

# 一种基于图像分析提取作物冠层生物学参数的方法与验证<sup>\*</sup>

雷咏雯<sup>1,2</sup>, 王娟<sup>2</sup>, 郭金强<sup>2</sup>, 危常州<sup>1,2\*</sup>

(1. 石河子大学农学院资源与环境系, 石河子 832000; 2. 石河子大学新疆绿洲生态农业重点实验室, 石河子 832000)

**摘要:** 利用图像分析方法, 通过准确识别冠层和背景像素进行棉花冠层生物学产量和叶面积系数估测。采用 Olympus C740 Ultra Zoom 数码相机拍摄棉花不同生育期冠层图像, 在棉花冠层数码照片特征分析的基础上提出了棉花冠层图片计算机自动判读的方法, 即混合采用图像色度(H)、绿光(G)、红光(R)灰度值构造提取条件, 通过多重判断识别棉花冠层和背景, 并编写了相应的计算机程序。利用该程序分析棉花不同施氮量下、不同生育期提取地面覆盖度参数与棉花生物学产量、叶面积系数间的关系, 发现棉花冠层地面覆盖度指标可以有效预测棉花生物学产量和叶面积系数, 二者间指数相关系数达到  $r=0.97$  以上, 为极显著相关。

**关键词:** 图像分析; 作物冠层; 生物学参数; 判别条件

中图分类号: TP79; S363

文献标识码: A

文章编号: 1004 1389(2006)03 0045 05

## Digital Image Analysis for Estimating Crop Canopy Parameters Based on Object Features and on Farming Validation

LEI Yong wen<sup>1,2</sup>, WANG Juan<sup>2</sup>, GUO Jin qiang<sup>2</sup> and WEI Chang zhou<sup>1,2\*</sup>

(1. Dept. of Resource and Environment, Agronomy College, Shihezi University, Xinjiang 832000, China; 2. The Key Oasis Eco Agriculture Laboratory of Xinjiang Production and Construction Group, Shihezi University, Xinjiang 832000, China)

**Abstract:** This research is attempted to buildup an innovational method based on object features of cotton canopy digital images to identify plant and background pixels and hence to estimate percent ground cover of vegetation (PGCV) and leaf area index (LAI). Cotton canopy was captured by using an Olympus C740 Ultra Zoom digital camera. A protocol mixed hue of HLS color space with R, G value was setup and multiple judgment process was designed to extract cotton canopy pixels from background noise. A computer program based on the protocol above was designed simultaneity. Relation between PGCV that derived from the designed program and cotton canopy biomass or LAI with different N application and in different growth stages suggested that PGCV could reliably evaluate both PGCV and LAI, empirical statistical showed that a significantly high coefficient between PGCV and cotton canopy biomass or LAI reached ( $r=0.97$ , respectively). The results indicated that image analysis is a promising method for quick diagnosis of crop growth and development.

**Key words:** Image analysis; Crop canopy; Biological parameters; Judgment protocol

多光谱、高光谱为手段的遥感在估计地面作物覆盖度、总生物量、产量、吸氮量等生物参数方面已有了大量研究, 取得了很多成果<sup>[3, 5, 6, 8, 10]</sup>, 但

是这些遥感手段仪器设备价格高, 且受分辨率、时间周期、地理、空域、气象条件等限制, 在我国目前的农业生产中难以投入实际应用。随着数码相机

\* 收稿日期: 2005 11 27 修回日期: 2005 12 29

基金项目: 国家自然科学基金资助项目(批准号: 30560069); 农业部“948”重大专项资助项目(批准号: 2003 Z53); 新疆兵团科技计划项目(NKB03SDXNK11ZP)资助项目。

作者简介: 雷咏雯(1968)女, 主要从事农业信息化和农业气象工作。电话: 0993 2306697。

\* 通讯作者: 危常州, 男, 39岁, 教授, 农学博士。从事植物营养学、信息农业研究。Email: weichangzhou@163.com, 电话: 0993-2058227。

