

基于光谱分析的罂粟识别研究

王浩^{1,2}, 秦来安^{1*}, 靖旭¹,
何枫¹, 谭逢富¹, 侯再红¹

(1 中国科学院安徽光学精密机械研究所中国科学院大气光学重点实验室, 安徽 合肥 230031;
2 中国科学技术大学, 安徽 合肥 230026;)

摘 要: 根据排查罂粟的需要, 利用光纤光谱仪在室内测量了小麦、罂粟、虞美人和狗尾草等植物在 450~1000 nm 波长范围内的光谱反射率。运用 SPSS 统计软件的逐步判别分析方法获得区分四种植物的 8 个特征波长, 以选定的特征波长点建立判别模型进行分析, 以罂粟错误率及非罂粟错误率为评价指标, 研究了识别效果随特征波长点及特征波长点数目的变化情况。结果表明“红边”附近波长对罂粟及非罂粟的识别有重要影响, 利用特征波长组合 684.0, 706.4, 725.2, 919.2 nm 或 684.0, 694.3, 706.4, 725.2 nm 时, 罂粟错误率及非罂粟错误率可以降低到 0%。

关键词: 光谱; 罂粟识别; 光谱分析; 逐步判别分析

中图分类号: O433

文献标识码: A

文章编号: 1007-5461(2019)02-0151-05

Research of identification of papaver based on spectral analysis

WANG Hao^{1,2}, QIN Lai-an^{1*}, JING Xu¹,
HE Feng¹, TAN Feng-fu¹, HOU Zai-hong¹

(1 Key Laboratory of Atmospheric Optics, Anhui Institute of Optics and Fine Mechanics,
Chinese Academy of Sciences, Hefei 230031, China;
2 University of Science and Technology of China, Hefei 230026, China;)

Abstract: According to the needs of investigation the illicit cultivation of opium poppy, the spectral reflectance of wheat, opium poppy, corn poppy and green bristlegrass in the wavelength range of 450~1000 nm was measured using a fiber optical spectrometer in the laboratory. Using the stepwise discriminant analysis method of SPSS statistical software, eight characteristic wavelength for distinguishing these four species of plants were obtained. The discriminant model is established with the selected characteristic wavelength point. The discriminant model takes poppy recognition error rate and the non-poppy recognition error rate as evaluation index. The variation of the recognition effect with the number of characteristic wavelength points and characteristic wavelength points was studied. The results showed that the wavelengths near the “red-edge” have an important influence on the identification of poppy and non-poppy. Using two combinations of the four characteristic wavelengths of 684.0, 706.4, 725.2, 919.2 nm or 684.0, 694.3, 706.4, and 725.2 nm, the poppy recognition error rate and non-poppy recognition error rate can be reduced to 0%.

Key words: spectroscopy; opium poppy identification; spectral analysis; stepwise discriminant analysis

1 引言

我国目前面临的禁毒形势依然严峻复杂, 禁毒工作任务十分艰巨^[1]。精确识别毒品原植物是禁毒工作的关键。近些年来, 随着遥感技术的发展, 卫星遥感、航空遥感等已经广泛应用到地形测绘、环境监测以

基金项目: Supported by National Key Research and Development Plan(国家重点研发计划, 2016YFC0800900)

作者简介: 王浩 (1990 -), 男, 博士生, 主要从事激光大气探测研究方面的研究。E-mail: 514262022@qq.com

导师简介: 侯再红 (1968 -), 男, 博士, 研究员, 博士生导师, 主要从事激光大气传输方面的研究。E-mail: zhhou@aiofm.ac.cn

收稿日期: 2018-04-13; **修改日期:** 2018-06-30

* 通信联系人。E-mail: laqin@aiofm.ac.cn

及精准农业等方面^[2~6]。国内外的学者对不同植物运用不同手段或者算法提取植物的形状、纹理、颜色和光谱特征进行识别的研究奠定了遥感的基础,这些研究结果也表明基于光谱分析技术的植物识别方法在实时性方面具有明显优势^[7,8]。陈树人等利用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 可以 100% 区分双子叶植物棉花和杂草刺儿菜;5 个特征波长 375,465,585,705 和 1035 nm 可以 100% 区分单子叶植物水稻和稗草^[9]。Vrindts 等在实验室内测量了甜菜、玉米以及七种杂草在 400~2000 nm 波长范围内的反射率,以 555,675,815,1265,1455 和 1665 nm 处的反射率作为分母,利用在很少波段处的反射率比在实验室内可以区分杂草,识别率高达 97%;在田间自然光照条件下测量了作物在 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%^[10]。王海青等利用近红外光谱仪在室内采集了黄瓜和其茎叶在 690~950 nm 波长范围内的反射光谱,在主成分分析基础上结合马氏距离建立识别模型剔除 7 个异常样本后,利用剩余样本进行偏最小二乘法建模,预测值和实际值的相关性达 0.9941,正确识别率达 100%^[11]。

