

基于支持向量机(SVM)的稻纵卷叶螟危害水稻高光谱遥感识别

石晶晶¹ 刘占宇^{1,2} 张莉丽³ 周 湾³ 黄敬峰^{1,2,*}

(¹浙江大学 农业遥感与信息技术应用研究所, 浙江 杭州 310029; ²浙江省农业遥感与信息技术重点实验室, 浙江 杭州 310029; ³杭州市植保土肥总站, 浙江 杭州 310020; * 通讯联系人, E-mail: hjf@zju.edu.cn)

Hyperspectral Recognition of Rice Damaged by Rice Leaf Roller Based on Support Vector Machine

SHI Jing jing¹, LIU Zhan yu^{1,2}, ZHANG Li li³, ZHOU Wan³, HUANG Jing feng^{1,2,*}

(¹Institute of Remote Sensing and Information Technology Application, Zhejiang University, Hangzhou 310029, China; ²Key Laboratory of Agricultural Remote Sensing and Information System, Zhejiang Province, Hangzhou 310029, China; ³General Station of Plant Protection and Soil fertilizer of Hangzhou, Hangzhou 310020, China; * Corresponding author, E mail: hjf@zju.edu.cn)

Abstract: The spectra of healthy leaves and leaves damaged by the rice leaf roller were measured and analyzed by the method of continuum removal. In the range of 430 - 530 nm and 560 - 730 nm, the band depth and slope were extracted. Then the extracted parameters were chosen as the input vector of the support vector machine (SVM) to design a support vector classifier for the recognition of the leaves damaged by the rice leaf roller. The results confirmed that the classification precision of the SVM with radial basis function(RBF) kernel function was as high as 100% when γ and C were 0.25 and 1, respectively. This could provide theoretic basis for farmers to recognize the rice leaf damaged by the rice leaf roller on time and control it effectively.

Key words: support vector machine; rice leaf roller; hyperspectral remote sensing; continuum removal; rice; insect damage

摘要: 对健康水稻叶片以及受稻纵卷叶螟危害后的水稻叶片进行了室内光谱的测定及分析。对 430 ~ 530 nm 和 560 ~ 730 nm 波段采用连续统去除的方法, 分别提取了波深、斜率参量作为径向基核函数支持向量机的输入变量, 利用 LIBSVM 软件包构建叶片高光谱识别模型。当参数 γ 和惩罚系数 C 分别取 0.25 和 1 时构建的径向基支持向量机模型的性能最佳, 识别精度达 100%。研究结果为实时水稻病虫害的早期监测以及田间管理提供了一定的理论基础。

关键词: 支持向量机; 稻纵卷叶螟; 高光谱遥感; 连续统去除; 水稻; 虫害

中图分类号: S127; S435.112+.1

文献标识码: A

文章编号: 1001-7216(2009)03-0331-04

稻纵卷叶螟 (Rice leaf roller, *Cnaphalocrocis medinalis* Guenée) 属鳞翅目、螟蛾科, 是危害水稻生长和影响产量的主要害虫之一, 广泛分布于全国稻区, 严重时全田叶片枯白, 产量损失严重^[1]。据统计, 2007 年全国稻纵卷叶螟发生面积达 2530 万 hm^2 , 较 2003 年增加 29.3%, 具有发生范围广、危害面积大等特点^[2]。传统的病虫害监测方法主要是基于田间取样、调查以及综合其他信息进行预测预报。这种方法不仅耗时、费力、预报滞后, 且不宜在大范围内展开, 无法进行实时、快速的监测。利用高光谱遥感技术对植被病害进行监测, 已成为病虫害监测的重要研究方向^[3-5]。

高光谱遥感技术已经广泛应用于水稻病虫害的监测中。Kobayashi 等^[6]筛选了稻瘟病敏感波段, 并提出利用黄熟期波段比值 $R550/R970$ 和 $R725/R900$ 估测病情严重程度。吴曙雯等^[7]研究了不同程度水稻叶瘟的反射光谱的变异特征。刘占宇等^[8]利用红边和绿峰等参数识别白穗和正常穗, 精度达到 100%。然而, 目前利用高光谱遥感技术对稻纵卷叶螟危害水稻进行识别的报道较少。

