

基于光谱特征点秩和检验的稻曲病发病程度检测

桑佳茂, 陈丰农*

杭州电子科技大学自动化学院, 浙江 杭州 310018

摘要 利用无人机航拍获得稻田高光谱图像并计算特征点的归一化植被指数(NDVI), 通过 NDVI 值秩和检验的方法对稻曲病发病程度进行检测。在中国水稻研究基地选取 28 块面积相同的相邻水稻试验区, 采用 4 种不同农田管理方式, 分别为自然生长和喷洒 3 种不同农药, 每种管理方式有 7 个不同播种日期, 相邻地块前后播种日期相差 1 周, 依次递减, 每个区域种植水稻 500 株左右。在稻曲病发病盛期, 先实地调查水稻的发病状况, 根据单位面积内稻穗的发病数目得到发病指数。然后用无人机载高光谱相机拍摄试验田得到所有试验区的高光谱图。对各个高光谱图像进行高质量拼接, 得到完整的覆盖整个试验区域的高光谱图像。对光谱数据预处理后从中提取最能反映稻曲病发病程度的归一化植被指数, 用箱型图清洗数据以去除异常值, 对不同水稻试验区的 NDVI 值进行秩和检验, 第一步对总样本进行秩和检验, 验证各组样本中是否有显著差异, 确定差异性来源于何组样本; 第二步对 4 组样本排列组合共得到 11 组不同组合方式的待检验样本, 对这 11 组样本数据分别进行秩和检验。每组得到的显著性水平均远 < 0.01 , 说明不同组间的样本数据存在极显著差异, 反映出此方法用于稻曲病发病程度检测的合理性。为了显示不同发病区域, 用不同的颜色对不同的稻曲病发病程度的种植区域进行标注。最后以大田水稻发病指数作为对照组, 与秩和检测的结果进行对比, 结果表明秩和检验对稻曲病的发病程度检测有可行性。

关键词 病害; 高光谱; 光谱分析; 水稻; 秩和检验; 稻曲病; 发病程度

中图分类号: S237 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2021)10-3214-06

引言

水稻是中国重要的粮食作物, 全国以水稻为主食的人口数量超过 50%。在水稻病害中, 稻曲病号称水稻的癌症, 目前该病有扩大趋势。稻曲病不仅影响水稻产量, 而且病原菌含有毒素, 对人有毒害作用。目前对稻曲病没有好的根治方法, 主要依赖于喷洒农药进行预防, 但农药对环境和人们的身体健康又会带来隐患。

近年来, 随着遥感技术的不断发展和普及, 植物病虫害领域的遥感监测研究不断增多并迅速发展。Tobias B. Hank 等^[1]用卫星获取的高光谱数据分析了稻曲病的光谱特征。马德贵等^[2]用高光谱成像技术对稻曲病的发病等级划分方法建立了分类模型。吴曙文等^[3]用遥感卫星技术, 在红外光谱波段研究了患有不同程度稻曲病的叶片和冠层反射光谱光谱特征变化。在水稻的类似病症上也有不少学者做了相关的研究, 有报道利用红外遥感技术确定了水稻黑条矮缩病的发病

等级判定依据, 最后选取了五个有效波段的光谱信息, 用支持向量机分类方法建立了预测模型。有研究将不同植被指数引入到植物病虫害的监测^[4-5]。不同形式的植被指数都有一定的适用环境和针对性, 是植物光谱信息的一种重要应用方式。

目前植物病虫害检测主要借助卫星遥感、无人机遥感、彩色图像算法和高光谱图像处理等技术。由于大部分卫星遥感技术采集的图片质量不是很高, 造成关键像素信息的丢失, 浑浊, 大部分病虫害检测系统不能发挥很好的作用, 主流的图像病虫害算法有主成分分析法(principal component analysis, PCA)和聚类方法等, 以及这些方法的结合所产生的方法。但是这些传统算法都是针对清晰图像像素本身的操作, 一旦图像不清晰, 就会造成病虫害检测效果不好、不够准确。

本研究提出一种利用高光谱图像获取水稻归一化植被指数(normalized vegetation index, NDVI), 并将其结合秩和检验^[6]应用在水稻稻曲病的发病程度检测。研究中对多种管理

收稿日期: 2020-09-10, 修订日期: 2021-01-05

基金项目: 国家自然科学基金项目(41671415), 浙江省基础公益研究计划项目(LGN18F030002)资助

作者简介: 桑佳茂, 1996 年生, 杭州电子科技大学自动化学院硕士研究生 e-mail: 925867825@qq.com

* 通讯作者 e-mail: fnchen@hdu.edu.cn

方式试验稻田的水稻生长状况进行监测,用无人机航拍获得试验田的高光谱图像,得到每个试验水稻种植区域的归一化植被指数,对不同稻区的 NDVI 值进行秩和检验,结果 p 值远 < 0.01 ,表明对稻曲病的发病程度检测有显著性。

