在这个区间内浮动, 选取这两种含盐量具有指导意义。由图 4 可以看出, 含盐量对于土壤光谱的反射率具有正影响, 即反射率会随着含盐量的增加而升高, 但影响程度明显弱于偏振角和探测角所带来的影响。实际上已有学者指出, 盐碱土在可见光 / 近红外 (VIS/NIR) 波段的反射率会随着含盐量的增加而升高, 但只在含盐量较高的土壤结壳时有较好的相关性。因此, 我们在进行相关性研究时要考虑这些影响因素。

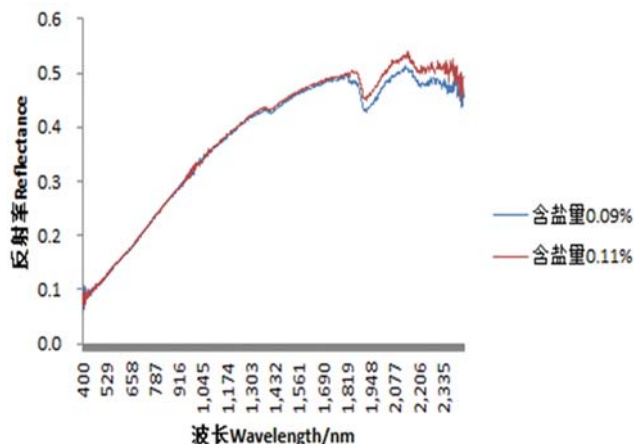


图 4 不同含盐量下盐渍化土壤的光谱曲线

2.2 SEM 路径运算的定量分析

对已设计好的初始模型进行路径分析,使路径系数完全标准化,得到各个路径系数,然后通过路径系数的大小得到各变量之间的影响强度。由图 5 可以看出,综合角度因子的影响强度为 -0.74 。负的路径系数代表变量之间的负影响,而理化性质或者说含盐量对所选取的特征波段的影响强度只有 0.14 。这说明角度因子的影响强度比理化性质的大。在进行单因素的影响强度分析时, A 对 B 的影响强度与 B 对 C 的影响强度的乘积为 A 对 C 的影响强度。因此可以得出,偏振角对特征波段的影响强度达到了 -0.62 ,为较大的负影响;探测角的影响则达到了 0.23 ,为一定程度上的正影响,而含盐量对特征波段的影响强度为 0.13 ,为不太明显的正影响,这与原始光谱给我们的直观感受一致。

通过各变量间路径系数的大小可以得出,影响吉林西部盐碱地光谱特征的最主要的因素是角度变化,其特征波段的反射率变化受综合角度因素的负影响(路径系数为 -0.74),土壤的理化性质对土壤偏振反射光谱的特征波段有一定的影响(路径系数为 0.14),但不如角度因素的影响显著。而在单因素的影响分析中,偏振角的影响大于探测天顶角的影响,含盐量的影响最小。

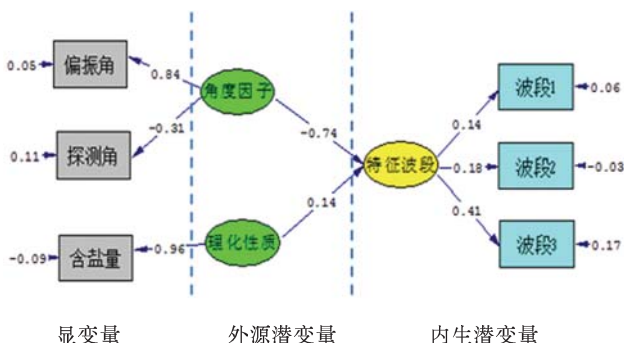


图 5 路径分析结果图注: (1) 路径系数为 0.1 代表自变量变化一个标准差将影响因变量 0.1 个标准差 (2) 对于单独因素的影响强度来说, A 对 B 的影响强度与 B 对 C 的影响强度的乘积则为 A 对 C 的影响强度。

2.3 交互作用下的定量分析

根据正交试验的设计方案,测量了各因素不同水平组合下的多角度高光谱偏振反射率,见表 3。其中, K_1 、 K_2 为各因素不同水平的均值, R 为极差,“1”和“2”代表各因素的水平,虽然 $A \times B$ 、 $A \times C$ 和 $B \times C$ 三列中的“1”和“2”没有实际意义,但其计算的极差 R 有统计意义^[11]。极差越大,则该因素对偏振反射的影响就越大。根据表 3 所展现出的极差大小,可以得出各因素与其交互作用的主次关系为

$$\text{主 } \overline{CBAA} \times \overline{BB} \times \overline{CA} \times \overline{C} \text{ 次}$$