支持向量机 (Support vector machine, SVM) 是建立在 VC 维理论和结构风险最小化原理基础上的机器^[9], 相对于

传统机器, 它最大程度地克服了学习过程中由于高维特征空间计算带来的维数灾难以及过学习等困难^[10-11], 对于模式识别具有良好的泛化能力。

连续统去除 (Continuum removal, CR) 又称基线归一法, 它可以有效地消除背景噪声的影响, 使得光谱吸收特征能够相对于一个共同的基线进行比较。Clark 等^[12]分析了几种不同农作物在 550 ~ 750 nm 连续统去除光谱反射率的差别, 黄木易等^[13]利用连续统去除提取的参数波深与小麦条锈病严重程度呈极显著负相关。

本研究在对健康水稻和受稻纵卷叶螟危害的水稻叶片进行高光谱特征分析的基础上, 利用连续统去除提取光谱吸收特征参数波深和斜率作为输入变量, 构建稻纵卷叶螟危害水稻的径向基支持向量机模型, 为稻纵卷叶螟的监测与及时、有效、针对性治理提供理论依据。

收稿日期: 2008-08-14; 修改稿收到日期: 2008-09-29。

基金项目: 国家 863 计划资助项目 (2006AA10Z203); 国家科技支撑计划资助项目 (2006BAD10A01)。

第一作者简介: 石晶晶 (1984 -), 女, 硕士研究生。

1 材料与方 法

1.1 田间试验描述

试验数据主要来源于 3 个不同的试验地点。分别为试验点 1 浙江省杭州市萧山区瓜沥镇运东村 (30°27' N, 120°12' E); 试验点 2 :浙江省杭州市萧山区蜀山街道章潘桥村 (30°06' N, 120°15' E); 试验点 3 :浙江大学现代化农业研究示范中心 (30°14' N, 120°10' E)。供试土壤为砂壤土 (砂粒 83.49%、粉粒 12.15%、黏粒 4.36%), 土壤全氮含量为 1.15 g/kg, 速效氮含量为 188.5 mg/kg, 全磷 1.21 g/kg, 全钾 72.7 mg/kg, 有机质 9.96 g/kg, pH 值 6.78。3 个试验点自然条件接近。种植水稻为杂交稻。

参考 GB/T 15793 - 1995《稻纵卷叶螟测报调查规范》, 对分蘖期水稻受害情况进行了调查, 试验点卷叶率分别为 19.8%、8.9%和 11.1%。稻叶受害程度为 1 级, 受害程度较轻。且本试验中采集的非健康叶片因受害形成的白色条纹长约占叶片面积的 5% ~ 25%。

1.2 试验仪器

本研究使用美国 ASD (Analytical Spectral Device) 公司生产的便携式光谱辐射仪 Field Spec Pro FR™ 光谱仪, 波段值为 350 ~ 2500 nm, 其中 350 ~ 1000 nm 光谱采样间隔为 1.4 nm, 光谱分辨率为 3 nm; 1000 ~ 2500 nm 光谱采样间隔为 2 nm, 光谱分辨率为 10 nm。模拟光源为与光谱辐射仪配套的 50 W 卤灯。

1.3 光谱测定

在水稻分蘖期从田间采集健康和受稻纵卷叶螟危害的水稻植株, 置于放有冰块的恒温箱中带回实验室进行光谱测定。将叶片样品置于反射率近似为零的黑色橡胶上, 方向与光源保持一致, 固定光谱仪探头, 并使光谱仪探头垂直向下, 正对待测叶片受害部位, 光谱仪视场角为 8°, 探头距叶片表面垂直距离 3 cm; 方位角 (光源与叶片表面的夹角) 45°。从叶片基部开始采集光谱至叶尖结束, 每片叶子在不同部位测定 4 次, 每次采集 5 条光谱, 以其平均值作为观测叶片的光谱反射值, 每次测量前后进行标准白板校正。