1 实验部分

1.1 试验地点与材料

本研究的试验地点在中国水稻研究所,位于浙江省杭州市富阳区(30°05'N, 119°95'E),亚热带季风气候,气候温和,日照充足。试验材料为甬优 12 号水稻,该品种生长整齐,植株较高,株型较紧凑,剑叶挺直而内卷,叶色浓绿,茎秆粗壮;分蘖力中等,穗大粒多,着粒密,穗基部枝梗散生。在研究区选取 28 个面积相同的相邻水稻试验区,区域内采用 4 种水稻管理方式,分别为自然生长和通过植保无人机喷洒 3 种不同农药,每种管理方式有 7 个不同播种日期,前后播种日期相差 1 周,依次递减,每个区域有水稻 500 株左右。试验时先调查水稻的发病状况,并用无人机载高光谱相机距地 30 m 高空拍摄试验田。28 个地块田间分布如表 1 所示。

表 1 2019 年度水稻试验田的种植分布

Table 1 Planting distribution of rice test field in 2019

种植日期				
第一期(06-15)	A	B	C	D
第二期(06-20)	A	B	C	D
第三期(06-25)	A	B	C	D
第四期(06-30)	A	B	C	D
第五期(07-5)	A	B	C	D
第六期(07-10)	A	B	C	D
第七期(07-15)	A	B	C	D

注: A, B, C 组采用不同农药抑制病虫害, D 组试验田作为对照组自然生长,未作任何病害抑制处理。A 组: 40% 氟硅唑 18 mL · 亩⁻¹, 50% 肟菌酯 15 g · 亩⁻¹; B 组: 75% 肟菌酯 · 戊唑醇 15 mL · 亩⁻¹; C 组: 40% 氟硅唑 18 mL · 亩⁻¹, 12.5% 烯唑醇 40 g · 亩⁻¹

Note: Groups A, B, and C used different pesticides to suppress diseases and insect pests, and the test field in group D grew naturally as a control group without any disease suppression treatment. Group A: 40% flusilazole 18 mL · mu⁻¹, 50% trifloxystrobin 15 g · mu⁻¹; group B: 75% trifloxystrobin · tebuconazole 15 mL · mu⁻¹; group C: 40% flusilazole 18 mL · mu⁻¹, 12.5% diniconazole 40 g · mu⁻¹

1.2 数据采集

高光谱相机为德国 Cuber 公司的 UHD185,可画幅式采集数据,采集速度可达 5 帧 · s⁻¹,光谱范围 450~1 000 nm,光谱波段分辨率为 8 nm,图像分辨率为 1 000×1 000 像素;在采集数据之前对高光谱相机进行白板和黑板校正,同时采用辐照度计校正,以避免因为光照变化而引起的数据不一致现象。该高光谱相机的无人机挂载平台为大疆经纬 M600 PRO(图 1)。航拍高度 30m,横向重叠度为 0.7,纵向重叠度为 0.8^[7]。高光谱图像拼接采用倾斜摄影三维建模软件 Pho-

toScan,根据航片坐标、高程信息和相似度对多个高光谱样本排序,进行高质量拼接^[10],最后得到一个较完整的覆盖整个试验区域的高光谱图像(图 2)。数据分析平台 ENVI(Vision 5.3, Exelis Inc, Boulder, CO, USA)是一个强大的遥感影像处理平台,可在集成开发环境(integrated development environment, IDL)的辅助下根据算法二次开发,而数据分析平台 Cube-Pilot(Cubert GmbH, Ulm, Germany)则用于光谱的重采样,该软件的交互性较强,方便光谱数据的管理。在 28 个地块中,每个地块随机选择 10 000 个点,最终获得 28 万个光谱数据。光谱数据处理与分析使用 MATLAB (Version 2019, The Mathworks Inc., Natick, MA),对光谱数据进行了平滑和光散射校正,以增强样本的可靠性。



