

基于地质统计学影像纹理的石河子地区化控期棉花长势监测

廖楚江¹, 王长耀¹, 李红², 杨朋润²

(1. 中国科学院遥感应用研究所遥感科学国家重点实验室, 北京 100101; 2. 新疆石河子市科技开发中心, 石河子 832001)

摘要: 化学调控是棉花生长期的一个重要环节, 在较大区域尺度上, 化控常常整齐划一, 依固定的量而行, 忽视其普遍存在的空间差异, 结果导致某些棉田投入过量或投入不足, 在精细农业所倡导的基于特定场的农业投入这一框架下, 为辅助决策者制定有针对性的棉田化控方案, 该文提出了应用地质统计学影像纹理进行对化控棉花长势监测, 通过借助地质统计学方法在分析空间结构变化上的优势, 强化遥感影像上不同长势棉田间的差别, 提高样本选择的分离度和长势空间分类精度。实验结果表明, 运用变异函数纹理结合光谱波段的最大似然分类方法能够很好地界定棉花长势的不同类型, 总分类精度达 90.53%, 证明了应用地质统计学影像纹理方法在针对化控的棉花长势监测上的有效性。

关键字: 地质统计学; 影像纹理; 变异函数; 棉花化控; 中巴卫星

中图分类号: TP79; S127

文献标识码: A

文章编号: 1002-6819(2006)08-0135-05

廖楚江, 王长耀, 李红, 等. 应用地质统计学影像纹理的化控期棉花长势监测研究[J]. 农业工程学报, 2006, 22(8): 135-139.

Liao Chujiang, Wang Changyao, Li Hong, et al. Cotton growth monitoring during chemical control stage using geostatistical image texture: a case study of Shihezi[J]. Transactions of the CSAE, 2006, 22(8): 135-139. (in Chinese with English abstract)

0 引言

传统农业耕作方式采用大片土地的平均化投入, 忽略了大多数土地都存在的空间差异, 如利用统一的耕作、播种、灌溉、施肥、喷药等农艺措施, 很少考虑对农田的盲目投入及过量喷药施肥造成的生产成本增加和环境污染。

目前, 基于特定场的农业管理系统, 即精细农业提供了减少成本, 优化作物产量和环境保护的方法, 主要利用传感器及监测系统来收集各种数据, 发现由于作物生长环境和生长条件的空间差异性, 指导必要的和精确的农业投入。化学调控是棉花生长期的一个重要环节, 通过化控可以控制棉花个体长势, 使整体生长和谐, 抑制棉株疯水权的发生和生长, 甚至可以不用掐尖和打疯水权^[1,2], 可大大减少田间管理用工量。但是化控的量度常常较难把握, 主要视棉花长势而定, 即弱苗应少控或不控, 而苗旺则要多控或加大药量^[1,3]。在较大区域尺度上, 如新疆各兵团垦区, 棉花化控常常都是一致的, 结果使得某些地方化控程度不够, 而另外一些地方则过分化控, 前者造成了棉花的减产, 而后者则造成了不必要的成本浪费和环境污染^[4], 因此, 监测棉花在空间上的长势差异, 提出特定场的棉花化控需求, 是解决这一系列问题的关键所在。

收稿日期: 2005-08-15 修订日期: 2005-12-09

基金项目: 中国科学院知识创新工程重大项目(KZCX1-SW-01-02); 国家高技术研究发展计划(863 计划 2003AA131130)项目

作者简介: 廖楚江(1979-), 男, 湖北鄂州人, 博士生, 主要从事遥感数字图像处理、农情遥感监测方面的研究。北京 中国科学院遥感应用研究所遥感科学国家重点实验室, 100101。

Email: eren_sheep@yahoo.com.cn

1 研究区域和数据描述

北疆是中国棉花的主要产区, 棉田主要分布于天山北麓的洪积平原上, 该区气候干燥, 光热资源丰富, 是优质陆地棉的重要基地之一。本文选择位于石河子市东北部的农八师 148 团、149 团和 150 团作为研究区域(如图 1), 其经度为 85°53'58"~86°23'26", 纬度为 44°43'2"~45°10'54", 面积 1994.67 km²。