其中, 在试验点 1 和 2 的水稻植株中, 测定了 38 个叶片样本 (健康和受害的叶片各为 19 个), 试验点 3 中测定了 70 个叶片样本 (健康叶片和受害叶片样本各 35 个), 总计 108 个样本。

1.4 支持向量机 (SVM) 简介

SVM 是从线性可分情况下的最优超平面发展起来的机器学习。输入向量 x 通过非线性变换映射到一个高维特征空间 Z , 以结构风险最小化为原则, 在这个高维空间中构造最优分类超平面实现分类^[9]。SVM 分类函数类似一个神经网络, 包括输入层、中间层和输出层三部分。要实现 SVM 分类需要通过定义内积函数实现输入向量的非线性变换, 每一个中间层节点对应于输入样本与一个支持向量的内积 $K(x, x_i)$, 输出是多个中间层节点的线性组合^[14], 输入向量通过决策函数 $f(x) = \text{sgn}[\sum_{i=1}^l y_i K(x_i, x) + b]$ 映射到一个高维空间

中实现样本分类。输出层代表了类别, 本文中将样本集分为健康和虫害两类, 分别用 +1、-1 作为输出值, 以径向基 (RBF) 函数 $K(x_i, x) = \text{Exp}(-\|x_i - x\|^2 / \sigma^2)$ 作为 SVM 中的核函数建立分类模型。SVM 分类器的构建以及预测都是基于林智仁教授开发的 LIBSVM 软件包实现的。构建模型的关键是要确定参数 (C, γ) 以达到对未知样本的准确预测。其中, 惩罚系数 C 有调和分类中追求最大化分类间隔和最小化错误之间矛盾的作用, C 值越大意味着更强调最小化训练错误。

1.5 连续统去除

连续统定义为用直线逐点连接光谱特征曲线上凸出的“峰”值点, 并使折线在“峰”值点上的外角大于 180°, 作为 100% 基线^[15]。连续统去除是实际光谱反射曲线上的反射率与连续统上相对应波段的反射率的比值^[4]。基于连续统的思想可以获取一些光谱吸收特征参数, 如波深和斜率等。

波深 (Band depth, BD) 可以理解为特征峰的极小值相对于 100% 基线的距离。吸收特征斜率 K 可以表示为: $K = (R_e - R_s) / (e - s)$, 其中, s, e 分别是吸收起点和终点处的波长, R_s, R_e 分别是 s, e 对应的光谱反射率。

2 结果与分析

2.1 水稻受稻纵卷叶螟危害后光谱特征分析

为了减少光谱数据中的环境和仪器噪声, 首先运用五步滑动平均法进行数据预处理。图 1 为健康和受稻纵卷叶螟危害的水稻叶片光谱特征, 从图中可以看出健康叶片和受害后叶片的光谱特征存在明显的差异, 近红外平台变化较可见光波段显著。其中, 在可见光范围内以及短波红外 2080 ~ 2350 nm 处受稻纵卷叶螟危害后的叶片比健康叶片的平均光谱反射率有明显升高, 但是在可见光波段不存在明显的以 500 nm 为中心的叶绿素蓝光吸收谷, 在 490 ~ 530 nm 蓝边波段处受害叶片的光谱反射率呈逐渐上升趋势, 而近红外波段 760 ~ 1300 nm 和短波红外波段 1560 ~ 1750 nm 处受胁迫的叶片反射率降低, 以近红外平台区的降低最为显著, 近红外陡坡效应明显削弱, 曲线趋势平缓。