技术快速发展,其空间分辨力和存贮空间迅速增加且价格低廉,是一种十分易于采用和普及的方法。目前采用数码照片分析进行作物生长分析和监测已开展了较多研究,其核心是作物冠层与背景噪音的识别与剔除。研究者采用的方法可以分为采用商业图像分析软件和自行设计软件两种,且均获得肯定的结果。如 Lukina 等<sup>[7]</sup>采用 Micrografx Picture Publisher 处理图像,采用植被覆盖比例(Percent Ground Cover of Vegetation, PGCV)的概念估算小麦冠层覆盖度并估计了冬小麦冠层生物量。Adamsen<sup>[2]</sup>采用 JPEG 格式采集小麦数码影像,然后采用 Paint Shop Pro version 3.12 软件将 JPEG 格式的文件转化成 GIF 格式,使颜色降到 256 色。然后裁减外围图幅,采用 G/R 表征小麦叶绿素含量。Ewing<sup>[4]</sup>设计了基于 DOS 的软件 RGBCal、DyEye,通过交互调节红绿蓝颜色,色调、明度和饱和度的阈值及红绿蓝的主成分分析估测了玉米的覆盖度。

这些成果表明,数字图像分析技术在农业遥感中具有潜在广泛的应用前景。但是其估测精度受光照条件的影响,如在阴影下精度较高,而在阳光下精度较低<sup>[7]</sup>,另外这些方法需要大量的人工干预,去除背景还带有主观性<sup>[1]</sup>。本文以棉花作物为例,通过数字影像获取和不同对象像素特征分析,并与田间实测参数对比,试图提出一种自动的更精确和更客观的估测作物覆盖度的技术,为估测作物生长参数提供一种快速可靠的方法。

## 1 试验设计和研究方法

### 1.1 田间试验设计

2003 年在石河子大学农学院试验站设置不同氮水平的棉花膜下滴灌试验,以获得不同氮素营养状态的棉花冠层图片。N 肥设置 0、180、270、360 kg/hm<sup>2</sup> 共 4 个水平,随机区组 3 次重复,小区面积 48.6 m<sup>2</sup>,共计 12 个小区,各小区均设有保护行,采用单独的滴灌设备以便于试验控制。30%氮肥作为基肥播前施入,70%随水滴施。供试品种为新陆早 13 号。采用“干播湿出”法播种,其他田间管理措施与当地常规栽培措施相同。

### 1.2 测试项目与方法

棉花生物学产量测定是在棉花不同生育时期随机采取 10~4 株棉花样品(随生育进程减少采样株数),称量鲜重,样品带回室内烘干法测定干重并根据棉花亩密度换算出棉花单位面积生物学

产量。

叶面积系数测定采用 ADC 公司生产的 AM2000 型光电叶面积仪测定并结合打孔结合烘干法校正。

### 1.3 棉花冠层图像摄取

棉花田间冠层数码图像摄取采用 Olympus C740 Ultra Zoom 型数码相机在晴天获取。拍摄时间选择在中午 12:00~13:00,拍摄时设定相机为强制曝光模式,快门速度 1/250 s,不设定白平衡。在棉花出苗后 40 d, 60 d, 70 d, 80 d, 90 d, 110 d 后分别拍摄棉花冠层图像,每小区每次拍摄 3 幅图像。影像采用 JPEG 格式存储,文件大小为 2048×1536 像素。由于数码相机采用中心投影,边缘投影随着视场角的增大畸变增大,在本研究中以影像中心为起点,放射状分析图幅 1/4 矩形面积的光谱特性。

## 2 冠层图像特征分析与程序设计

本研究采用 HLS 颜色空间结合 RGB 颜色空间分析棉花冠层光谱特征,以色调(H)作为对象识别的基本指标,以红光(R)、绿光(G)值的相对大小构造辅助识别指标,实现棉花冠层准确、快速自动识别,并且建立生物学产量、叶面积系数(LAI)与冠层像素指标间的数学模型。

### 2.1 棉花冠层的影像特征分析

基于棉花冠层光照条件和光反射特性,可将影像分成 7 个子类:①没有热点、阳光直射的棉花;②有“热点”的棉花;③遮阴的棉花叶片;④白色的花;⑤红色的花;⑥阳光直射的土壤;⑦遮阴土壤。所谓“热点”是棉花在阳光直射下叶片如果与相机成 90°方位角时由于光反射强烈而在照片上使绿色叶片表现为白色的现象。研究发现地面死亡的植物残余物如上季遗留的枯枝落叶、地膜等与土壤的光谱反射特征相同,故不需要对它们设立新的子类,可以全部归入土壤背景中。