对这些因素及其交互作用因素进行方差分析,可以反映各因素影响的显著程度。参照统计学的方法,检验一个因素对实验结果有无显著影响,先从 F 分布表中查出临界值 F_{α} 值,然后比较 F 值与临界值的大小。若 F 值大于临界值,则该因素对实验结果有显著影响。一般来说, F 值与临界值之间的差距越大,说明该因素或交互作用对实验结果的影响越显著^[12]。分别选取了 443 nm 、 670 nm 和 865 nm 的反射率进行交互作用分析,得到的三个波段显著性分析结果一致。表 3 是波长 670 nm 处的正交试验分析结果。通过表 3 的方差结果可以看出,偏振角对土壤光谱特征的影响最大,显著性为特别显著;其次是探测角和含盐量的影响,较为明显。同时可以发现,虽然不同因素的交互作用对吉林西部盐渍化土壤有一定的影响,但程度不如单因素的影响显著。

表 3 不同偏振角、探测天顶角和含盐量下的正交试验计算结果

试验号	A	B	A × B	C	A × C	B × C	空列
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2
K1	75.6597	75.1600	76.3199	73.3971	76.0401	75.9652	76.1214
K2	76.6145	77.1142	75.9543	78.8771	76.2341	76.3090	76.1529
R	0.9548	1.9542	0.3656	5.4800	-0.1940	0.3438	0.0315
F	48.45	328.43	19.58	2354.70	0.1610	17.94	
显著性	一般显著	显著	一般显著	特别显著	不显著	显著	

2.4 正交试验与 SEM 分析的对比

在分析单因子对吉林西部盐碱土偏振高光谱特征的影响时, 采用两种分析方法得到的结果相同, 即偏振角的影响最大, 探测角的影响次之, 土壤含盐量的影响最小。这一点可以一定程度上证明 SEM 在光谱分析中的应用是可行的, 同时通过它们的对比也可以得出两种分析方法各自的侧重点。

SEM 的优势主要是, SEM 可以快速有效地分析多个变量间的因果关系, 将不可测量的变量变为可测, 并且可以给出明确的影响强度大小, 而传统的针对多变量研究的统计方法没有足够的灵活性; 同时, SEM 也可以对存在的误差进行有效处理; 相比于其他的统计分析方法, SEM 更简单, 参数更少, 更加直观灵活; 在特别复杂的系统中, 例如存在几十个变量时, SEM 分析的优势特别明显。

正交试验则能进行交互作用的分析, 交互作用是要素之间相互影响、互相关联的关系。影响因素的交互作用在自然界中是普遍存在的。在医药学、农学和化工等领域, 交互因素都是被重点研究的对象。交互作用分析能使我们对影响事物的因素有更清晰的认识。而且正交试验的

结果准确, 能给出各个因子影响的显著程度, 这是其与 SEM 分析相似的地方。

目前, SEM 对交互作用的分析正是统计学研究的前沿。相信随着技术的成熟, 该方法在遥感应用中会起到更大的作用。

3 结论与展望

通过对吉林西部盐碱土的偏振高光谱的测量分析, 发现: (1) 偏振角、探测天顶角及含盐量都会对吉林西部盐碱土偏振高光谱的特征造成影响, 影响主要体现在光谱反射率上。0° 偏振时的反射率比 90° 偏振时的大, 30° 探测天顶角下的反射率要大于 20° 时的反射率。随着含盐量的上升, 土壤反射率也呈上升趋势, 但不是特别明显。(2) 利用 SEM 对影响吉林西部盐碱土偏振反射特征的因素进行了分析。结果表明, 综合的角度影响强度要大于土壤自身的一些理化性质的影响, 且偏振角的影响强度大于探测角的影响强度, 土壤含盐量的影响强度最小。(3) 上述三个因素以及它们的交互作用都会对盐渍化土壤的偏振反射产生一定的影响, 但交互作用的影响不如单因子的影响程度显著。在单因子的影响程度中, 偏振角的影响最大, 探测天顶角的影响次之, 土壤含盐量的影响最小。(4) 正