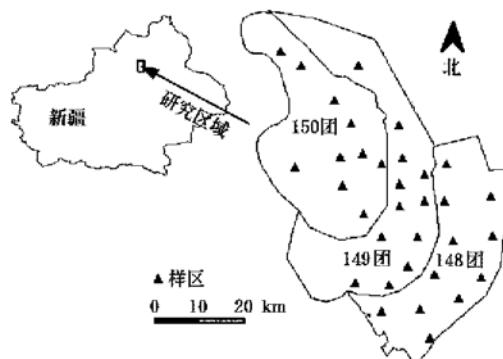


图 1 研究区域

Fig. 1 Study area

区域内棉花品种主要为新陆早 7 号和 8 号, 其生育进程为 4 月中下旬播种, 5 月上中旬出苗, 6 月中下旬现蕾, 7 月下旬至 8 月初开花, 9 月初开始吐絮。由于种植密度较高, 所以化控的次数一般为 3~4 次, 第一次化控时间一般选择在苗蕾期, 即 6 月中旬左右, 化控时间间隔控制为 10 d 到 15 d 左右。鉴于棉花长势空间差异在半个月之内变化不大, 因此利用遥感监测指导化控可以选择在第一次和第三次化控的时段, 从影像分辨率、成本和回归周期三方面考虑, 本文选用 2004 年 6 月 13 日

中巴卫星(CBERS)CCD 影像(分辨率 19.5 m, 回归周期 26 d)作为研究数据(4 个波段, 1981×2648 像素)。另外, 为配合方法和精度的验证, 于卫星过境时段的 2004 年 6 月 12 日~17 日在实地调查了 33 个样区, 每个样区平均选取 5 个 20 m×20 m 大小的样本点, 共调查了 165 个样本点的棉花平均株高、果枝台数、果枝节数和真叶数, 样区分布图如图 1 所示。

2 方法

棉花长势主要从株高、叶片数和果枝节数 3 个方面来衡量。在棉花种植密度基本相同的情况下, 由于土壤条件和生长条件的不同而长势不同, 长势好的棉田株高较高, 果枝节数较多, 叶片较密集, 实地调查时这些差异较易分辨出来, 但要从光谱影像上解译它们却十分困难, 分类精度无法达到要求, 因此在监测作物长势, 进行长势空间分类时, 人们一般寻求其它的途径。

传统的棉花长势监测主要以服务估产为目的, 方法上以遥感反演植被指数为基础, 如刘爱霞等(2003)基于 TM 和 CBERS 数据通过反演归一化植被指数(NDVI)监测棉花长势^[5], 吕建海等(2004)基于 MODIS 数据通过叶面积指数(LAI)监测棉花长势^[6]。此类方法的原理是长势好的棉田由于果枝节数更多, 叶片更密集, 其植被指数相对长势差的棉田更大, 据此可以得到棉花产量在时间轴和空间轴上的差异^[7,8]。但该类方法用于指导化控则存在不足, 因为 NDVI 和 LAI 随着棉花的生育期呈现明显的非单调等级变化, 使得化控主要针对的棉花旺长区域可能比其它区域 NDVI 和 LAI 还要低, 即它们不能识别旺长的棉田^[6,8]。

在单纯的光谱分类方法较难实现的情况下, 加入影像纹理是一个可取的办法^[9,10]。地质统计学方法一直是分析空间结构变化的最佳选择, 一些研究者尝试将其用于遥感分类, 结果发现, 加入了地质统计学影像纹理的分类方法在分辨地表同类地物的空间分布结构上有显著的优势, 如 Mcgwire 等(1993)利用地质统计学影像纹理探讨美国加州 Madera 郡橡树空间分布情形^[10], Miranda et al. 等(1998)应用地质统计学纹理基于 JERS-1 数据对巴西热带雨林地区浓密植被和稀疏植被的分类^[11], 周天颖等(2004)应用地质统计学纹理基于华卫 2 号数据对台湾南部水稻田进行分类^[12], 鉴于不同长势棉田在空间结构上的差异, 本文试图在原始光谱影像的基础上加入地质统计学的纹理影像来监测棉花长势的空间差异, 提高棉花长势分类的精度。