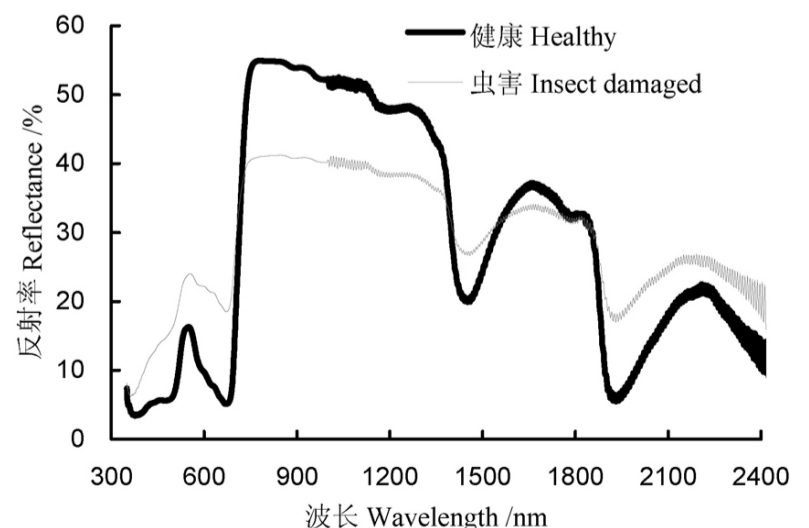


图 1 健康和受稻纵卷叶螟危害后叶片的光谱特征曲线
Fig. 1 Reflectance spectra of healthy rice leaf and leaf damaged by rice leaf roller.

植物光谱特征的变化受光合色素以及细胞结构变化影响^[16]。在水稻受害初期,稻纵卷叶螟的幼虫将叶片缀合成管状虫苞,幼虫取食叶片上表皮导致了叶片表皮细胞受损,而叶片的表皮结构的变化可以改变光的反射^[17],叶片变薄,细胞层数的减少,细胞间隙改变,导致了近红外波段反射率的降低。随着幼虫进一步啃食叶肉,叶肉细胞中的叶绿体结构遭到破坏,叶片表现为白色条斑,使得叶绿素对蓝、红光的吸收降低,表观发射率增强。由于上表皮细胞遭受破坏叶片的蒸腾作用增强以及叶肉细胞破坏同时共同引起的叶片含水量下降是造成短波红外处光谱曲线变化的主要原因。

2.2 光谱吸收特征参数选取

利用高光谱数据进行水稻叶片的识别主要是基于光谱吸收特征参数的选取,光谱吸收特征参数主要包括吸收波段的波深和斜率等。光谱吸收特征参数可以通过连续统去除的方法对数据进行归一化处理后得出。根据光谱特征分析,选择470 nm和670 nm叶绿素吸收峰附近的波段进行连续统去除,其连续统的具体参数见表1。

通过对水稻叶片在430~530 nm和560~730 nm波段内的反射光谱进行连续统去除处理,将光谱反射率归一到同一个基线,扩大了叶绿素蓝光和红光吸收谷特征。表2中的数据显示,稻纵卷叶螟危害以后的水稻叶片在蓝光和红光吸收谷的吸收深度均小于健康叶片,分别为健康叶片的29.3%和51.1%,在430~530 nm之间光谱吸收特征参数斜率变化较小。然而,随着受稻纵卷叶螟危害后的叶片近红外陡坡效应的削弱,叶片在以670 nm为中心的叶绿素a吸收波段斜率是健康叶片的51.7%。通过已选择的光谱吸收特征参数

表1 连续统去除的参数设置

Table 1 .Parameters of continuum removal .

特征吸收峰 Characteristic absorption peak/nm	连续统起点 Start point of continuum /nm	波深中心 Band center /nm	连续统终点 End of continuum /nm
470	430	490	530
670	560	671	730

表2 水稻叶片光谱吸收特征参数统计分析

Table 2 .Statistical analysis on spectral absorption characteristic parameters .

