

基于成像光谱技术的作物杂草识别研究

刘波^{1,2}, 方俊永¹, 刘学¹, 张立福¹, 张兵³, 童庆禧¹

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 中国科学院研究生院, 北京 100049

3 中国科学院对地观测与数字地球科学中心, 北京 100080

摘要 杂草识别是变量喷雾和物理方法精确除草的前提。利用自主设计的地面成像光谱系统在自然环境下获取了胡萝卜幼苗以及马齿苋、牛筋草和地锦等杂草在 380~760 nm 波长区间的高光谱数据, 通过对数据归一化消除光照条件的影响之后, 运用逐步法进行波段选择, 采用 Fisher 线性判别方法对杂草与胡萝卜幼苗进行了识别。结果表明, 当把每种杂草都作为一类加以精细区分时, 运用选择的 8 个波段建立模型对杂草和胡萝卜幼苗的识别率达 85% 左右; 当把杂草整体作为一类与胡萝卜幼苗进行区分时, 运用选择的 7 个波段识别率高于 91%。同时为了设计低成本的杂草识别系统, 通过穷举法选择最优的 2 和 3 波段组合, 其中最 3 波段组合对杂草胡萝卜幼苗的识别能力与逐步法选择的 5 个波段相当, 整体识别率达 89%。此外发现, 红边波段对杂草有着显著的识别能力。

关键词 成像光谱; 光谱分析; 杂草识别; 作物

中图分类号: S543 文献标识码: A DOI: 10.3964/j.issn.1000-0593(2010)07-1830-04

引言

随着人们对环境污染、食品安全问题的日益重视, 高效、安全、经济的精准农业成为现代化农业的发展方向。传统的大面积除草剂使用不仅会带来环境污染, 而且会造成极大浪费; 而基于变量喷雾和物理手段的精确除草无论在环境效应、安全生产还是经济节约上, 都有着明显的优势, 因而成为当前的重要研究方向。精确除草的前提和关键在于杂草和作物的自动快速识别。国内外学者尝试了利用不同手段进行杂草与作物识别, 诸如基于机器视觉方法和光谱分析方法等, 其中基于光谱信息的技术方法因简单、快速而受到了广泛的关注^[1-7]。

在人工光照条件下, Borregaard 等利用成像光谱仪进行了甜菜、马铃薯和 3 种杂草的识别, 经过波段选择利用 694 和 970 nm 波段对甜菜和杂草的识别正确率达到了 89%, 686 和 856 nm 波段对马铃薯和杂草的识别正确率达到 94%; 当利用选择的 6 个波段对杂草和作物以及杂草之间进行区分, 整体识别率下降为 60%~71%^[8]。Feyaerts 等利用光谱分辨率为 35 nm 的多光谱系统, 通过人工移植杂草于机器视场范围内, 获取不同波段图像对甜菜和 5 种杂草进行区分, 通过

最大化可分性函数选出 441, 446, 459, 883, 924 和 988 nm 作为特征波段, 整体识别率达 85%^[9]。Piron 等利用自行研制的多光谱装置, 在人工光照条件下进行了室外试验, 通过穷举法获取的最优 3 波段(450, 550, 700 nm)对胡萝卜幼苗和杂草的整体识别率达到了 72%^[10]。毛文华等利用傅里叶变换光谱仪测量分析了小麦、小藜和芥菜等几种杂草在 700~1100 nm 波长范围内的反射率, 选择 686, 708, 722, 795, 929, 956 和 1122 nm 7 个波段建立判别模型识别小麦和杂草, 识别率达到了 97% 以上。陈树人等利用 ASD 光谱仪在室内测量了棉花、刺儿菜、水稻、稗草的 350~2500 nm 的反射率, 通过逐步选择的 3 个特征波长 385, 415, 435 nm 区分双子叶植物棉花和杂草刺儿菜, 识别率达 100%; 5 个特征波长 375, 465, 585, 705 和 1035 nm 区分单子叶植物水稻和稗草, 识别精度为 100%^[11]。

当前大多数研究都集中在室内或者人工光照条件下, 无论样本准备还是测量条件都比较理想, 影响因素较少; 而实际田间环境复杂, 多种因素都会影响植物光谱特性; 此外, 非成像的光谱分析由于测量方式的限制很难用于田间作物杂草幼苗时期, 而这个时期恰恰是除草的最佳时机。因此在田间条件下进行杂草作物识别有待于进一步深入研究。本文利用自行研发的地面成像光谱仪, 在实际田间环境下获取了幼