2.1 变异函数

从统计学的角度, 可以用两个与 DN 值相关联的概念表达纹理: 局部/全局的变化性和空间自相关, 前者是通过计算一个局部层次上的方差来达成, 即先计算一个移动窗口的均值, 后计算 DN 值与该均值的偏离程度^[13]; 而空间自相关是假定在影像中 DN 值不是完全随机分布的, 每种土地覆盖类别有其独特的空间变异性(空间依赖)结构, Lark 指出像素之间的变异函数的值依赖于它们的空间自相关, 该值可以作为每种土地覆盖

类别的纹理描述器^[14]。如下所述, 地质统计学可以将这两方面联合建模。

从地质统计学的角度, 可以把遥感影像的 DN 值看作为场变量, 既是随机分布的, 同时也存在一定的空间相关性, 在固有假设条件下, 这两方面可以通过变异函数连接起来, 方程表示为^[6]

$$\gamma(h) = \frac{1}{2}E[DN(x + h) - DN(x)]^2 \quad (1)$$

式中 E ——数学期望; $\gamma(h)$ ——相距为 h 的一对像素 DN 值平方差的数学期望的 $1/2$, 即半变异值, $\gamma(h)$ 是一个依赖位于 $x + h$ 和 x 处的两个像素之间步长矢量 h 的模和角度的矢量函数。变异函数表达的是空间依赖性, 不同的地表结构有其独特空间变异性(空间依赖)结构。

2.2 空间变异性的纹理计算

将地质统计学纹理信息引入到影像分类过程中, 必须计算一系列半变异值, 本文使用的变异函数有两种, 一种是方向变异函数, 另一种称为交叉变异函数, 前者用于计算单个波段的空间变异性, 而后者用于研究两个波段联合产生的空间变异性^[15]。

方向半变异函数可以参考式(1) 得到,

$$\gamma_k(h) = \frac{1}{2P(h)} \sum_{\alpha=1}^{P(h)} [DN_k(x_\alpha) - DN_k(x_\alpha + h)]^2 \quad (2)$$

式中 $P(h)$ ——相距步长 h 的像素对的数目; $DN(\cdot)$ ——像素 x_i 和 $x_i + h$ 处的 DN(data number) 值; k ——传感器的波段号。

交叉变异函数对两个波段之间的联合变异性(交叉相关)进行量化。如式(3) 所示, j, k 分别代表两个不同的波段。

$$\gamma_{jk}(h) = \frac{1}{2P(h)} \sum_{\alpha=1}^{P(h)} |DN_j(x_\alpha) - DN_j(x_\alpha + h)| \cdot |DN_k(x_\alpha) - DN_k(x_\alpha + h)| \quad (3)$$

2.3 变异函数参数

地质统计学影像纹理是通过计算移动窗口内的相邻像元的半变异值得到的, 计算方法如图 2 所示, 在计算的过程中需考虑下列几个因子, 计算窗口大小、像元取样距离(步长)、计算方向和波段组合。计算窗口(移动视窗)大小, 主要是依据在计算每个像元的纹理值时, 需要同时考虑到该像元周围多大的范围来决定, 一般是在反复实验的基础上得到的, 一个理想的窗口既不能太大, 否则将使得相邻的类型互相干扰, 但也不能太小, 否则将减小变异函数分类器的使用价值, 本文选择的考察窗口包括 3×3 、 5×5 和 7×7 窗口, 通过对比分类精度决定最终的窗口大小。像元取样距离(步长)依据窗口大小的改变而改变, 它代表了像元与像元之间的关系, 图 2 中窗口大小为 3, 步长可为 1 或者 2, 一般情况下, 计算结果取所有步长的平均值。考虑计算方向是因为地表景物的排列顺序与方向并不固定, 因此在计算纹理特征时, 需考虑计算东-西向(0°)、南-北向(90°)、东北-西南向(45°)、西北-东南向(135°) 4 个方向的纹理, 当

空间方向性不明显时, 可将4个方向的变异函数值加总取平均, 成为等向变异元。波段组合是针对交叉变异函数和虚拟变异函数的一个参数, 选择不同的波段组合, 将对地表景物的区分产生重要影响。

