

基于数码照片的草地植被覆盖度快速提取方法

胡健波¹, 张璐², 黄伟¹, 吴世红¹, 刘长兵¹

(1. 水路交通环境保护技术交通行业重点实验室 交通运输部天津水运工程科学研究所, 天津 300456;

2. 河南理工大学测绘与国土信息工程学院, 河南 焦作 454000)

摘要:草地植被覆盖度是草地健康与否的重要参数,提高草地植被覆盖度的计算效率和精度具有重要的意义。本研究提出了一种利用过绿特征植被指数和半自动阈值设定算法(半自动阈值法)的从数码照片中快速计算草地植被覆盖度的方法,并将该方法与最大似然法监督分类方法(最大似然法)和色度饱和度法进行了比较。对 32 张草地数码照片的植被覆盖度估算结果和准确结果进行回归分析,分别得到斜率、截距和回归系数这 3 个参数。结果表明,3 种方法都未能考虑植物的非绿色组分而存在低估的问题。半自动阈值法的准确度堪比最大似然法,且没有后者存在的低覆盖率高估问题;而且该方法人工干预少,计算结果准确客观,适用性强;但是对绿色特征不明显的植物(如灰绿色植物)效果不佳。

关键词:数码照片;植被覆盖度;过绿特征;半自动阈值

中图分类号:S812

文献标识码:A

文章编号:1001-0629(2011)09-1661-05

*¹ 我国西北地区气候干旱、多风,生态环境十分脆弱。草地植被具有改善气候、防止水土流失等重要生态功能。植被覆盖度是指包括乔、灌、草和农作物在内所有植被的冠层、枝叶在生长区域地面的垂直投影面积占生长区域面积的百分比^[1],是生态系统健康与否的重要指示因子^[2]。植被覆盖度测量方法的好坏,直接决定了测量结果的准确度和可信度。草地植被覆盖度的实地测量方法较多,根据采用手段不同,可大致分为目估法、采样法和仪器法^[3]。

目估法是传统的快速植被覆盖度估测方法,虽然简单易行,但是主观随意性较大,测算结果跟目估对象的实际覆盖度大小及测量人的经验密切相关;有研究表明,个人目估最大绝对误差可达 40.4%^[4]。采样法是通过设定样方,将一根根样针在草地中垂直投下,记录草叶被样针击中的数目,计算击中数目占总样针数的比值作为草地覆盖度^[5],这种方法虽然准确,但是操作复杂,且十分耗时。仪器法是使用专门的测算草地植被覆盖度的仪器来进行作业,如空间定量计、移动光量计和照相法,测量速度较快,结果客观;其中照相法因其具有简单、经济等优势,更是应用广泛^[6]。随着数码相机的普及以及数字图像处理技术的快速发展,照片法已越来越多地应用于草地覆盖度的实地调查工作中。

池宏康等^[7]利用 Photoshop 图像处理软件手动

提取数码照片中绿色像素,虽然和目估法相比工作量有所降低,但是仍然较费时费力,且同样存在主观随意性较大的缺点。因此,快速自动的植被覆盖度估算是近年来照相法的研究热点。Zhou 和 Robson^[8]利用数码照片,通过结合光谱和纹理的非监督分类方法提取了草地覆盖度,结果证明该方法比单纯的利用 k-means 非监督分类和最大似然法监督分类,在准确程度上有了很大的提高。宋雪峰等^[9]用数码相机获取草地样方照片,从样方照片数据中提取 6 项指标,通过分析建立逻辑判别模型计算植被覆盖度,总体精度达到 94.7%。张清平等^[10]利用颜色分析软件 WinCAM,选定一致的植被与非植被的颜色标准,通过颜色比对提取植被并求得植被覆盖度。张学霞等^[11]通过选取植被和非植被的感兴趣区域,分析各自的光谱信息规律,利用线性混合像元分解的方法来获取植被覆盖度。钱金波^[12]计算过绿特征($\text{Excess Green} = 2 \times G - R - B$)植被指数,并主观设定阈值,区分植被和非植被,从而求得植被覆盖度。张云霞等^[13]把数码照片中的绿色大