参数 Parameter	430 ~ 530 nm				560 ~ 730 nm			
	波深 Band depth(BD ₁)		斜率 Slope(K ₁)		Band depth(BD ₂)		斜率 Slope(K ₂)	
	健康叶片 Healthy leaves	受害叶片 Insect damaged leaves	健康叶片 Healthy leaves	受害叶片 Insect damaged leaves	健康叶片 Healthy leaves	受害叶片 Insect damaged leaves	健康叶片 Healthy leaves	受害叶片 Insect damaged leaves
最小值 Min	0.44163	0.02795	0.00078	0.00079	0.84264	0.21312	0.00123	0.00047
最大值 Max	0.54841	0.35672	0.00123	0.00120	0.88430	0.73508	0.00169	0.00109
标准差 SD	0.02747	0.06475	0.00011	0.00010	0.00843	0.11191	0.00009	0.00014
平均值 Mean	0.48810	0.14299	0.00097	0.00102	0.85998	0.43916	0.00152	0.00077

进行稻纵卷叶螟危害水稻叶片的高光谱识别,避免了由于高光谱数据的信息量大带来的运算上的困难。

2.3 基于SVM的水稻高光谱识别

本研究共测定了108个样本的光谱数据用于识别模型构建,其中70个样本数据作为训练样本集,38个作为测试集。基于连续统去除处理后的光谱特性分析,选择吸收峰波段深度BD₁、BD₂,斜率K₁、K₂等4个特征值,分别得到70×4的训练矩阵和38×4的测试矩阵。

首先对输入向量进行归一化处理,将数据缩放在区间[0,1]范围内,避免一些特征值范围过大而另一组特征值范围过小,在计算内积时引起数值计算的困难。然后调用LIBSVM软件中的svmtrain命令对训练样本采用3次(3 fold cross validation)交叉验证法,即将数据集随机分为3组,每1组依次作为测试集,剩余部分作为训练集,计算3次的平均分类精度,通过不断尝试改变参数(C,)的取值比较分类精度^[18]。经过反复筛选,确定了当参数C=1、=0.25,其他参数使用系统默认值时,交叉验证的平均精度为100%。利用筛选的最优参数对训练样本进行训练建模,所需要支持向量数为9个,经检验该模型分类精度可以达到100%,实现了健康样本和虫害样本的识别,从而构建了适用于稻纵卷叶螟危害水稻的识别SVM模型。对于已经训练好的网络,参数权值是固定不变的,为了检验所构建模型分类性能验证其适用性,最后利用svmpredict命令根据已建立的模型对测试样本进行预测。分类结果见表3,表中数据显示已构建的SVM对测试样本可以准确无误地进行识别,表明利用高光谱数据建立的稻纵卷叶螟初期危害水稻识别模型具有良好的分类效果。

3 讨论

稻纵卷叶螟是我国的一种重要性迁飞性害虫,对水稻产量和品质影响较大。因此,在农业生产中实现对稻纵卷叶螟危害初期受害水稻的识别具有重要的意义,它可为植保人员采取针对性地预防和治理提供科学的决策支持。

本研究对健康叶片和受害初期叶片的光谱特征进行了分析,结果表明在可见光波段水稻叶片受到稻纵卷叶螟危害后的光谱反射率高于健康叶片,且在500nm处没有显著的

表 3 基于 SVM 的测试样本分类结果

Table 3 . Classification results of testing data set with SVM .

实际分类 Actual classification	SVM 分类 Classification results with SVM			分类精度 Classification accuracy/%
	健康 Healthy	虫害 Insect damaged	合计 Total	
健康 Healthy	19	0	19	100
虫害 Insect damaged	0	19	19	100
合计 Total	19	19	38	100

叶绿素吸收谷 而近红外平台区域的变化趋势与可见光波段相反。连续统去除是一种有效去除背景的方法,通常用于光谱特征提取,选择 430~530 nm 和 560~730 nm 两个特征吸收波段区间的波深和斜率作为 SVM 的输入向量,当 $\gamma = 0.25$ 、 $C = 1$ 时,构建了径向基 SVM 模型,并运用测试样本对模型的性能进行了检验,分类精度高达 100%。表明 SVM 具有适应高维特征空间、小样本学习、泛化能力强、简便快捷等优点,进一步证实了 SVM 在水稻病虫害的高光谱监测中的可行性。