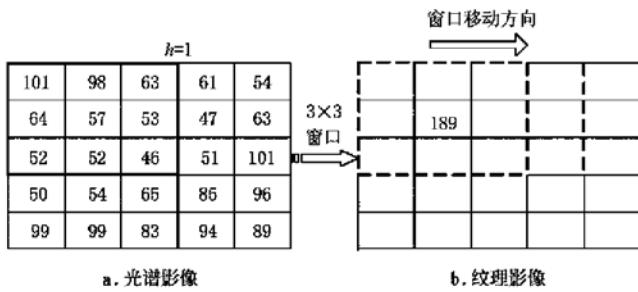


图2 地质统计学纹理计算示意图
Fig. 2 Calculation of geostatistical texture

2.4 不同长势棉田纹理特征

为了证明变异函数区分几种不同长势棉田的能力, 在影像上选择了通过实地调查得到的代表不同长势棉田的3块图斑, 大小为 17×17 像素, 分别为长势好, 较好和差, 利用方向变异函数公式进行计算, 得到的变异函数图如图4所示。在研究变异函数图时, 一般关注3个参数^[12], 变差距离、基台值和块金值(如图3), 从变异函数定义可以看出, 当一对样本距离增加时, 所对应的变异函数值一般也增加, 即当变异函数达到一个平台时的样本间距离为变差距离或相关尺度, 它表明当样本间距等于或大于此距离时, 样本就变得完全独立。在变差距离处变异函数所达到的平台叫做基台值。块金值是在极短的样本距离($h \approx 0$)之间变异函数从原点的跳升值(不连续性), 它是由于样本误差和短距离的变异性引起的。

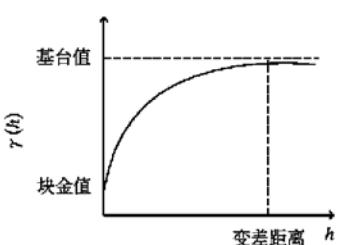


图3 变异函数图参数意义
Fig. 3 Parameter meaning of variogram

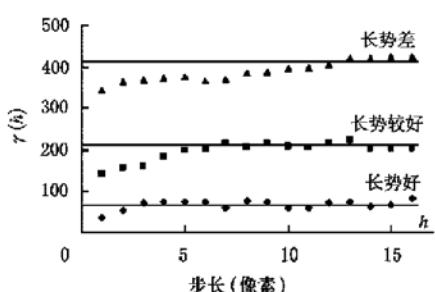


图4 不同类别变异函数图
Fig. 4 Variograms of different types

从图4中可以看出, 长势好的棉田变异函数曲线非常平缓, 当步长距离大于3时呈现小至于无的空间自相关, 而长势较好棉田变异函数在步长大于7时才变得平缓, 其基台值也比长势好的棉田更大, 说明长势较好的棉田比长势好的棉田有着更大的变异性, 长势差的棉田显示最大的变异性, 变异函数曲线在步长大于13个像素时才趋于平稳。

3 结果与分析

为了集中在棉花研究上, 首先对非棉花类型进行掩膜处理, 研究区域的主要地物类型有盐碱地、荒草地、居民地、水体、小麦和棉花, 以4, 3, 2波段假彩色合成进行解译, 其中盐碱地为白色, 居民地为蓝绿色, 水体为蓝黑色, 几种植被类型中小麦此时已经封垄, 显示为较棉花更深的红色, 荒草地显示为与居民地相似的蓝绿色, 因此较易区分, 采用最大似然分类法将上述类型分成两个类别, 棉花类和其它类, 分类后以棉花作为感兴趣区域进行掩膜处理, 其它类别赋以0值, 得到棉花的影像图。

3.1 研究变量选择

CBERS卫星CCD影像的4个波段前3个波段为可见光波段, 第4波段为近红外波段, 为了突出几种不同类型长势棉田的光谱和形态特征, 并减少用于纹理分析的变量, 需要首先对CBERS影像前4个波段进行了主成分分析, 得到两个用于研究的波段。