利用遥感手段进行罂粟的排查不仅能够降低所需的人力成本,而且效率较高。但是到目前为止对罂粟植株的相关光谱特性研究较少。陆永帅等人提出了基于深度置信网络模型的机载多光谱数据罂粟识别算法,利用现有的 550 nm,850 nm 和 1200 nm 三通道多光谱相机对罂粟种植田和周围的高粱、啤酒花、树木等进行识别,验证了在小样本的情况下用随机隐退的深度置信网络模型识别罂粟的可行性^[12]。

了解罂粟的光谱特性及其与其他植物的区别对罂粟的遥感识别或者激光铲除都有重要的指导意义。本文结合实时排查非法种植罂粟的需要,测量了罂粟和同时期较为常见的虞美人、小麦、狗尾草等四种植物在 450~1000 nm 波长范围内的反射率,并利用 SPSS 软件筛选出特征波长后建立判别模型进行判别。

2 材料与方法

2.1 实验材料

实验所使用的罂粟和狗尾草植物来自甘肃省农垦农业研究院的罂粟种植基地罂粟田,虞美人来自自种,其叶片大小与罂粟基本相同。小麦从农田中移植到花盆中种植一个星期后进行实验。实验仪器是美国海洋光学公司的 USB2000+ 光纤光谱仪,光源采用卤钨灯,表 1 和表 2 分别为光谱仪和卤钨灯参数。

Table 1 Parameters of USB2000+ fiber optic spectrometer

Parameters	Value
Detector	2048 pixel CCD
Detection rang	350-850 nm or 350-1000 nm
Effective pixels	2048pixel,size 14μm×200 μm
Sensitivity	400 nm : 75 photon/count ; 600 nm : 41photon/count
Light aperture	50 μm
Spectral resolution	1.3 nm
Stray light	at 600 nm<0.05%,at 435 nm<0.10%

Table 2 Halogen lamp parameters

Parameters	Value
Power	6.5 W
Wavelength range	360 ~2500 nm
Stability	0.2%
Drift	0.1%/h
Light bulb color temperature	3100 K

2.2 实验方法

在室内采集了罂粟和虞美人等四种植物叶片的光谱数据,光谱仪经过白板校准后进行测量,光纤探头置于叶片表面下方约 5 mm 处,探头视场约为 12.7 度。叶片采用悬空放置,上方无遮挡,实验时叶片正面朝下,叶面与探头基本垂直,实验室采用全黑背景。每种植物叶片每个部位测量 10 次反射率,每种植物测得 170 组样本数据。

3 数据处理

3.1 压缩数据

由于在光谱仪两端光谱数据信噪比较低，因此实验中仅使用信噪比较高的 450~1000 nm 波段。对每次测量的小麦、罂粟、虞美人和狗尾草的反射率数据首先进行小波去噪处理，然后将同一部位测量的 5 次结果平均来进一步平滑数据，这样就将 680 组数据压缩到了 136 组。光谱仪软件实际输出波长数据间隔为 0.38 nm，在 450 到 1000 nm 之间的原始光谱包含了 1463 个数据点，为了提高计算机的处理速度，采用每 5 个波长合并为 1 个波长点，将 1629 个数据点压缩为 326 个数据点，合并后光谱分辨率约为 1.9 nm。

3.2 寻找特征波长点

以 293 个相应波长的反射率作为每种植物的变量，从每种样品中抽出 24 组数据组成训练数据集，用标记变量把他们分为四类。利用 SPSS 的 Discriminate 中的逐步判别分析方法 (stepwise method) 统计分析数据，寻找判别四种植物的特征波长点。

3.3 建立判别模型

将所有数据放在一起组成工作数据，以筛选出的特征波长点作为变量，应用 SPSS 统计软件的 Discriminate(enter independents together) 判别分析方法建立判别模型。为了寻找全局最优的特征波长点组合，先验证用单个特征波长变量的罂粟和非罂粟的识别情况，选择罂粟错误率 (把罂粟识别为非罂粟的概率) 与非罂粟错误率 (把非罂粟识别为罂粟的概率) 之和最小的波长与其它特征波长点组合成两个变量，然后把罂粟错误率与非罂粟错误率之和最小的两个组合分别与其它特征波长点组合成 3 个变量，如此循环，直至罂粟错误率与非罂粟错误率之和为 0 或 8 个特征波长点全部参与建模判别分析为止。