收稿日期: 2009-09-02, 修订日期: 2009-12-08

基金项目: 中国科学院重大科研装备研制项目, 遥感科学国家重点实验室自由探索项目和国家科技支撑计划项目(2007BAH15B01)资助

作者简介: 刘波, 1984年生, 中国科学院遥感应用研究所博士研究生, e-mail: boxueyu_liu@hotmail.com

苗期胡萝卜和几种杂草高光谱数据,进行了胡萝卜与杂草以及杂草之间的识别研究。

1 实验部分

1.1 仪器设备

实验仪器为自行设计的地面成像光谱仪,由计算机系统、光机系统和电子学系统构成。其中光机系统主要包含 4 部分,扫描镜、光学镜头、分光器件(ImSpector V9, Spectral Imaging, Ltd, Finland)以及 CCD 相机。该系统在 400~850 nm 之间分为 344 个波段,光谱采样间隔约为 1.4 nm;通过扫描镜转动来实现空间维数据获取,空间分辨率随作业平台高度而变化,最优高于 2 mm, DN 量化位数为 12 位,详细介绍见文献[12]。

1.2 数据获取

2009 年 7 月 30 日,在北京大兴区务各庄农田中采集了高光谱数据。作物为胡萝卜,种植时间为 7 月 14 日,正值除草的理想时段,除草时期的选择对胡萝卜产量有着显著的影响^[10]。田间杂草主要为马齿苋(*Portulaca leraacea*)、牛筋草(*Eleusine Indica*)和地锦(*Euphorbia humifusa*),都为北方地区常见的杂草。数据采集时天气多云,光照条件随时间变化较大,采集时间为 9:30-13:00,总共获取了 41 幅图像。

2 数据处理

2.1 数据预处理

对于采集的数据,波长小于 410 nm 的数据由于信噪比较低而舍弃;为了减少计算数据量和去除噪声,对于剩下的 320 个波段每隔相邻 4 个波段求取平均,得到平均后的 80 个波段。由于测量时光照条件变化,不便实时测量标准板以获取反射率。为了减小和消除光照条件变化的影响,对数据做了归一化处理,即对每个像元每个波段的 DN 值除以其 80 个波段的平均 DN 值,图 1 为胡萝卜光谱 DN 曲线归一化前后的对比。

先通过计算 NDVI(normalized difference vegetation index)设定阈值,把土壤背景分离出去;然后对于每种杂草和胡萝卜,分别随机采取训练样本和验证样本各 3 000 个像元,为了保证数据的独立性,训练样本和验证样本来自不同的 20 幅图像,这样得到总体训练样本和验证样本各 12 000 个,其中训练样本用于波段选择和建模,验证样本用于检验模型。

2.2 波段选择与分类方法

在验证样本集上,采用 Wilks' lambda 逐步方法进行波段选择,在 SPSS 软件中完成。需要指出的是逐步选择方法得到的波段组合未必是最优的波段组合。选出波段之后,本文采用 Fisher LDA 法(linear discriminant analysis)进行杂草作物区分。Fisher LDA 的基本思想是根据类间距离最大、类内距离最小的原则确定线性判别函数,根据建立的线性判别方程判定未知样本的类别,LDA 在 Matlab 6.5 中编程实现。

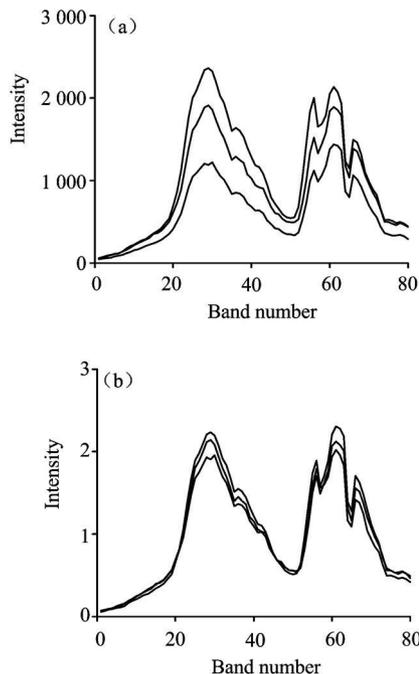


Fig 1 Normalization of hyperspectral data

(a): Raw data; (b): Normalized data

3 结果与分析

对于物理除草,只需把杂草识别出来即可,无需进一步区分不同种类的杂草;对于变量喷雾有选择性地使用不同除草剂时,则希望不仅进行杂草与作物的区分,而且需要识别出每种杂草。因此,本文针对这两种情况分别进行了研究,即物理除草对应于杂草整体和作物的两类区分,变量喷雾对应于杂草之间以及杂草与作物的多类区分问题。