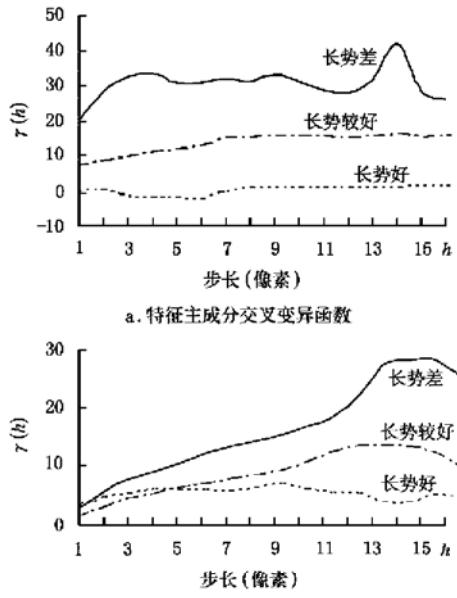


图5 两种主成分交叉变异元分析方法对比
Fig. 5 Comparison of two principal component analyses

主成分的选择直接牵涉到后续纹理计算的效果, 本文以交叉变异函数 17×17 窗口为例, 取长势好, 较好和差3种类别的变异函数进行分析, 对比两种主成分分析方法的优劣, 一种是直接对4个波段求取第一和第二主成分, 将其作为研究波段, 另一种是特征主成分分析方法^[14], 即将4个波段分成两组, 第一组为可见光波段(1, 2, 3波段), 第二组为红外波段(第4波段), 将分析

得到第一组的第一主成分与第二组的第四波段作为研究波段。图 5 是分别采用两种方法得到的变异函数图,由图可见,使用特征主成分分析法,3 种类型地物特征差异更为明显,故以可见光波段组的第一主成分(记为 PC1)和红外波段(B4)作为纹理计算的基础。

3.2 纹理波段求取与解译

如上所述,分别采用 3×3 、 5×5 和 7×7 窗口计算 PC1 和 B4 的变异函数纹理波段,包括方向变异函数(V)和交叉变异函数(CV),结果采用步长为 1 个像素(19.5 m)在 4 个方向(N-S、E-W、N45E 和 N45W)的平均值。

从该幅影像计算得到 9 个纹理波段,即 3 种不同窗口下 PC1 和 PC2 的两个单变量(V)变异函数纹理波段和一个双变量(CV)变异函数纹理波段。图 6 中 4 幅图是经过实地调查的 2 块区域(50×50 像素),a 和 c 为截取的 CBERS4,3,2 波段合成的假彩色影像,而 b 和 d 则为 7×7 窗口的 PC1 和 B4 的方向变异函数纹理波段和二者的交叉变异函数纹理波段合成的影像。图 6 可见,尽管 a 图红色和浅紫色代表了不同的棉花长势,但是差别不明显,而且等级区分太少,与之对应的 b 图则将其分成了 4 种颜色,紫色和深红色对应长势好的棉田,而黄色代表长势较好的棉田,绿色代表长势一般的棉田,c 图无法分辨长势的差异,其中黑色代表的是其它类别,红色代表的是棉田,长势区分十分不明显,与之对应的 d 图则将其分成了 4 种主要颜色,红色和紫色代表长势好的棉田,黄色代表长势较好的棉田,而分布于棉田周围的蓝色,代表的是长势差的棉田。

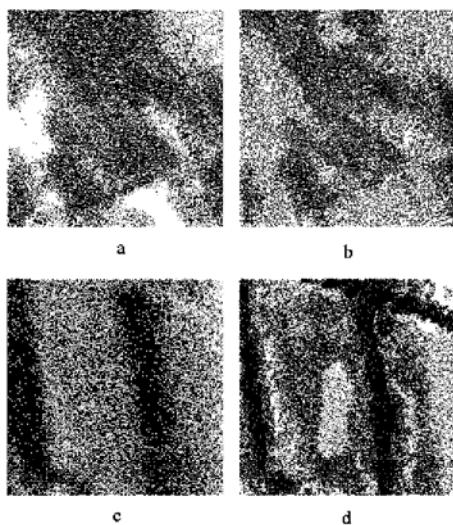


图 6 单纯光谱波段与纹理波段合成对比

Fig. 6 Comparison between combination of spectral bands and combination of texture bands