### 2.2 棉花冠层光谱特征提取算法

光谱特征提取技术的核心是尽量提取与棉花冠层有关光谱信息,而扣除无关背景值。根据棉花冠层影像特点,采用下述原理提取光谱参数。

#### 2.2.1 没有热点、阳光直射的叶片和背景的识别

阳光直射下棉花叶片在红光(R)和蓝光区(B)强烈吸收而在绿光区(G)吸收较少,反射较强,故 G 值大于红区和蓝区反射值,棉花冠层叶片像素满足  $G > R$  且  $G > B$ ,但是  $R \ll 220$ ,  $G \ll 220$

(表 1)。

表 1 阳光照射下土壤和背景图像像素的光谱特征值( $n=60$ )

Table 1 Characteristic spectral value as cotton leaf and other background radiated directly by sunlight ( $n=60$ )

对象 Object	R	G	B	H
叶片 Leaf	151±49	164±47	160±48	77±4
枯叶和其他有机残余物 Organic residues	175±35	158±38	148±42	32±4
地膜膜面 Surface of plastic film	145±13	144±9	162±18	29±3

注: \* SE 是对 60 幅图片, 900 个像素点的统计结果。由于叶片方位角不同, 其 RGB 值不同, 故 SE 差别较大。对于单独的像素点, 阳光直射的叶片严格符合  $G > 0$  而阳光直射的背景严格符合  $G < 0$ 。表 2, 表 3 同。

Note: \* SE stands for the average value of 900 pixels from 60 images. Due to azimuth angle difference, RGB value of different objects were varied, but to a single pixel, cotton leaf perfectly fit  $G > 0$  and background meet  $G < 0$  if directly radiated by sunlight. Values in table 2 and table 3 had the same case.

背景(土壤、秸秆、枯叶或地膜)在苗期因为裸露面积大, 土壤干燥呈现土黄色, 在其他时期成灰或灰褐色, 虽然它们 RGB 值也较高, 有时还大于冠层的 RGB 值, 但其像素值满足  $G < R, G < B$ 。因此采用  $G > 0$  即可有效识别阳光直射的叶片和背景,  $G > 0$  为冠层像素,  $G < 0$  为背景。

将 RGB 颜色空间转化成 HSB 空间, 发现封行前棉花冠层色调(H)介于 60~80 之间, 而残留物 H 在 10~40 之间, 考虑到 60~120 是 HLS 颜色空间的标准绿色区间, 且其它子类 H 区间均不会与此范围重叠, 因此采用  $H = \{60, 120\}$  和  $G > 0, G > 0$  作为没有热点、阳光直射的叶片和背景的认识条件。研究发现此条件可以完美识别和分离阳光直射的植物冠层与土壤背景。

2.2.2 有“热点”的棉花叶片 热点的 H 值一般均满足  $120 > H > 60$ , 但有少数“热点”的 H 值小于 60。形成典型“热点”的叶片一般是棉花的倒 3 叶或倒 4 叶, 这时叶片 RGB 像素值均大于 240, 多数情况下接近 255。在阳光非常强烈的情况下, 下部叶片也出现类似于“热点”的部位, 这些点符合  $120 > H > 60, R > 0$  没有规律, 对于此类像素点采用  $G > 0$  不能正确判别。但其 R、G 值大于 220 小于 240。因此采用统一的识别特征  $H = \{60, 120\}$  和  $G > 220, G > 220$  可以正确识别“热点”和类似“热点”的像素点。

由于“热点”的 RGB 值都不能反应叶片的实际颜色值, 故以棉花倒三叶和倒四叶 RGB 平均值取代图像中的值, 经过多幅图像分析, 认为设为  $R = 120, G = 140, B = 90$  较为合理, 但不同生育期存在差别。

表 2 “热点”和花朵像素的光谱特征值

Table 2 Characteristic spectral value of “Hot dot” and flowers

对象 Object	R	G	B	H	n
典型“热点” Typical “Hot dot”	254±1	254±1	238±1	67±4	78
类似“热点” “Hot dot” alike	228±7	229±14	209±12	65±4	92
白色花 Flower in white	251±2	249±3	245±3	63±2	150
红色花 Flower in red	252±2	176±26	210±14	337±6	150

注: n 是统计样本数。

Note: n stands for number of statistics samples.