交试验与 SEM 分析在光谱分析中均能取得可信的结果,两者既存在差异又有各自的优势。SEM 在光谱分析中的应用具有一定的可行性。

由于实验条件的限制,没有对土壤的理化性质进行更多指标的测量,最后得出的结论不够全面,但仍然能反映一些问题。实际上,影响土壤反射率的因素非常多,如土壤类型、有机质含量、水波段分含量、颗粒大小和表面粗糙程度等,所以在 VIS/NIR 波段反射率升高的原因也特别多,并不只有盐碱化这一种解释,而且盐碱化导致反射率升高的本质原因可能是水分含量减少、有机物减少和表面结壳等因素共同作用的结果。因此,在利用遥感对土壤盐碱化进行研究时一定要考虑多种因素,对于一个事物的描述,其描述指标越多则表述就越准确。在今后的实验研究中,若能综合多种因素的影响对盐碱土的光谱特征进行描述,会取得不错的成果。

参考文献

- [1] 赵虎,晏磊,赵云升. 橄榄岩的多角度偏振反射光谱研究[J]. *地质与勘探*, 2004, **40**(2):51-54.
- [2] Li X W, Strahler A H. Geometric Optical Modeling of a Coniferous Forest Canopy[J]. *IEEE Trans. Geosci. Remote Sensing*, 1985, **23**:207-211.
- [3] 刘强,柳钦火,肖青,等. 机载多角度遥感图像的几何校正方法研究[J]. *中国科学 D 辑*, 2002, **32**(4):299-306.
- [4] 李小文, Strahle A H, 朱启疆,等. 基于颗粒构成的粗糙表面二向性反射-相互掩蔽效应的几何光学模型[J]. *科学通报*, 1993, **38**(1):86-89.
- [5] 韩阳,赵云升,赵乃卓,等. 森林土壤多角度高光谱偏振反射影响初探[J]. *光谱学与光谱分析*, 2009, **29**(3):702-706.
- [6] 韩阳,赵云升,张莉莉,等. 丁香叶片叶绿素含量偏振高光谱数学模型反演研究[J]. *光谱学与光谱分析*, 2009, **29**(6):1595-1598.
- [7] 韩阳,李潜,赵云升,等. 三因素及其交互作用对植物多叶片多角度偏振高光谱特征的影响[J]. *红外与毫米波学报*, 2010, **29**(4):316-320.
- [8] 王酉石,储诚进. 结构方程模型及其在生态学中的应用[J]. *植被生态学报*, 2011, **35**(3):337-344.
- [9] 吴林海,侯博,高申荣. 基于结构方程模型的分散农户农药残留认知与主要影响因素分析[J]. *中国农村经济*, 2011(3):35-48.
- [10] 宋开山,赵云升,张柏. 土壤偏振反射特性的多角度测量与研究[J]. *土壤通报*, 2004, **35**(4):420-425.
- [11] 方开泰,马长兴. *均匀实验与正交试验*[M]. 北京:科学出版社, 2011, 35-77.
- [12] 李云雁,胡传荣. *实验设计与数据处理*[M]. 北京:化学工业出版社, 2009, 147.

(上接第 29 页)

参考文献

- [2] Wang W, Yang L J, Xie Y T, et al. Edge Detection of Infrared Image with CNN DGA Algorithm[J]. *Optik - International Journal for Light and Electron Optics*, 2014, **125**(1): 464-467.
- [3] Jing Y, An J B, Liu Z X. A Novel Edge Detection Algorithm Based on Global Minimization Active Contour Model for Oil Slick Infrared Aerial Image[J]. *Geoscience and Remote Sensing, IEEE Transactions on*, 2011, **49**(6):2005-2013
- [4] 王志刚,王伟,许晓鸣. 基于塔型结构的多分辨率模糊边缘检测[J]. *红外与毫米波学报*, 2002, **21**(4):305-309.
- [5] 王巍,安友伟,黄展,等. 基于 CNN 的红外图像边缘检测算法的 FPGA 实现[J]. *光子学报*, 2012, **41**(11):1354-1358.
- [6] Morrone M C, Owens R A. Feature Detection from Local Energy[J]. *Pattern Recognition Letters*, 1987, **6**(5):303-313.
- [7] Venkatesh S, Owens R A. An Energy Feature Detection Scheme[C]. *In The International Conference on Image Processing*, 1989: 553-557.
- [8] Kovesi P. Invariant Measures of Image Features from Phase Information[D]. Univ. of Western Australia, 1996.
- [9] Kovesi P. Image Features from Phase Congruency[J]. *A Journal of Computer Vision Research*, 1999, **1**(3):1-27.