3.1 杂草之间以及杂草与作物的多类区分

图 2 显示了随着入选波段数目的增加,3 种杂草和胡萝卜识别率以及整体识别率整体趋势上快速增加,当波段数目达到 8 个时,识别率均位于 81%~88% 之间;之后随着波段数目的增加识别率增加幅度变得十分微小,精度趋于稳定。因此利用所选的前 8 个波段进行胡萝卜与 3 种杂草的精细识别可以达到比较理想的效果,8 个波段对应的中心波长依次为 585, 714, 608, 732, 434, 827, 696, 661 nm。其中 714, 732 和 696 nm 3 个波段位于植物反射光谱的红边区域,主要反映了不同植物叶片结构的差异,为绿色植物的重要的诊断性特征,这与植被遥感的理论完全相符^[3]。

图中显示马齿苋和牛筋草较难识别,而胡萝卜和地锦较易识别,当只有 4 个波段时识别率可达 80%。此外当波段数目小于 5 时,部分类别识别精度并非完全随着波段数目增加而提高,而是出现波动,这是由于波段选择是以所有类别的整体可分性为依据的,因此并不能保证每一类的识别率都随着波段数目增加而提高;图中整体识别率则随着波段数目增加而逐渐提高很好地说明了这个问题。

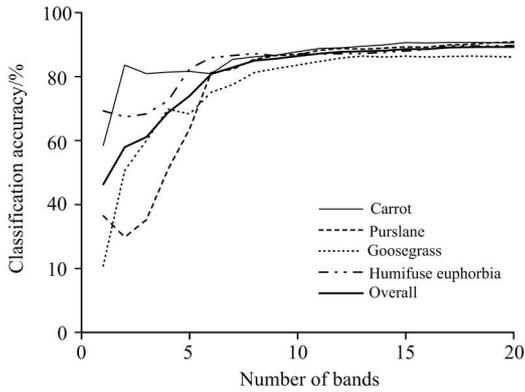


Fig 2 Classification results of weed and crop weed using selected bands

3.2 杂草与作物两类区分

图 3 为杂草与胡萝卜的识别率以及整体识别率随着入选波段数目增加的变化情况。同样地, 无论是杂草、胡萝卜还是整体识别率都随着波段数目增加而整体提高; 与图 2 不同的是, 这里没有出现明显的拐点, 识别率精度基本上稳步提高。当只有 1 个波段时, 各类识别率都超过了 80%, 该波段为 714 nm 位于红边区域, 是所有波段中区分胡萝卜与杂草能力最强的波段, 再次说明红边是反映植物状况的重要特征光谱段。可以有效地区分不同种类植物, 这点与陈树人等的研究结果相一致^[12]。当入选波段增加到 2 个和 3 个的时候, 整体精度有所提升, 但幅度比较小。当入选波段增加到 7 个 (714, 579, 608, 696, 597, 661, 423 nm) 时, 各种识别率都达到了 91% 以上。值得注意的是, 其中 714, 608, 696, 661 nm 4 个波段同时也出现在 3.1 节 4 类精细识别的 8 个波段组合中, 显示了这 4 个波段是进行杂草识别的重要波段; 且 2 种情况下出现的先后次序相同, 则揭示了他们对分类贡献大小的排列顺序。两种情况下选择的波段并非完全相同, 则说明了不同波段组合适应于不同的分类问题。