* 收稿日期:2010-11-10 接受日期:2010-12-27
基金项目:中央级公益性科研院所基本科研业务费专项资金(TKS090303)

作者简介:胡健波(1982-),男,浙江宁波人,助理研究员,博士,主要从事环境遥感应用和图像处理算法研究。
E-mail: whogamble@qq.com

通信作者:刘长兵 E-mail: l. c. b@263.net

于红色和蓝色并且绿色大于特定阈值的像素划分为绿色植被。然而目前的这些草地植被覆盖度方法均存在参数设定主观,对照片拍摄时间和光照条件要求较高等缺点,限制了数码照片在草地植被覆盖度调查中的推广应用。本研究提出了一种利用过绿特征植被指数和半自动阈值设定算法(半自动阈值法)的从数码照片中快速计算草地植被覆盖度的方法,并将该方法与最大似然法监督分类方法(最大似然法)和色度饱和度法进行了比较,旨在提出一种更加客观、适用性更强的草地植被覆盖度提取方法。

1 材料与方法

1.1 数码照片采集方式 以西藏阿里地区国道219日土至区界段公路改建项目的环境影响评价工作为依托,项目地处西北干旱地区,公路沿线以低覆盖度草地为基调,少量高覆盖度的草甸只分布在河流湖泊周边。试验照片以公路沿线两侧的自然草地植被为主要拍摄内容,也有部分村镇的人工植被。

本研究提出的方法对数码相机的性能要求很低,任意一款能够自动调焦的普通家用数码相机即可满足要求。本研究中所有照片和数据均来自索尼

DSC-V1数码相机,照片尺寸为2 592像素×1 944像素。拍摄时两臂向前持平拿稳相机,保持垂直向下的拍摄角度和大约1.5 m的拍摄高度。拍摄时注意:当拍摄地点存在叶片细小或狭长的植物时,可适当降低拍摄高度;当拍摄时光照条件不好时(如阴天或者非正午前后的一定时间内),利用闪光灯作为补充光源,以减少植物叶片互相遮挡产生的阴影。

1.2 数码照片植被覆盖度提取算法

1.2.1 算法原理 数码照片中的植被提取主要分两个步骤:首先,构建一个能够突出植被特征并抑制非植被背景的植被指数(图1);然后,设定最佳的阈值,将植被指数大于阈值的像素归为植被,小于阈值的像素归为背景。植物在可见光波段附近的反射光谱呈现出蓝色、红色强烈吸收,近红外和绿色强烈反射的特征;基于此特征,目前植被遥感领域已提出大量的植被指数,如归一化植被指数(NDVI)等^[14]。鉴于普通的家用数码相机拍摄的照片不存在近红外波段,本研究选择过绿特征(Excess Green=2×G-R-B)作为植被指数^[15]。确定植被指数后,最终分类结果的好坏取决于第2步的最佳阈值设定。

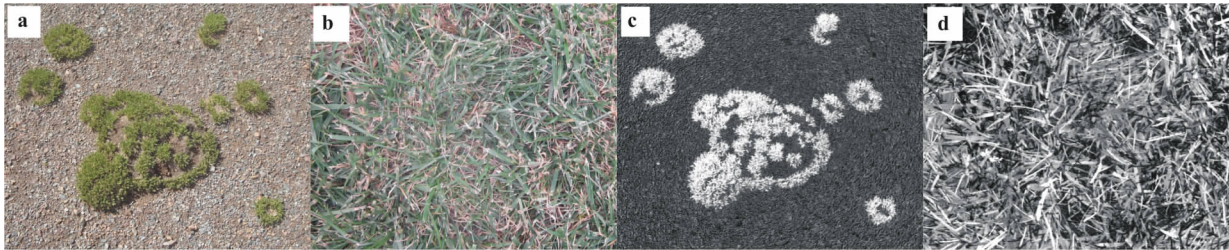


图1 样片a、b及对应的过绿特征植被指数

注:a、b分别是在晴天和阴天两种光照条件下的照片,c、d分别为a、b的过绿特征植被指数图。

在植被指数图中指定一条直线,得到一条过绿特征剖面曲线(图2)。以图2为例,在理想情况下,只有在植被和背景交界处才存在变化梯度,阈值的设定十分简单,阈值在-15~60随意设定均能得到几乎同样的准确结果;而在实际情况下,无论是植被还是背景,均存在锯齿状的随机噪声,合理的阈值应当设定在20~30。在绝大多数实际情况下,背景噪声凌乱,人工主观设定最佳的阈值十分困难,需要不断的试错才能逐步逼近最佳值;即便如此,植被覆盖度提取结果也会因人而异。本研究提出一种半自动阈值设定方法设定最佳阈值。