3.3 精度分析

采用最大似然分类法进行分类,在 6 月中旬的实地调查数据中,从样本点的平均株高、果枝数、果节数和真叶数 4 个方面对棉花长势进行衡量,计算时株高的权重

为 0.4,其他 3 个指标的权重为 0.2,将棉花长势分为好(数值大于 25),较好(20~25),一般(15~20)和差(小于 15)4 个级别,如株高为 31.7 cm ,真叶 10.2 张,果枝 3.4 枝,果节 4.2 个,其长势的数值结果是 16.24,长势即为一般,以此为基础,将影像分成上述 4 个类别,共取 521 个样本,其中训练样本 389 个,检验样本 132 个,因为化控对象主要针对的是旺长区域(长势好一类),故检验样本设有 40 处为旺长区域,使得对其的区分能力可以直接左右最后总的分类精度,进行监督分类时,所选择的训练样本各类别之间并非完全分离的,大都会有重叠发生,在光谱基础上加入纹理的一个重要意义就是能提高训练样本选择的分离度,以式(4)和(5)计算的样本分离度(TD)对训练样本选择进行测度^[14]。

$$D_{ij} = \frac{1}{2} \operatorname{tr}[(C_i - C_j)(C_i^{-1} - C_j^{-1})] \quad \frac{1}{2} \operatorname{tr}[(u_i - u_j)(C_i^{-1} - C_j^{-1})(u_i - u_j)^T] \quad (4)$$

$$TD_{ij} = 2000[1 - \exp(-D_{ij}/8)] \quad (5)$$

式中 i, j ——待比较的两个类别; C_i —— i 类别的协方差矩阵; u_i ——类别的平均值; tr ——矩阵对角线的总和。2000 代表最佳的分离度。大致上 1700 以上表示较好分离效果,表 1 显示的是采用 7×7 窗口的 3 个纹理波段解译的最终选取的样本分离度矩阵表,影像的样本分离度很高。分类精度(使用者精度)矩阵如表 2 所示,尽管 7×7 窗口在长势较好和一般类别上精度偏低,但是其长势好和差较其它窗口却具有更高的分类精度,由于针对化控,主要对象是处于两个端点的类别,因此采用 7×7 窗口计算得到的纹理波段作为分类的基础,最终影像的总体分类精度为 90.53%,部分分类结果如图 7 所示,4 种类别的面积比例为 5:3:13:4。

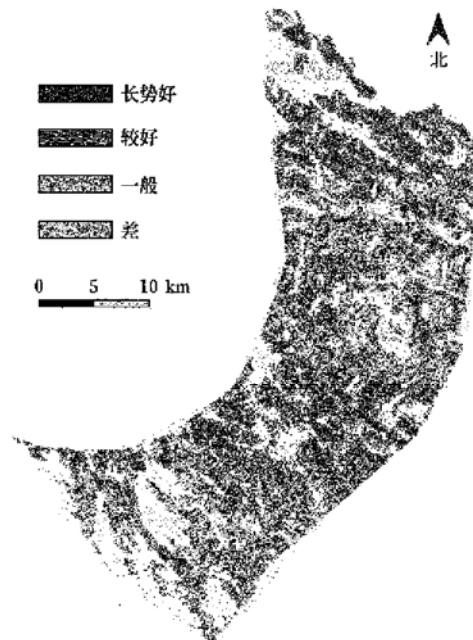


图 7 分类结果

Fig. 7 Result of classification

表1 分离度矩阵表

Table 1 Matrix of transformed divergence

类别	好	较好	一般	差
好	0	1690	1870	1930
较好	1690	0	1830	1800
一般	1870	1830	0	1760
差	1930	1800	1760	0

表2 分类精度矩阵表

Table 2 Matrix of classification accuracy

视窗	好	较好	一般	差
3×3	86.31	91.06	89.33	90.29
5×5	87.47	90.13	88.15	91.34
7×7	91.75	89.33	88.55	92.51

4 结 论

化学调控是棉花生育期内一项重要的投入,合理的投入是节约高产和环保的保证,基于特定场的棉花化控建立于对棉花长势的监测上,与传统的用于估产评价的棉花长势监测不同,它不需要详细获取棉田之间的具体对比值,而只需要得到大区域尺度上棉田长势的空间差异分布结构,因此,应用传统的植被指数方法进行监测非但没有必要,而且由于其不能很好的判别棉花旺长,可能造成化控所要特别针对的旺长棉田投入不够,本文所采用的加入地质统计学影像纹理的分类方法无疑更适合针对化控的棉花长势监测,它能较精确地对不同长势的棉田进行分类,从而反映出整个区域内棉田长势的空间差异,在此基础上,决策者只需要鉴别分类结果图像上几处不同长势类别的棉田喷药量,即可指导制定整个区域内的有层次的喷药方案,极大地减少了人工调查的投入,使得大区域尺度上的化控更为精确合理。