与前面类似, 随着波段数增加, 杂草和胡萝卜的识别精度有所波动, 而整体精度一直提高。

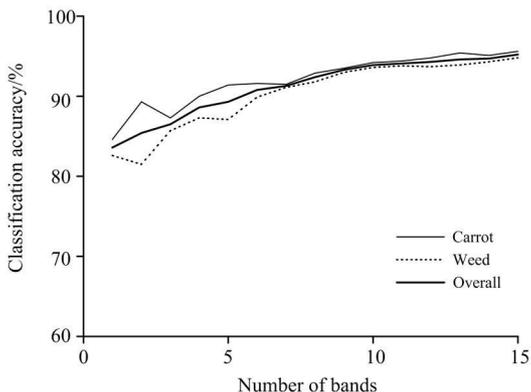


Fig 3 Classification results of crop weed using different selected bands

3.3 最佳 2 和 3 波段组合

考虑到高光谱成像仪价格比较昂贵, 实际应用中希望能够利用较少的波段达到较为满意的识别结果, 以便设计开发

低成本的仪器, 比如可以利用滤光片技术容易实现的 2 波段或者 3 波段的系统。但是逐步法选出的波段组合未必最优, 因此本研究采用穷举法选择了最优的 2 波段组和 3 波段组合, 可分性准则采用散布矩阵 (scatter matrix)^[14]。选择的波段组合只有 2 和 3 个的情况, 因此计算量在可接受范围之内。对于每种杂草都要识别的 4 类精细区分且波段数目小于 5 时, 部分杂草识别精度不高, 难以达到实用水平, 因此本文没有列出这种情况下的最优波段选择结果, 以下结果和分析只针对杂草整体和胡萝卜的两类识别情况。

表 1 为穷举法得到的最优 2 波段、最优 3 波段与 2 2 中用逐步选择法得到的 2 波段和 3 波段组合的识别精度对比情况。为了说明光谱法的有效性, 同时也给出了通常数码相机 RGB 对应的 3 个波段的识别精度。

Table 1 Classification results of crop weed and the corresponding selected bands using different band selection methods

	波段组合 /nm	胡萝卜 识别率/%	杂草 识别率/%	总体 识别率/%
穷举法	568, 714	87.9	83.1	85.5
逐步法	579, 714	89.3	81.5	85.4
穷举法	597, 614, 708	89.6	88.5	89.1
逐步法	579, 608, 714	87.3	85.7	86.5
RGB	450, 550, 638	79.1	75.6	77.3

由表可以看出, 与逐步法所得 2 波段相比, 最优 2 波段的整体识别精度并未提高; 而最优 3 波段组合无论是杂草、胡萝卜还是整体识别率都高于相应的逐步法 3 波段结果, 其中整体识别率达 89.1%, 与逐步法的 5 波段的识别精度相当。因此在考虑低成本的仪器设计时, 可以考虑 568 和 714 nm 与 597, 614, 708 nm 的波段设置。同样我们可以发现, 穷举法得到的最优波段组合中都至少含有一个波段位于红边区域, 证实了红边波段提取植物信息的重要性。有趣的是, 无论穷举法还是逐步法得到的结果, 都远远优于常规 RGB 波段的识别率 (整体为 77.3%)。实际上— 优选的单波段 (714 nm) 的识别精度 (83.6%) 就已远远高于 RGB 3 波段的识别精度, 这充分说明了不同波段的杂草识别能力相差巨大, 同时也说明了多 (高) 谱分析方法的优势。

4 结论

在自然环境下, 利用自行设计研制的地面成像光谱系统进行了杂草与作物的识别。

(1) 利用 585, 714, 608, 732, 434, 827, 696, 661 nm 波段可以有效地实现对杂草之间以及杂草与作物的精细识别, 整体识别精度优于 85%。利用 714, 579, 608, 696, 597, 661, 423 nm 波段对杂草和作物进行区分, 精度高达 91% 以上。

(2) 考虑到实际应用中的成本问题, 利用穷举法得到的最优 3 波段组合 (597, 614, 708 nm), 其整体识别精度高达 89.1%, 高于逐步法得到 3 波段结果, 相当于逐步 5 个波段的精度。这为以后设计研发廉价简易的杂草识别仪器提供了

参考。

(3) 位于红边区域的 710 nm 左右的波段对杂草与作物有着显著的区分能力, 因此在逐步法中通常作为第一个或者第二个入选波段, 并且是最优波段组合中不可缺少的波段。

(4) 随着入选波段的增加, 识别精度整体呈上升趋势, 个别类别出现波动。对于不同种类杂草和作物的精细区分, 精度曲线的拐点出现在 8 个波段, 之后精度增加微弱, 而杂草整体与作物的识别精度曲线则没有出现明显的拐点, 稳步增加。