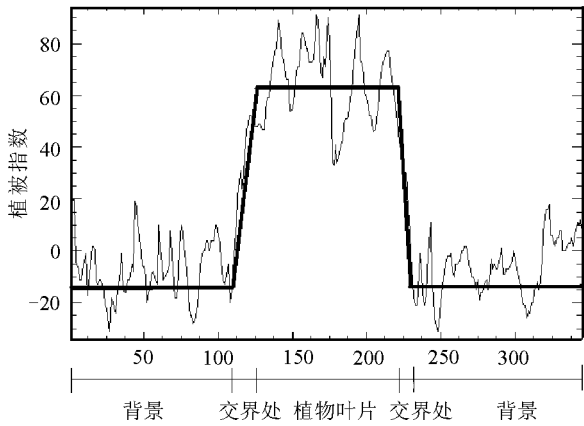
1.2.2 植被指数半自动阈值设定 植被指数半自动

阈值设定方法基于如下假设:植被像素的植被指数较大,背景像素的植被指数较小,而且各自的随机噪声引起的平均波动梯度小于植被和背景交界处的变化梯度(图2)。当设定阈值等于t时,得到的结果是一幅二值图(1代表植被、0代表背景),处于1和0像素交界处的像素为边缘像素;图2中,边缘像素为剖面曲线与平行于x轴的直线(植被指数=t)交点。边缘像素的平均梯度强度计算公式如下:

g=f_v×g_v+f_b×g_b+f_j×g_j (1)

式中,g表示整张照片的边缘像素的平均梯度强度,f_v、f_b、f_j和g_v、g_b、g_j分别表示落在植被、背景和边界区域的边缘像素百分比以及边缘像素的梯度强

度。 f_v 、 f_b 、 f_j 总和等于 1, g_j 大于 g_v 和 g_b (根据前面的假设)。



—表示无任何噪声的理想情形, —表示实际情形。

图 2 植被指数剖面曲线示意图

最佳的 t 应当使边缘像素落在植被和背景的交界处, 同时会尽可能避免落在背景和植被区域, 即 f_j 尽可能的大, 而 f_v 和 f_b 尽可能的小。以一定的步长不断尝试不同的 t , 计算 g ; 根据公式(1)可知, t 最佳时即为 g 最大时。

然而, 考虑到图像中存在不可预料的噪声, t 值只能在某个范围内选取, 而不能在植被指数最小值至最大值之间选取。如图3所示, 当 t 小于 -50 或

者大于 100 的时候, 边缘像素数量极少, 但是平均梯度强度却很大; 说明这个范围内的个别像素与周边像素的植被指数差异很大而数量又很小, 从而产生较大的 g , 属于不应考虑的噪声点。因此需要人为设定 t 的选取范围, 本研究建议设定在边缘像素平均梯度强度分布图的凹谷处, 如样片 a 的选取范围为 $0 \sim 80$, 样片 b 的选取范围为 $-30 \sim 80$ 。本研究的草地植被覆盖度提取方法之所以不是一种全自动的方法, 就是因为需要人为设定这个阈值选取范围进行主观干预。

最佳阈值确定之后, 即可得到植被和背景分类图(图 4)。植被的像素数与整张数码照片的像素数之比即为该数码照片的植被覆盖度。样片 a、b 的植被覆盖度分别为 12.50% 和 67.01%。

