

基于光谱分析技术的作物中杂草识别研究

陈树人, 栗移新, 毛罕平, 沈宝国, 张玉珠, 陈斌

江苏大学现代农业装备与技术省部共建教育部重点实验室, 江苏省重点实验室, 江苏 镇江 212013

摘要 除草剂的精确喷施、物理方法精确除草皆依赖于杂草的自动识别。光合色素和结构差异导致作物、杂草的光谱反射率不同, 因此不同植物可以利用光谱特性来区分。利用 ASD 光谱仪在室内分别测量了棉花、刺儿菜、水稻、稗草等四种植物在 350~2 500 nm 波段范围内的光谱反射率。运用 SAS 统计软件的 STEP-DISC 过程筛选能够区分作物和杂草的波长; 判别模型中加入筛选所得特征波长, 利用 Discrim 过程进行判别分析。实验结果表明, 利用 3 个特征波长 385, 415 和 435 nm 有效地从双子叶植物棉花中识别出双子叶杂草刺儿菜, 其识别率为 100%, 波长 415 和 435 nm 的组合对识别模型的贡献最大; 利用 5 个特征波长 375, 465, 585, 705 和 1 035 nm 可有效地从单子叶植物水稻中识别出单子叶杂草稗草, 其识别率也为 100%, 黄色到橙色的过渡波长 585 nm 和“红边”内的波长 705 nm 的组合对识别模型的贡献最大。

关键词 光谱分析; 双子叶植物; 单子叶植物; 杂草识别

中图分类号: S341.6 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2009)02-0463-04

引言

粗放式的大面积喷洒除草剂虽然达到了除草的效果, 但带来了诸多问题, 比如农业环境的污染、杂草对除草剂产生抗性、杂草群落变迁等。为减缓化学除草带来的环境污染、引发粮食和农产品安全问题, 变量喷雾、物理方法精确除草成为现阶段的研究热点。精确除草的前提条件是杂草的自动识别。与利用图像处理的杂草识别方法相比, 基于光谱分析技术的杂草识别方法处理过程相对简单, 并且在实时性方面具有明显的优势。

Brown 等运用分光辐射仪在自然光照条件下测量了免耕棉田中的兰草、狗尾草、小藜等七种杂草在 400~900 nm 范围内的冠层反射率, 通过分析得出 440, 530, 650 和 730 nm 四个特征波长^[1]。Goel 等运用机载成像分光计以 1 m 的空间分辨率测量了有三种不同组合杂草的玉米和土豆作物在 475.12~970.01 nm 范围内 24 个波长的数据, 三类不同杂草组合分别是以苘麻、反枝苋和藜为主的阔叶类, 以金色狗尾草为主包括稗草在内的禾本科杂草和自然发生的优势杂草混合。分析结果表明, 在红光 675.98 和 685.17 nm 处及近红外光 743.93~830.43 nm 波段, 能够用于区分玉米田中的杂草; 而对于土豆作物, 仅 811.40 nm 处有效^[2]。Vrindts 等运

用光学频谱分析仪在实验室内测量了甜菜、玉米以及七种杂草在 400~2 000 nm 波长范围内的反射率, 以 555, 675, 815, 1 265, 1 455 和 1 665 nm 处的反射率作为分母, 利用在很少的波段处的反射率之比在实验室内可以区分作物与杂草, 识别率高达 97%; 在田间自然光照条件下, 运用线性摄谱仪测量了作物和杂草在 480~820 nm 波长范围内的冠层反射率, 判别模型利用有限的波段处反射率的比值。当周围光照条件一致时, 田间的作物和杂草正确识别率高达 90%^[3]。Slaughter 等研究了用可见光和近红外光反射光谱(400~2 500 nm)区分西红柿和同属茄科的杂草龙葵和毛茄, 利用窄波段超光谱模式和标准判别分析, 发现在 2 120~2 320 nm 波长范围内的光谱数据能达到最好的区分率(100%)^[4]。毛文华等利用傅里叶变换红外光谱仪测量并分析了小麦、小藜和芥菜等几种杂草在 700~1 100 nm 波长范围内的反射率, 运用选定的特征波长建立判别模型识别小麦和杂草的正确识别率达到了 97%^[5]。

利用光谱特性对单子叶作物和单子叶杂草、双子叶作物和双子叶杂草的分类研究很少。结合江苏省本地实际情况^[6, 7], 本文以棉田中的刺儿菜和稻田中的稗草为对象, 对作物与杂草形状和颜色相似的情况下, 利用光谱分析技术研究了杂草的识别方法。