2.2.3 阴影下的棉花和土壤的识别 不同时期棉花冠层和土壤 RGB 值差异较大, 蕾期叶片符合  $G > 0$ , 土壤符合  $R < 0$  的特征, 但是在其他生育期该 2 类对象均满足  $G > R, G > B$  且符合  $120 > H > 60$ (表 3)。但是研究发现遮阴叶片所有生育期平均  $G > 6$ , 而背景  $G < 2$ , 这一分类特征十分稳定, 因此, 加入限制条件  $G > 2$  可以很好识别遮阴的土壤和叶片,  $G > 2$  为遮阴叶片,  $G < 2$  为遮阴土壤。该指标也可满足蕾期的判别, 故在全生育期可采用相同的识别特征。

2.2.4 白色花的判别 白色花的特征是其 H 值介于 60~120 之间, 这表明其内部还是含有一定的绿色素, 但肉眼不能识别。其 RGB 值均大于 240(表 2)。其特点和“热点”的光谱特征相似, 可以采用与“热点”相同的指标判别。

2.2.5 红色花的判别 红花的特征是其 H 值介于 300~360 之间, 这是其和其他子类明显的区别(表 2)。简单以  $360 > H > 300$  即可以很好区别其与其他子类。

表 3 阴影下叶片和土壤的光谱特征值 ( $n=60$ )

Table 3 Characteristic spectral value of shaded leaf and soil ( $n=60$ )

		R	G	B	H
蕾期 Bud	叶片 Leaf	106±34	120±25	92±53	76±12
	土壤 Soil	120±13	105±12	108±15	68±7
盛花期 Flourishing flower	叶片 Leaf	29±9	35±9	27±7	78±5
	土壤 Soil	38±3	39±3	34±2	67±7
盛铃期 Flourishing boll	叶片 Leaf	21±6	27±7	19±5	82±12
	土壤 Soil	27±2	29±3	27±1	61±18

综合以上分析,虽然棉花冠层分类成员和其他作物相比较多,但是如果采用色度判别结合RGB辅助判别,则可以很好地解决背景干扰的问题,从而解决冠层像元提取的问题。

### 2.3 通用棉花冠层提取计算机算法

根据上述提取特征,构造了通用棉花冠层自动提取算法:①在 Visual Basic 下采用循环语句结合 point 函数逐像元读取像元 RGB 值,存储在数组中。②对 RGB 值进行转换,获得个像素对应的 HSB 值,方法如下: $L(\text{ight}) = (\text{PixMax} \times 100) / 255$

$S(\text{aturation}) = ((\text{PixMax} - \text{PixMin}) / \text{PixMax}) \times 100$

$H(\text{ue}) = (2n + (\text{Red} - \text{Green}) / (\text{PixMax} - \text{PixMin})) \times 60$

$\text{PixMin}) \times 60$

如果  $H < 0$  则  $H = 360 + H$

$n$  的取值方法如下:

If  $\text{PixMax} = \text{Red}$ , then  $n = 0$ ,

If  $\text{PixMax} = \text{Green}$ , then  $n = 1$

if  $\text{PixMax} = \text{blue}$ , then  $n = 2$

上式中,  $\text{PixMax}$  为 RGB 空间中灰度最大像元值,  $\text{PixMin}$  为 RGB 空间中像元灰度最小值,  $n$  取决于 RGB 空间中哪类像元获得了最大值。

按照以下通用算法提取棉花冠层像元信息(图1),如果像元符合某判断条件,则计入相应的类别。(1)到(5)子类合并构成了棉花冠层像素信息,(6)和(7)子类构成了土壤背景“噪音”信息,予以剔除。