1.3 精度验证与方法比较 利用 32 张依托公路沿线的草地数码照片作为样本对色度饱和度法^[16]、最大似然法监督分类方法(简称最大似然法)以及本研究的半自动植被指数阈值设定法(半自动阈值法)这 3 种植被覆盖度提取方法进行精度验证。对每一张数码照片, 手动勾勒出植被像素, 并将所得植被覆盖度作为准确结果。对所有样本数码照片, 将 3 种方法的计算结果分别与准确结果进行线性回归分析, 比较回归系数、斜率和截距这 3 个参数, 分析 3 种方法的优缺点。

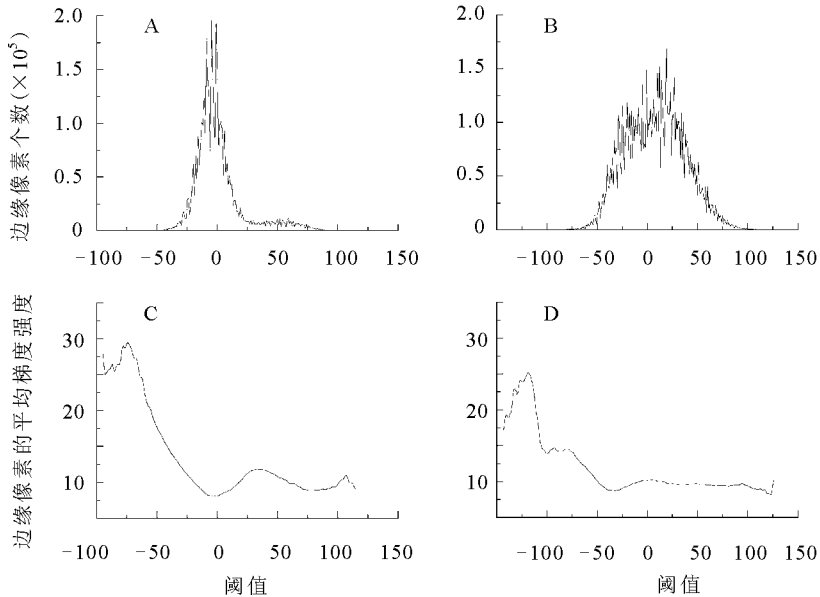


图 3 样片 a、b 的边缘像素个数及平均梯度强度分布数

注: A、B 为图 1 样片 a、b 的边缘像素个数图; C、D 为样片 a、b 的边缘像素平均梯度强度分布图; 样片 A 的最佳阈值等于 37, 样片 B 的最佳阈值等于 -3。

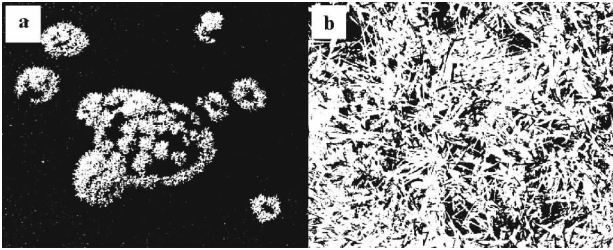


图4 样片a、b的植被提取结果

注:样片a、b的植被覆盖度分别为12.50%和67.01%。

2 结果与分析

从回归斜率来看,3种方法的回归斜率均小于1,存在低估的现象(图5)。这主要是因为人工目视解译时勾画的是植物的外轮廓,包括占植物绝大多数的绿色叶片和少量的非绿色的花朵、叶片互相遮挡的阴影等;而其他3种方法只是提取植物的绿色

部分。低估现象无法避免,低估程度存在差异。色度饱和度法的低估现象比较严重,回归斜率为0.6446;半自动阈值法和最大似然法比较接近,回归斜率分别为0.8617和0.8633。

从回归截距来看,最大似然法最差,偏离0最远,达到4.65198;半自动阈值法和色度饱和度法偏离0较小,分别-0.7655和-0.2530。最大似然法存在一个较严重的问题,即低植被覆盖度高估现象。这主要是因为植被覆盖度低的时候,植物往往较小,往往找不到很纯的植物叶片(混合像元问题)用于分类前的训练,训练的像素中带有一定的非植被信息,使得部分颜色接近植物的背景也被归为植被。