4 结果与分析

实验所得每种植物样品的平均反射率如图 1。从图中可以看出，罂粟的反射率高于其它三种植物，而狗尾草的反射率低于其他三种植物。四种植物在可见光波段反射率差别较大，在近红外波段罂粟、小麦和虞美人的反射率非常接近。用 Wilk's lambda^[13] 逐步判别分析方法选定的判别罂粟、虞美人、小麦和狗尾草四类植物的特征波长点为 532.3,550.7,629.7,684.0,694.3,706.4,725.2,919.2 nm，其中 684.0,694.3,706.4,725.2 nm 四个特征波长位于 "红边"^[14] 附近。

图 2 为四种植物的平均识别率随入选的特征波长点数目的变化情况。从图中可以看出，四种植物的平均识别率随着特征波长点的增加整体上呈上升趋势。如果选用 4 个特征波长点将样品分为小麦、罂粟、虞美人和狗尾草四类，平均识别率分别为 76.2%,68.4%,93.5%,100%，罂粟的平均正确识别率最低，而四种植物的最高正确识别率分别为 100%,100%,94.4%,100%。

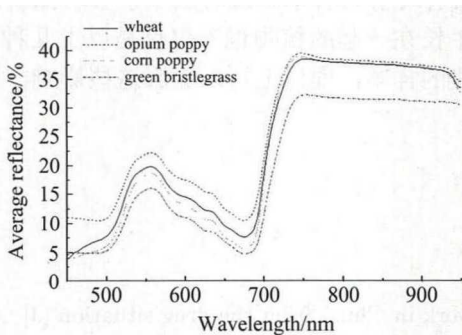


Fig.1 The mean reflectance of samples

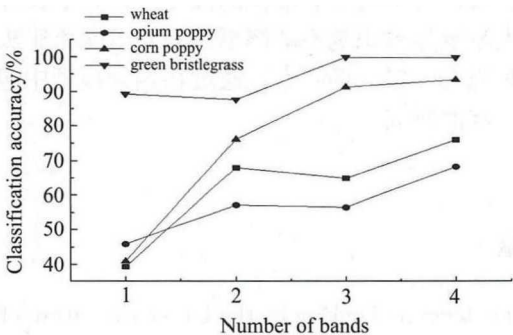


Fig.2 Classification results of the sample

对于非法种植罂粟的排查，只需识别所探测的植物是否为罂粟即可，无需进一步区分是何种植物。用不同特征波长点建立判别模型识别罂粟和非罂粟两类的结果如表 3 所示，从表中可以看出，罂粟识别错误率与非罂粟识别错误率随着特征波长点数的增加，下降较为缓慢，但在四个特征波长点 684.0,706.4,725.2,919.2 nm 或 684.0,694.3,706.4,725.2 nm 组合上时，罂粟识别错误率与非罂粟识别错误率已经降到 0%，而 684.0,694.3,706.4 和 725.2 nm 四个特征波长均位于“红边”附近，可见“红边”附近的波长对罂粟的识别有非常大的作用。

Table 3 The result of discriminant analysis used different numbers of the selected wavelength

Characteristic wavelength (nm)	Poppy misunderstanding rate (%)	Non-poppy misunderstanding rate (%)	Characteristic wavelength (nm)	Poppy misunderstanding rate (%)	Non-poppy misunderstanding rate (%)
532.3	44.4	45.3	684.0,919.2,725.2	44.4	14.3
550.7	55.6	50.8	684.0,725.2,532.3	44.4	23.8
629.7	44.4	45.3	684.0,725.2,550.7	44.4	19
684.0	44.4	4.8	684.0,725.2,629.7	44.4	23.8
694.3	44.4	35.7	684.0,725.2,694.3	44.4	23.8
706.4	55.6	65.9	684.0,725.2,706.4	44.4	14.3
725.2	77.8	53.2	684.0,919.2,629.7,532.3	33.3	23.8
919.2	66.7	96	684.0,919.2,629.7,550.7	44.4	23.8
684.0,532.3	44.4	19	684.0,919.2,629.7,694.3	44.4	19
684.0,550.7	44.4	19	684.0,919.2,629.7,706.4	44.4	14.3
684.0,629.7	33.3	23.8	684.0,919.2,629.7,725.2	33.3	19
684.0,694.3	44.4	23.8	684.0,919.2,725.2,532.3	44.4	9.5
684.0,706.4	44.4	23.8	684.0,919.2,725.2,550.7	33.3	9.5
684.0,725.2	44.4	9.5	684.0,919.2,725.2,694.3	22.2	4.8
684.0,919.2	44.4	0	684.0,,919.2,725.2,706.4	0	0
684.0,919.2,532.3	44.4	19	684.0, 725.2,706.4,532.3	33.3	14.3
684.0,919.2,550.7	44.4	23.8	684.0,725.2,706.4,550.7	44.4	14.3
684.0,919.2,629.7	33.3	23.8	684.0,725.2,706.4,629.7	33.3	4.8
684.0,919.2,694.3	44.4	19	684.0,725.2,706.4,694.3	0	0
684.0,919.2,706.4	44.4	19			