(5) 无论穷举法还是逐步法得到的 2~3 波段组合, 其识别能力都远远高于常规数码相机的 RGB 3 波段; 即使是选出的单波段(714 nm), 其分类精度为 83.6%, 远远高于 RGB 的 77.3%。

本研究虽然在田间实际环境下初步进行了杂草识别研究, 取得了较好的结果, 但由于实验条件限制, 依然未能全面考虑各种因素。应在不同天气条件下、作物幼苗期的不同阶段以及试验中增加杂草和作物的种类, 进一步验证和改进模型。这也是我们的下一步研究工作。

参 考 文 献

- [1] Wang N, Zhang N, Dowell F E Y S, et al. Transactions of the ASAE, 2001, 44(2): 409.
- [2] Alchanatis V, Ridell L, Hetzroni A, et al. Computer and Electronics in Agriculture, 2005, 47: 243.
- [3] Goel P K, Prasher S O, Patel R M, et al. Transactions of the ASAE, 2002, 45(2): 443.
- [4] Onyango C M, Marchant J A. Computer and Electronics in Agriculture, 2003, 39: 141.
- [5] Tillett N D, Hague T, Miles S J. Computer and Electronics in Agriculture, 2001, 32: 229.
- [6] Perez A, Lopez F, Benloch J V, et al. Computer and Electronics in Agriculture, 2000, 25: 197.
- [7] Virndts E, De Baerdemaeker J, Ramon H. Precision Agriculture, 2002, 3: 63.
- [8] Borregaard T, Nielsen H, Norgaard L, et al. Journal of Agricultural Engineering Research, 2000, 75(4): 389.
- [9] Fevaerts F, Van Gool L. Pattern Recognition Letters, 2001, 22: 667.
- [10] Piron A, Leemans V, Kleynen O, et al. Computer and Electronics in Agriculture, 2008, 62: 141.
- [11] CHEN Shuren, LI Yixin, MAO Harping, et al (陈树人, 栗移新, 毛罕平, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2009, 29(2): 463.
- [12] FANG Jurong, QI Hongxing, LIU Xue, et al(方俊永, 齐洪兴, 刘学, 等). Proceedings of the 7th Technology and Application of Imaging Spectroscopy(第七届成像光谱技术与应用研讨会文集), 2009. 148.
- [13] TONG Qingxi, ZHANG Bing, ZHENG Larfen(童庆禧, 张兵, 郑兰芬). Hyperspectral Remote Sensing Principle, Technology and Application(高光谱遥感原理、技术与应用). Beijing: Higher Education Press(北京: 高等教育出版社), 2006. 22.
- [14] Theodoridis S, Koutroumbas K. Pattern Recognition. Beijing: China Machine Press(北京: 中国机械出版社), 2006.

Research on Crop Weed Discrimination Using a Field Imaging Spectrometer

LIU Bo^{1,2}, FANG Jurong¹, LIU Xue¹, ZHANG Lifu¹, ZHANG Bing³, TONG Qingxi¹

1. State Key Lab of Remote Sensing Science, Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing 100101, China

2. Graduate University of Chinese Academy of Sciences, Beijing 100049, China

3. Center for Earth Observation and Digital Earth, Chinese Academy of Sciences, Beijing 100080, China

Abstract Discrimination of weeds from crop is the first and important step for variable herbicides application and precise physical weed control. Using a new field imaging spectrometer developed by our group, hyperspectral images in the wavelength range 380-870 nm were taken in the wild for the investigation of crop weed discrimination. After normalizing the data to reduce or eliminate the influence of varying illuminance, stepwise forward variable selection was employed to select the proper band sets and fisher linear discriminant analysis (LDA) was performed to discriminate crop and weeds. For the case of considering each species as a different class, classification accuracy reached 85% with eight selected bands while for the case of considering overall weed species as a class, classification accuracy was higher than 91% with seven selected bands. In order to develop a low-cost device and system in future, all combinations of two and three bands were evaluated to find the best combinations. The result showed that the best three bands can achieve a performance of 89% comparable to the performance achieved by five bands selected using stepwise selection. The authors also found that "red edge" could afford abundant information in the discrimination of weed and crop.

Keywords Imaging spectrometer; Spectral analysis; Weed discrimination; Crop

(Received Sep. 2, 2009; accepted Dec. 8, 2009)