从回归系数来看,最大似然法与准确结果最接近半自动阈值法次之;色度饱和度法最差。

综上所述,最大似然法因为存在人工干预选择训练像素这个过程,包含了一定的主观纠错功能在里面,故而计算结果与真实情况很接近,但是存在低植被覆盖度高估现象。色度饱和度法虽然也有人工干预选择参数这一步骤,但是只适用于植被简单、背景单一的情形^[16],植被和背景の色度饱和度重叠程度较高的复杂情形下,效果较差。这两种方法的主观干预部分都需要较为复杂且费时的操作。相比之下,本研究的半自动阈值法无论是从结果精度还是人工干预的简便程度上,都更值得推荐。

当然,本研究的半自动阈值法同样存在一定的缺陷,即无法准确提取绿色特征不明显的植物,比如叶片呈灰绿色的驼绒藜(*Ceratoides latens*)以及各种蒿类植物。根据经验,这类植物的大部分叶片会被错分成背景(图6),造成植被覆盖度低估。

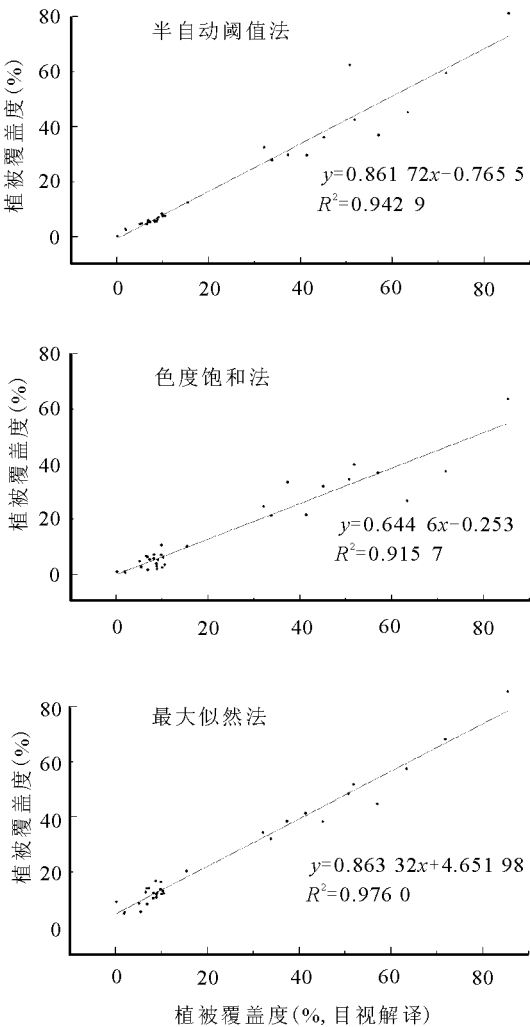


图5 植被覆盖度提取方法效果比较

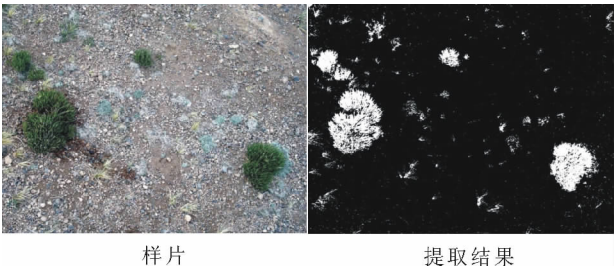


图6 存在蓝绿色植物的草地样片及其植被提取结果

3 结论

本研究提出了一种利用过绿特征植被指数和半自动阈值设定算法的快速从数码照片中计算草地植被覆盖度的方法,并将该方法和最大似然法监督分

类方法和色度饱和度法进行了比较。本研究提出的方法人工干预极少,计算过程客观,结果比较准确可靠;而且对拍摄器材和天气条件等外在因素要求不高,适用性强。本研究采用过绿特征作为植被指数,故而该方法无法准确提取出绿色特征不明显的植物,比如灰绿色植物。如能提出更好的可见光波段的植被指数,或者数码相机中增加近红外波段的感光元件,可能能够得到更准确结果。