收稿日期: 2007-10-06, 修订日期: 2008-01-06

基金项目: 江苏省博士后科研计划项目(0601014B), 国家自然科学基金项目(60575020)和江苏大学现代农业装备与技术建教育部重点实验室培育点开放基金项目(NZ200708)资助

作者简介: 陈树人, 1965年生, 江苏大学现代农业装备与技术建教育部重点实验室教授 e-mail: srchen@ujj.edu.cn

1 材料与方法

1.1 实验材料

实验所采用的水稻(*Oryza sativa* Linn.)、稗草(*Echinochloa crusgalli*)、棉花(*Gossypium hirsutum*)和刺儿菜(*Cephalanoplos segetum*)均来自于江苏大学农业工程研究院温室大棚;水稻和稗草为 2007 年 7 月 1 日于培养钵中培育,棉花和刺儿菜样本来自于农业工程研究院实验田。实验仪器是美国 ASD 公司的 Analytical Spectral Device Full Range FieldSpec Pro(全光谱便携式光谱分析仪),在 350~1 000 nm 范围采样间隔是 1.4 nm,1 000~2 500 nm 范围采样间隔是 2 nm。光源采用与光谱仪配套的卤素灯。

1.2 实验方法

2007 年 9 月 7 日,在室内采集了棉花、刺儿菜、水稻、稗草四种植物的光谱数据。光谱仪经白板校准后进行测试,光纤探头置于叶片表面上方 10~20 mm,探头视场角为 8°。为了不影响植物的反射率,实验采用黑色背景。光谱仪设置为一次测量得到 5 组数据,每种样本测量 40 次 200 组数据。

2 数据处理

对每次测量的 5 组数据取平均,即每种样本的 200 组数据压缩为 40 组。以其中 20 组为建模数据集,总体 40 组数据为训练数据集。测量仪器导出的数据包括 350 到 2 500 nm 间隔为 1 nm 的各波长处的反射率。为了提高计算机工作效率,按 675~815 nm 波段(675 nm 是红光最小值,815 nm 是在 NIR 高原区最大值)每隔 5 nm 选取一个数据点,在 1 950~2 500 nm 波段内每 20 nm 选取一个数据点,其余波段每 10 nm 选取一个数据点,这样把 2 151 个数据点压缩为 202 个数据点。

以 202 个相应波长处的反射率作为每种植物的变量,利用 SAS 的 STEPDISC 过程筛选变量。其筛选根据是:(1)将已选择的变量作为协变量,所考查变量作为因变量进行协方差分析,分析结果中 F 检验的显著性检验结果;(2)固定已选择变量的效应后,从分类变量预测所考查变量的平方偏相关系数。采用 STEPDISC 过程对变量进行筛选,可以避免无关变量对判别分析结果的影响或指标变量过多导致判别结果的不稳定性。用筛选所得波长作为判别模型,利用 SAS 统计软件的 DISCRIM 过程对相应的数据集进行判别分析。

3 结果与分析

实验所得光谱曲线如图 1。在 780~1 350 nm 的近红外波段内,双子叶植物棉花和刺儿菜的反射率要高于单子叶植物水稻和稗草的反射率。这是由于叶片内部结构的差异造成的。双子叶植物的叶片多是双面叶,具有腹面(近轴面)和背面(远轴面)。腹面的栅栏组织具有相当的细胞空隙,背面的海绵组织则形成了较大的细胞间隙。单子叶植物水稻和稗草是等面叶,叶肉组织由均匀的薄壁细胞构成,细胞间隙比较

小。而细胞间隙是影响叶片近红外区反射率的重要因素。早期的研究证实,叶片在近红外波段处的反射率受暴露的叶肉细胞表面积(Ames)与单位叶面积表示的细胞间隙(IAS)的比值影响^[8-10]。Gausman 实验发现,夹竹桃在近红外区的反射率比玉米高,这与夹竹桃叶肉组织中细胞间隙比玉米叶肉组织中多相关^[11]。图中水稻和稗草的光谱曲线比较接近,棉花和刺儿菜的比较接近。因此单子叶植物与双子叶的植物间的区分容易,而单子叶作物和单子叶杂草、双子叶作物和双子叶杂草的区分则相对复杂。