高光谱遥感作为一种非接触的、无伤的技术在病虫害监测中显示了其发展潜力,结合 SVM 模型实现稻纵卷叶螟早期危害水稻的准确快速识别,省时省力,同时可以减少监测中存在的人为误差。然而本研究结论是基于室内光谱数据得出的,由于室内条件控制严格,是否适用冠层以及航天航空遥感监测有待于下一步研究,但本文的研究思路为病虫害航空航天遥感监测奠定了一定的理论基础。

参考文献:

- [1] 苏祖芳,周纪平,丁海红.稻作诊断.上海:上海科学技术出版社,2007:211-213.
- [2] 刘宇,王建强,冯晓东,等.2007 年全国稻纵卷叶螟发生实况分析与 2008 年发生趋势预测.中国植保导刊,2008,28(7):33-35.
- [3] 陈鹏程,张建华,雷勇辉.高光谱遥感监测农作物病虫害研究进展.农业信息科学,2006,20(2):388-389.
- [4] 浦瑞良,宫鹏.高光谱遥感及其应用.北京:高等教育出版社,2000:3-93.
- [5] 邢素丽,张广录.我国农业遥感的应用现状与展望.农业工程学报,2003,19(6):174-176.
- [6] Kobayashi T, Kanda E, Kitada K, et al. Detection of rice panicle blast with multispectral radiometer and the potential of using airborne multispectral scanners. *Phytopathology*, 2001, 91:316-323.
- [7] 吴曙雯,王人潮,陈晓斌,等.稻叶瘟对水稻光谱特性的影响研究.上海交通大学学报:农业科学版,2002,20(1):73-76.

- [8] 刘占宇,孙华生,黄敬峰.基于学习矢量量化神经网络的水稻白穗和正常穗的高光谱识别.中国水稻科学,2007,21(6):658-662.
- [9] Vapnik V N.统计学习理论的本质.张学工,译.北京:清华大学出版社,2000:95-96.
- [10] 王海光,马占鸿,王韬,等.高光谱在小麦条锈病严重度分级识别中的应用.光谱学与光谱学分析,2007,27(9):1811-1814.
- [11] 张录达,苏时光,王来生,等.支持向量机(SVM)在傅里叶变换近红外光谱分析中的应用研究.光谱学与光谱学分析,2005,25(1):33-35.
- [12] Clark R N, King T V V, Ager C, et al. Initial vegetation species and senescence/stress indicator mapping in the San Luis Valley, Colorado using imaging spectrometer data//Posey H H, Pendelton J A, van Zyl D. Proceedings of the Summitville Forum 95. Colorado Geological Survey Special Publication 38. Colorado: Colorado Geological Survey, Department of Natural Resource, 1995:64-69.
- [13] 黄木易,黄文江,刘良云,等.冬小麦条锈病单叶光谱特性及严重度反演.农业工程学报,2004,20(1):176-180.
- [14] 张耀波,张迁.基于 SVM 的遥感影像的分类.地理空间信息,2005,3(4):24-26.
- [15] 田庆久,宫鹏,赵春江,等.用光谱反射率诊断小麦水分状况的可行性分析.科技通报,2000,45(24):2645-2650.
- [16] Blakeman R H. The identification of crop disease and stress by aerial photography//Steven M D, Clark J A. Applications of Remote Sensing in Agriculture. London: Butterworths, 1990:229-245.
- [17] Barnes J D, Cardoso Vilhena J. Interactions between electromagnetic radiation and the plant cuticle//Kerstiens G. Plant Cuticles: An integrated functional approach. Oxford: BIOS Scientific Publishers, 1996:157-170.
- [18] Chang C C, Lin C J. LIBSVM: A library for support vector machines [CP/OL]. [2008 08 01]. <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.