图 1 大疆经纬 M600 PRO

Fig. 1 DJI Matrice 600 PRO

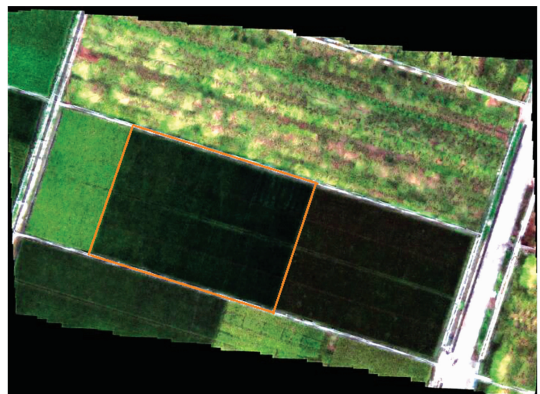


图 2 试验田无人机(UAV)航拍高光谱图

注:图 2 采集时间为 2017 年 9 月 18 日;水稻处于黄熟期;气候多云,无风;航拍路线为从西往东往返。

Fig. 2 Aerial photograph of Unmanned Aerial Vehicle (UAV) with hyperspectral device in test field

Note: Figure 2 was collected on September 18, 2017; the rice is in the yellow ripening period; the climate is cloudy and there is no wind; the aerial photography route is from west to east.

1.3 水稻病情指数等级划分

施辰子等^[9]将稻曲病分级标准分为 5 级,其中 0 级表示没有病害、1 级表示发病最轻、5 级表示发病最严重,以单位面积内的植株发病率作为病害等级的分类标准,具体的描述如下:

- 0 级: 植株稻穗患病率为 0;
 1 级: 植株稻穗患病率为 0~5%(含 5%);
 2 级: 植株稻穗患病率为 5%~10%(含 10%);
 3 级: 植株稻穗患病率为 10%~20%(含 20%);
 4 级: 植株稻穗患病率为 20%~50%(含 50%);
 5 级: 植株稻穗患病率为 50%以上。

按此标准, 根据田间病情调查结果, 计算病情指数(disease index, DI)如式(1)

$$DI = \frac{(\sum x_i \times n_i)}{(N \times k)} \times 100 \quad (1)$$

式(1)中, DI 为病情指数; x_i 为各级发病穗数(单位/株); n_i (1~5)为各级代表值; N 为调查总穗数(单位/株); k ($k=5$) 最高级代表值。为了计算方便, 最后将 28 个试验田按病情指数的大小分为相应的等级。

1.4 检验方法

显著性检验要先对总体样本提出一个假设, 然后通过各种检验方法的计量分析处理样本数据, 依据处理结果判断这个假设是否合理。根据检验样本总体的分布情况, 可分为参数检验和非参数检验。参数检验要求样本服从正态性分布, 并具有相同的方差。当数据不满足正态性假设和方差齐次性假定时, 不适用参数检验^[10]。

在实际工作中, 由于假定统计数据的不可知性, 在一个完整的统计工程中, 必须先检验数据的正态性和方差齐性。使用 MATLAB 的 *lillietest* 正态检验函数和 *vartestn* 方差齐性检验得知本试验样本集不符合正态性假设。考虑到试验数据集需要满足的要求和本次检验属于单因素一元方差分析, 本研究采用基于秩和的非参数检验。因为秩和检测不依赖于总体分布的具体形式, 应用时可以不考虑被研究对象为何种分布以及分布是否已知。

1.5 分析模型

水稻光谱和发病等级之间的关系属于非确定问题。高光谱图像具有波段数目多、光谱数据量大和分辨率高等特点, 导致数据维数较高, 如果计算全光谱波段, 将会产生较多冗余数据, 降低结果的准确性。本研究用归一化植被指数建立水稻光谱和水稻发病等级之间的关系, 用显著性检验验证分类的合理性。

归一化植被指数(NDVI)是植物生长状况以及植被空间分布密度的指示因子, 能反映植物冠层的背景影响, 且与植被覆盖有关, 与植被分布密度呈线性相关, 在使用遥感图像进行植被研究中得到广泛应用。另外 NDVI 值在植物叶绿素吸收较强的波谱区间蓝光波段计算得到, 选取适当的波段计算可以将植被的地物特征加强, 突出植被的冠状特征。本研究通过植被指数确定特征波段, 再选择对应的光谱点作为反映稻曲病等级识别的特征。