5 结论

本文在室内进行了罂粟及其同属植物虞美人、作物小麦和杂草狗尾草在 450-1000nm 波段光谱测定分析实验.利用 Wilk’slambda 逐步判别分析方法得到了判别时中植物的 8 个特征波长,利用 684.0,706.4,725.2,919.2 nm 或 684.0,694.3,706.4,725.2 nm 这两种组合的特征波长点能够有效的在四种植物当中识别罂粟和非罂粟,正确识别率都达到了 100%。“红边”附近的波长对罂粟的识别有非常大的作用。

上述结果通过对四种植物的分析,验证了运用光谱特征识别罂粟的可能性。但是罂粟的反射特性可能会因生长阶段的不同而有细微的改变。在其生长过程中,叶面有时会附着水滴或者灰尘也会对测得的反射光谱产生影响导致识别率的降低。另外可能和非法种植罂粟生长在一起的植物也不仅仅是以上几种,所以目前的研究有一定的局限性。通过在判别模型中包含更多类别的样本,理论上可以克服这些影响,这还有待于进一步的研究。

参考文献：

[1] Wang Jianwei. Looking at the long-term nature of anti-drug work in China from the drug situation [J]. *Journal of Yunnan Police Officer Academy* (云南警官学院学报), 2014(3):13-18 (in Chinese).

[2] López-Granados F, Torres-Sánchez J, Serrano-Pérez A, *et al.* Early season weed mapping in sunflower using UAV technology: variability of herbicide treatment maps against weed thresholds [J]. *Precision Agriculture*, 2016,17(2): 183-199.

[3] Alsalam B H Y, Morton K, Campbell D, *et al.* Autonomous UAV with vision based on-board decision making for remote sensing and precision agriculture [C]. *Aerospace Conference. IEEE*, 2017: 1-12.

[4] Jia Y, Su Z, Shen W, *et al.* UAV remote sensing image Mosaic and its application in agriculture [J]. *International Journal of Smart Home*, 2016, 10(5):159-170.

- [5] Aries R C, Utomo S. The low-cost UAV-based remote sensing system capabilities for large scale cadaster mapping [C]. *IOP Conference Series: Earth and Environmental Science*, 2016:012005.
- [6] Bhardwaj A, Sam L, Akanksha, *et al.* UAVs as remote sensing platform in glaciology: Present applications and future prospects [J]. *Remote Sensing of Environment*, 2016, 175:196-204.
- [7] Yuan Tianjun, Wang Jiajun, *et al.* The review of application and standards of near infrared spectrometry [J]. *Chinese Agricultural Science Bulletin* (中国农学通报), 2013, 29(20): 190-196 (in Chinese).
- [8] Chai Ying, Ruan Renzhong, Chai Guowu, *et al.* Species identification of wetland vegetation based on spectral characteristics [J]. *Remote Sensing for Land & Resources* (国土资源遥感), 2016, 28(3):86-90 (in Chinese).
- [9] Chen Shuren, Li Yixin, Mao Hanping, *et al.* Research on distinguishing weed from crop using spectrum analysis technology [J]. *Spectroscopy and Spectral Analysis* (光谱学与光谱分析), 2009, 29(2): 463-466 (in Chinese).
- [10] Vrindts E, Jde B, Ramon H. Weed detection using canopy reflection [J]. *Precision Agriculture*, 2002, 3(1):63-80.
- [11] Wang Haiqing, Ji Changying, Chen Kunjie. Research on identification of cucumber, stem and leaf based on spectrum analysis technology [J]. *Spectroscopy and Spectral Analysis* (光谱学与光谱分析), 2011, 31(10): 2834-2838 (in Chinese).
- [12] Lu Yongshuai, Li Yuanxiang, Peng Xishuai. Poppy detection in airborne multispectral data based on deep belief network [J]. *Remote Sensing Information* (遥感信息), 2017, 32(4): 98-103 (in Chinese).
- [13] Alrawashdeh M J, Radwan T. Wilk's lambda based on robust method [C]. *Ism International Statistical Conference. AIP Publishing LLC*, 2017:51-67.
- [14] Zhou Hongyu, Zheng Hongping. The effect and method of quantitative analysis of “Red-Edge” of vegetation [J]. *Remote Sensing Information* (遥感信息), 2010, 2010(4): 112-116 (in Chinese).