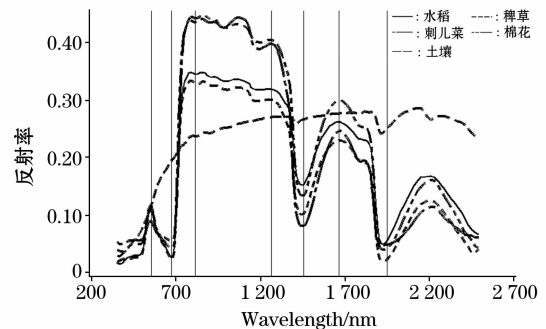


Fig. 1 Spectral distribution characteristic of cotton, spine-greens, rice, barnyard-grass, and soil

3.1 双子叶刺儿菜和棉花数据集的区分

筛选所得的波长为 385, 415, 435, 1395, 1 455 和 1 465 nm。利用以上波长处的反射率对训练数据集的判别得到了很好的效果,总的错误率为 0%。水分吸收波段 1 455~1 950 nm 受环境影响较大,排除此波段内的波长点,剩下的特征波长为 385, 415 和 435 nm。

Table 1 Classification results for spine-greens and cotton using different numbers of the selected wavelength

特征波长 /nm	刺儿菜 错误率/%	棉花 错误率/%	总体 错误率/%
385, 415, 435	0	0	0
415, 435	0	0	0
385, 415	10	5	7.5
385, 435	15	15	15
415	20	10	15
435	30	10	20
385	30	10	20

由表 1 可见,识别模型仅用一个特征波长点时,特征波长 415 nm 比 385 和 435 nm 的效果好,这与利用 STEPDISC 过程筛选变量时 F 检验的显著性检验结果一致。利用单个特征波长时错误率太高,因此判别模型宜联合使用多个特征波长。判别模型中入选全部 3 个特征波长时,总体错误率为 0%。由表也可以看出特征波长 415 和 435 nm 对识别模型的贡献较大,缺少特征波长点 415 和 435 nm 时,总体错误率分别为 15% 和 7.5%;相反,识别模型中仅有特征波长点 415 和 435 nm 时可以使错误率减少到 0%。

3.2 单子叶稗草和棉花数据集的区分

筛选所得的波长为 365, 375, 385, 455, 465, 495, 585, 605, 615, 675, 705, 1 035, 1 045, 1 305, 1 045, 1 305, 1 415, 1 435, 1 445, 1 555, 1 905, 2 130, 2 150, 2 170, 2 190, 2 330, 2 450 和 2 470 nm。利用以上这些波长处的反射率对训练数据集的判别得到了很好的效果, 总的错分率为 0%。而考虑到野外环境对水分吸收波段的影响很大, 所以仅选择利用 350~1 350 nm 波段的特征波长对训练集进行判别。结合 STEPDISC 过程筛选变量时的 F 检验的显著性检验结果, 选择了 375, 465, 585, 605, 705 和 1 035 nm 波长。表 2 为利用这些波长的反射率判别稗草和水稻的结果。

Table 2 Classification results for barnyard-grass and rice using different numbers of the selected wavelength

特征波长点/nm	稗草错 识率/%	水稻错 识率/%	总体错 识率/%
375, 465, 585, 705, 1 035	0	0	0
375, 465, 585, 705	0	0	0
375, 465, 585, 1 035	10	5	7.5
375, 465, 705, 1 035	5	10	7.5
375, 585, 705, 1 035	0	0	0
465, 585, 705, 1 035	0	0	0
585, 705	0	0	0
375	40	10	25
465	45	45	45
585	30	50	40
705	25	35	30
1 035	30	35	32.5

由表 2 可见, 当判别模型仅用一个特征波长时, 识别效果都比较差, 特征波长 375 nm 效果略好, 这与利用 STEPDISC 过程筛选变量时 F 检验的显著性检验结果基本一致, 但总体错识率太高。判别模型选入全部 5 个特征波长时, 总体错识率为 0%。由表 2 也可以看出各个特征波长对识别模型的贡献。当判别模型中仅缺少特征波长 705 nm 时, 稗草的错识率高达 10%, 水稻的错识率为 5%; 当仅缺少特征波长 585 nm 时, 稗草的错识率高达 5%, 水稻的错识率为 10%; 相反, 用特征波长 585 和 705 nm 时, 总体错识率可以减少为 0%。

4 结 论

在实验室条件下利用光谱分析技术能对颜色和形状非常相近的作物和杂草进行识别。

(1) 利用可见光波段的 3 85, 415 和 435 nm 3 个特征波长可有效地从双子叶植物棉花中识别出双子叶杂草刺儿菜, 其识别率高达 100%。

(2) 利用可见光波段范围内的 375, 465 和 585 nm 特征波长, “红边”范围的 705 nm 特征波长、和近红外高原区 1 035 nm 特征波长, 可有效地从单子叶植物水稻中识别出单子叶杂草稗草, 其识别率也能达到 100%。