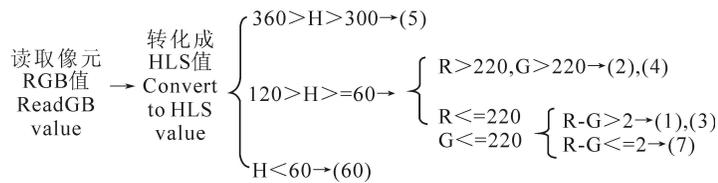


图1 棉花冠层识别的通用算法

Fig. 1 The general protocol of cotton canopy identification

### 2.4 棉花冠层信息自动提取程序的设计

在上述算法基础上用 VB 开发了一个自动提取棉花数码照片像元信息的计算机程序。该程序可以自动进行颜色的单色反演,可以按照影像的饱和度(S)、色度(H)、亮度(L)或像素点 R、G、B 色阶值的任意一种或多种构建过滤条件或构建复合过滤条件,以方便比较不同过滤条件的识别效果,也能根据研究的需要选择不同影像区域进行分析识别。

程序包括 3 个部分。第 1 步采用 RGB 和 HSB 函数读取图像像元信息和存储,第二步采用通用像元过滤条件对各像素进行筛选,对 RGB、HSB 值进行累进统计。累进数据采用数据文件输出到外部文件,可以利用 Office 办公软件或其他数据处理软件的强大数据计算功能处理输出的数据。第三步采用绘图语句对提取的有效像素(冠层像素点)采用 Pset 语句在新窗口绘图,检验过滤效果,如果不能令人满意,则重新开始构造过滤条件。

该程序是图形用户界面,操作简便。操作流程为:①运行程序,②打开一个影像文件,③构造过滤条件,④选择“执行”命令执行操作,同时数据文件被储存到一个数据文件中。重复上述 4 个步

骤,得到一系列数据,在数据处理软件中进行数理统计和比较分析。

### 3 利用地面覆盖度预测叶面积系数和生物学产量

采用有效像素(冠层像素总量)与图像像素总量的比值作为地面覆盖度(Percent Ground Cover of Vegetation, PGCV)<sup>[7]</sup>,这样可以克服由于光照条件不同导致 RGB 值的差异。由于各影像的像素总量均为  $2048 \times 1536$  像元,且均采用以成像中心为原点的 1/4 面积计算,因此 PGCV 的变化趋势实际上等同于冠层总像素变化趋势。地面覆盖指数与实测对应的棉花冠层叶面积指数、干物质生产量的关系见图 2、图 3。

地面覆盖指数与棉花冠层叶面积系数、生物学产量为指数相关,相关系数达到 0.971 和 0.975,均为极显著相关,从数据离散程度看,PGCV 低于 0.6 时其拟合精度较高,而大于 0.6 以后离散程度增加,这与 PGCV 的最大值为 1,棉花封行以后枝叶互相交叠识别度降低有关。这一结果与 Mages<sup>[9]</sup>在小麦上的研究结果吻合,但相关系数更高。因此,采用图像判别分析可以很好地预测叶面积系数和棉花生物学产量,但是可以