[参 考 文 献]

- [1] 喻树迅,张存信.中国短季棉概论[M].北京:中国农业出版社,2003: 200- 231.
- [2] 马富裕,程海涛,李少昆,等.高产棉花打顶调控的群体适宜性研究[J].中国农业科学,2004,37(12): 1843- 1848.
- [3] 李少昆,张旺峰,马富裕,等.北疆超高产棉花生理特性研究[J].作物学报,2000,26(4): 508- 512.
- [4] 慕自新,马富裕,郑重,等.膜下滴灌棉花随水根施多效唑的效果分析与化控技术的探讨[J].农业工程学报,2004,20(4): 54- 57.
- [5] 刘爱霞,王长耀,刘正军,等.基于RS与GIS的干旱区棉花信息提取及长势监测[J].地理与地理信息科学,2003,19(4): 101- 104.
- [6] 吕建海,陈曦,王小平,等.大面积棉花长势的MODIS监测分析方法与实践[J].干旱区地理,2004,27(1): 118- 123.
- [7] Moran M S, Clarke T R, Inoue Y. Estimating crop water deficit using the relation between surface-air temperature and spectral vegetation index [J]. Remote Sensing of Environment, 1994, 49(3): 246- 263.
- [8] 武建军,杨勤业.干旱区农作物长势综合监测[J].地理研究,2002,21(5): 593- 598.
- [9] F. Lopez-Granados, M. Jurado-Exposito. Using geostatistical and remote sensing approaches for mapping soil properties[J]. European Journal of Agronomy, 2005, 20(1): 1- 11.
- [10] McGwire K, Friedl M, Estes J E. Spatial structure, sampling design and scale in remotely-sensed imagery of a California savanna woodland[J]. International Journal of Remote Sensing, 1993, 14(11): 2137- 2164.
- [11] Miranda F P, Fonseca L E N, Carr J R. Semivariogram textural classification of JERS-1(FUYO-1) SAR data obtained over a flooded area of the Amazonian rainforest[J]. International Journal of Remote Sensing, 2000, 19(3): 549- 556.
- [12] 周天颖.多元尺度影像与半变异元纹理法于都市水稻田分类之研究[D].台湾逢甲大学土地管理系,2004: 51- 66.
- [13] Carr J R, Miranda F P. The Semivariogram in comparison to the co-occurrence matrix for classification of image texture[J]. IEEE Trans. On Geo. And Remote Sensing, 1998, 36(6): 1945- 1952.
- [14] Lark R M. Geostatistical description of texture on an aerial photograph for discriminating classes of land cover [J]. International Journal of Remote Sensing, 1996, 17, 2115- 2133.
- [15] M Chica-Olmo, F Abarca-Hernandez. Computing geostatistical image texture for remotely sensed data classification[J]. Computer and Geosciences, 2000, 26: 373- 383.

Cotton growth monitoring during chemical control stage using geostatistical image texture: a case study of Shihezi

Liao Chujiang¹, Wang Changyao¹, Li Hong², Yang Pengrun²

(1. State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing 100101, China; 2. Shihezi General Institute of Technology Development, Shihezi 832001, China)

Abstract: The chemical control is an important input during the course of the growth of cotton, for a great scale region, the date of chemical control usually refers to some solid season, regardless of the spatial difference among cotton fields. Under the framework of precision farming, this paper recommends using geostatistical image texture to monitor the growth condition of cotton, guiding the spray of chemical medicine. Comparing with tradition method based on vegetation index, although it could not calculate the concrete value of cotton growth condition, it could precisely obtain the spatial distribution of growth condition, what is more, it could identify the cotton field which is at the vigorous growth stage, and which is not able to be distinguished by the method based on vegetation index, so it is a better method towards guiding to chemical control of cotton.

key words: Geostatistics; image texture; variogram; chemical control of cotton; CBERS