参考文献

- [1] 秦伟,朱清科,张学霞,等. 植被覆盖度及其测算方法研究进展[J]. 西北农林科技大学学报(自然科学版), 2006,34(9):163-170.
- [2] Gutman G, Ignatov A. The derivation of the green vegetation fraction from NOAA/AVHRR data for use numerical weather prediction models [J]. International Journal of Remote Sensing, 1998, 19(8):1533-1543.
- [3] 温庆可,张增祥,刘斌,等. 草地覆盖度测算方法研究进展[J]. 草业科学, 2009, 26(12):30-36.
- [4] 章文波,符素华,刘宝元. 目估法测量植被覆盖度的精度分析[J]. 北京师范大学学报(自然科学版), 2001, 37(3):402-408.
- [5] Chapman S B. 植物生态学的方法[M]. 阳含熙,译. 北京:科学出版社,1980.
- [6] White M A, Asner G P, Nemani R R, *et al.* Measuring fractional cover and leaf area index in arid ecosystem; Digital camera, radiation transmittance, and laser altimetry methods [J]. Remote Sensing of Environment, 2000, 74(1):45-57.
- [7] 池宏康,周广胜,许振柱,等. 草地植被盖度的近距离遥感测定[J]. 草业学报, 2007, 16(2):105-110.
- [8] Zhou Q, Robson M. Automated rangeland vegetation cover and density estimation using ground digital images and a spectral-contextual classifier [J]. International Journal of Remote Sensing, 2001, 22(17):3457-3470.
- [9] 宋雪峰,董永平,单丽燕,等. 用数码相机测定草地盖度的研究[J]. 内蒙古草业, 2004, 16(4):1-6.
- [10] 张清平,章山山,陈露,等. 应用 WinCAM 软件判识分析草坪盖度[J]. 草业科学, 2010, 27(7):13-17.
- [11] 张学霞,朱清科,吴根梅,等. 数码照相法估算植被覆盖度[J]. 北京林业大学学报, 2008, 30(1):164-169.
- [12] 钱金波. 祁连山狼毒盖度地面测量与卫星遥感估算 [D]. 兰州:兰州大学, 2009.
- [13] 张云霞,李晓兵,张云飞. 基于数码相机、ASTER 和 MODIS 影像综合测量植被盖度[J]. 植物生态学报, 2007, 31(5):842-849.
- [14] 顾祝军,曾志远. 遥感植被盖度研究[J]. 水土保持研究, 2005, 12(2):18-21.
- [15] 龙满生,何东健. 玉米苗期杂草的计算机识别技术研究[J]. 农业工程学报, 2007, 23(7):139-144.
- [16] Richardson M D, Karcher D E, Purcell L C. Quantifying turfgrass cover using digital image analysis[J]. Crop Science, 2001, 41:1884-1888.

Quickly determining grassland cover using the digital image

HU Jian-bo¹, ZHANG Lu², HUANG Wei¹, WU Shi-hong¹, LIU Chang-bing¹

(1. Laboratory of Environmental Protection in Water Transport Engineering,

Tianjin Research Institute for Water Transport Engineering, Ministry of Communications, Tianjin 300456, China;

2. School of Surveying and Land Information Engineering, Henan Polytechnic University, Henan Jiaozuo 454000, China)

Abstract: The improving efficiency and accuracy of vegetation cover is very important because the cover is a necessary parameter to estimate the health of grassland ecosystem. This study proposed a quickly calculating vegetation cover method by Excess Green Index and semi-automatic threshold (ST). Efficiency of the algorithm was evaluated by comparing with Maximum-likelihood supervised classification (MC) and Hue-Saturation method (HS). 32 digital images were used to test the efficiency of three methods. Vegetation cover estimated by each method was regressed with accurate cover obtained by visual interpretation via slope, intercept, and regression coefficient. The three methods did not consider non-green component of vegetation, resulting in slightly underrating problem. In conclusion, ST not only achieved similar accuracy with MC but also avoided overrating problem of MC when cover was low. ST showed little human intervention, and achieved accurate cover; however, it failed to extract grey-green vegetation.

Key words: digital image; vegetation cover; Excess Green; semi-automatic thresholding