大幅度降低田间工作量, 工作效率明显提高。

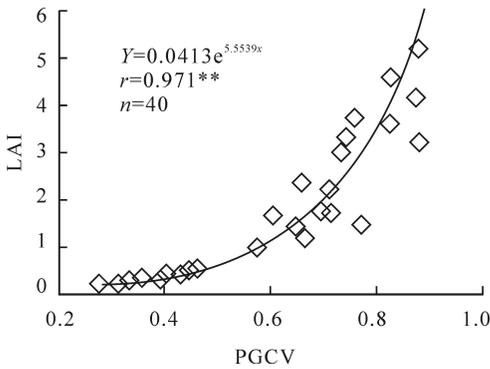


图2 地面覆盖度与叶面积系数间的关系

Fig. 2 Relation between PGCV and LAI

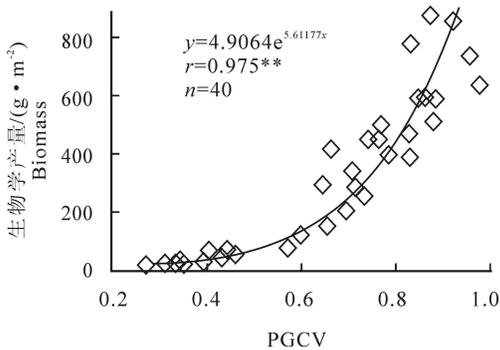


图3 地面覆盖度与生物学产量的关系

Fig. 3 Relation between PGCV and cotton biomass

## 4 讨论

本研究提出对棉花冠层数字图片进行计算机自动判读的方法, 即混合采用色度(H)、RGB 指标构造提取条件, 通过多重判断识别棉花冠层和背景, 并编写了相应的计算机程序。利用该程序分析棉花不同施氮量下、不同生育期提取冠层光谱参数与棉花生物学参数的关系, 发现棉花冠层地面覆盖度指标可以有效预测棉花生物学产量和叶面积系数, 二者间指数相关系数达到  $r = 0.97$  以上, 为极显著相关。该成果对棉花作物生长发育快速监测具有重要的价值。

本研究中对于阴影下的土壤和作物的识别主要依靠 G-R 差值大小, 这个指标在本研究条件下十分稳定, 判别效果很好, 但对于其他作物或其他生态区是否同样适用尚没有检验。

本研究提出的算法对于“热点”的识别较为精确, 但对于少量  $H < 60$  的“热点”会错误识别为背景加以扣除。对于白花和红花的识别也基本可

靠, 但在红花的边缘有少量像素 H 值小于 300, 在白花中则有个别点 H 为 0 或很小的值, 造成误判, 该部分占棉花冠层面积不大, 对地面覆盖度识别的影响不大, 但对于其形成的原因尚没有发现合理的解释, 应该努力寻找原因并加以解决。另外对于白花的识别采用和“热点”相同的识别条件, 使判断条件简化, 但是如果在细致的冠层成员分析中应该设计精确的判别条件对它们进行区别。

## 参考文献:

- [1] 李存军, 王纪华, 刘良云, 等. 基于数字照片特征的小麦覆盖度自动提取研究[J]. 浙江大学学报(农业与生命科学版), 2004, 30(6): 650~656.
- [2] Adamsen F J, Pinter P J, Barnes J E M., *et al.* Measuring wheat senescence with a digital camera[J]. *Crop Sci.*, 1999, (39): 719~724.
- [3] Dowell F E. Automated Color Classification of Single Wheat Kernels Using Visible and Near Infrared Reflectance[J]. *Cereal Chemistry*, 1998, 75(1): 142~144.
- [4] Ewing R P, Horton R. Quantitative Color Image Analysis of Agronomic Images[J]. *Agron. J.*, 1999 (91): 148~153.
- [5] Gerard B, Buerkert A. Aerial photography to determine fertiliser effects on pearl millet and *Guiera senegalensis* growth[J]. *Plant and Soil*, 1999(210): 167~177.
- [6] Li H, Lascano R J, Barnes E M, *et al.* Multispectral reflectance of cotton related to plant growth, soil water and texture and site elevation[J]. *Agron. J.*, 2001 (93): 1327~1337.
- [7] Lukina E, Stone M, Raun W. Estimating vegetation coverage in wheat using digital images[J]. *J. Plant Nutri* 1999, 22(2): 341~350.
- [8] McMurtrey J E, Chappelle E W, Kim M S, *et al.* Distinguishing nitrogen fertilization levels in field corn (*Zea mays* L.) with actively induced fluorescence and passive reflectance measurements[J]. *Remote Sens. Environ.*, 1994 (47): 36~44.
- [9] Moges S M, Raun W R, Mullen R W, *et al.* Evaluation of green, red, and near Infrared bands for predicting winter wheat biomass, nitrogen uptake, and final grain yield[J]. *J. plant nutri* 2004, 27(8): 1431~1441.
- [10] Stenberg B, Jensen L S, Nordkvist E, *et al.* Near infrared reflectance spectroscopy for quantification of crop residue, green manure and catch crop C and N fractions governing decomposition dynamics in soil[J]. *J. Near Infrared Spectrosc* 2004 (12): 331~346.