因为植物光谱反射特性的影响因素很多, 如土壤中的养分以及环境的污染程度^[12-14]。利用光谱特征波长对杂草的识别还需进一步研究, 找到受环境影响较小的波段, 提高判别模型的稳定性和有效性。

参 考 文 献

- [1] Brown R B, Steckler J-P G A, Anderson G W. Transactions of the ASAE, 1994, 37(1): 297.
- [2] Goel P K, Prasher S O, Patel R M, et al. Transactions of the ASAE, 2002, 45(2): 443.
- [3] Vrindts E, De Baerdemaeker J, Ramon H. Precision Agriculture, 2002, 3: 63.
- [4] Slaughter D C, Lanini W T, Giles D K. Transactions of the ASAE, 2004, 47(6): 1907.
- [5] MAO Wen-hua, WANG Yue-qing, WANG Yi-ming, et al(毛文华, 王月青, 王一鸣, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2005, 25(6): 984.
- [6] QIANG Sheng, WEI Shou-hui, HU Jin-liang(强 胜, 魏守辉, 胡金良). Journal of Nanjing Agricultural University(南京农业大学学报), 2000, 23(2): 18.
- [7] LI Dong-po, WU Zhi-jie(李东坡, 武志杰). Chinese J. Rice. Sci. (中国水稻科学), 2002, 16(4): 378.
- [8] Knipling E B. Remote Sensing of Environment, 1970, 1: 155.
- [9] Terashima I, Saeki T. Plant and Cell Physiology, 1983, 24: 1493.
- [10] DeLucia E H, Nelson K, Vogelmann T C, et al. Plant Cell and Environment, 1996, 19: 159.
- [11] Gausman H W. Photogrammetric Engineering and Remote Sensing, 1974, 40: 183.
- [12] CHI Guang-yu, LIU Xin-hui, LIU Su-hong, et al(迟光宇, 刘新会, 刘素红, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2006, 26(7): 1272.
- [13] FENG Lei, FANG Hui, ZHOU Wei-jun, et al(冯 雷, 方 慧, 周伟军, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2006, 26(9): 1749.
- [13] CHEN Shu-ren, LI Yi-xin, MAO Han-ping, et al(陈树人, 栗移新, 毛罕平, 等). Transactions of the Chinese Society for Agricultural Machinery(农业机械学报), 2008, 39(a): 96.
- [14] CHEN Shu-ren, LI Yi-xin, MAO Han-ping, et al(陈树人, 栗移新, 毛罕平, 等). Transactions of the Chinese Society for Agricultural Machinery(农业机械学报), 2008, 39(a): 96.

Research on Distinguishing Weed from Crop Using Spectrum Analysis Technology

CHEN Shu-ren, LI Yi-xin, MAO Han-ping, SHEN Bao-guo, ZHANG Yu-zhu, CHEN Bin

Key Laboratory of Modern Agricultural Equipment and Technology, Ministry of Education & Jiangsu Province, Jiangsu University, Zhenjiang 212013, China

Abstract Automatic detection of weeds is necessary for site-specific application of herbicides or precise physical weed control. Leaf reflectance is mainly determined by photosynthetic pigments, leaf structural properties and water content, so spectral reflectance characteristics can be used for weed discrimination. The spectral reflectance of cotton, rice and weeds was determined in the range from 350 to 2 500 nm using the Analytical Spectral Device Full Range FieldSpec Pro (ASD) in laboratory. The discrimination analysis was done using the statistical software package SAS. The characteristic wavelengths were selected by using STEPDISC procedure. With the selected characteristic wavelengths, discriminant models were developed using the DISCRIM procedure in SAS. For distinguishing spine-greens from cotton, three characteristic wavelengths, 385, 415, and 435 nm, were selected, and good classification performance (100% accuracy) was achieved. The combination of characteristic wavelengths 415 and 435 nm has the biggest contribution to discrimination model. For distinguishing barnyard-grass from rice, five characteristic wavelengths, 375, 465, 585, 705, and 1 035 nm, were selected, and also good classification performance (100% accuracy) was obtained. The transition point from yellow to orange wavelength (585 nm) and the wavelength 705 nm in the red edge contributed more to discrimination model.

Keywords Spectrum analysis; Dicotyledon; Monocotyledon; Weed discrimination

(Received Oct. 6, 2007; accepted Jan. 6, 2008)