NDVI 的值由近红外波段(R_{Nir})和红光波段(R_{Red})的反射率决定。NDVI 值范围在 $[-1, 1]$ 之间, 负值表示地面覆盖为云、水、雪等; 0 表示有岩石或裸土等覆盖; 正值表示有植被覆盖, 且随覆盖度增大而增大。其计算如式(2)所示

$$NDVI = \frac{R_{Nir} - R_{Red}}{R_{Nir} + R_{Red}} \quad (2)$$

式(2)中, R_{Nir} 为近红外波段的反射值; R_{Red} 为红光波段的反射值。

本研究采集光谱数据波段范围为 450~946 nm。对于近红外波段和红光波段的取值问题, 国内外专家和学者也有研究, 王福民等^[11]研究表明, 红光波段的反射率对 NDVI 的影响更大, 而在波段宽度 <50 nm 时, 近红外波段基本不受波段位置和宽度的影响; 张竞成等^[12]研究表明, 归一化植被指数 NDVI 中的红光波段影响较大, 红光波段的反射值如果在 660~680 nm 为中心的窄波段范围内选取, 可以更好地反应植物的实际生长状况, 近红外波段的反射值在 740~980 nm 范围内选取。

根据罗红霞等^[13]综合分析, 红光波段的位置对 NDVI 结果的影响较大, 且红光波段位置接近红谷极值, 即 670 nm 附近时影响尤为显著。本研究所选的光谱波段间隔 4 nm, 查阅文献综合考虑将 670 nm 处的取值作为红光波段的数值, 对比多组患病水稻和健康水稻的光谱图(如图 3), 通过区分度比较, 将 834 nm 处的取值作为近红外波段数值。

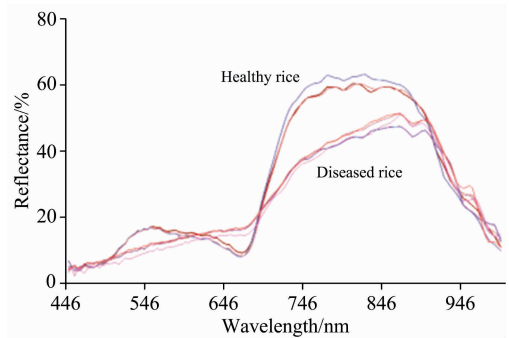


图 3 未患稻曲病水稻和患稻曲病水稻光谱
 Fig. 3 Spectra of rice without rice smut disease and rice with rice smut disease

本研究具体实施过程: (1)无人机载高光谱拍摄 28 个水稻试验区的高光谱图像, 提取各区域指定波段的光谱数据,

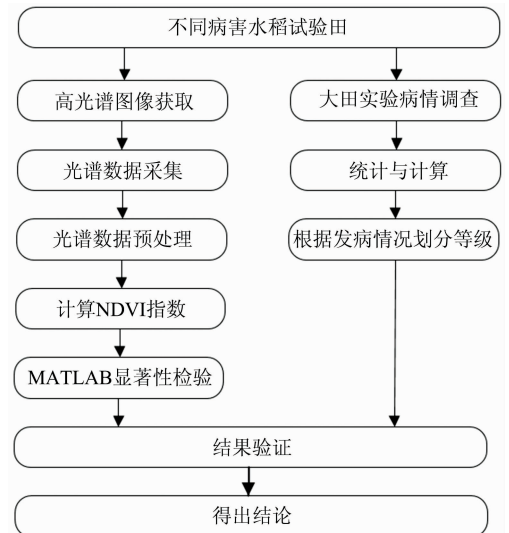


图 4 稻曲病秩和检验技术路线

Fig. 4 Technical route of rank sum test for rice smut disease

每个试验区各选择 10 000 个采样光谱；(2)光谱数据预处理，增强光谱特征；(3)计算每个试验区的 NDVI 值；(4)用得到的 NDVI 值做秩和检验，根据输出结果 p 值决定分类的显著性；(5)将检验结果与实际大田情况作对比，验证结论可靠性。

稻曲病和检验技术路线如图 4 所示。

2 结果与讨论

2.1 稻曲病发病等级划分

作为对照组，水稻的病害等级通过实地的大田调查得到。本研究统计了试验地块的发病稻穗数量，并参照 1.3 节中的方法对不同患病程度水稻进行区分和统计，计算了每块区域的发病指数(表 2)。通过病情指数对发病等级分为 1, 3,

表 2 不同试验区的稻曲病发病情况

Table 2 Disease grade of Rice False Smut in different experimental areas

水稻标号	施辰子稻曲病发病等级划分法						病情指数
	0级	1级	2级	3级	4级	5级	
雨优十二一期 A	500	0	0	0	0	0	0
雨优十二一期 B	500	0	0	0	0	0	0
雨优十二一期 C	499	1	0	0	0	0	0.022
雨优十二一期 D	498	1	1	0	0	0	0.089
雨优十二二期 A	499	0	0	1	0	0	0.111
雨优十二二期 B	498	0	2	0	0	0	0.133
雨优十二二期 C	500	0	0	0	0	0	0
雨优十二二期 D	499	0	0	0	0	0	0
雨优十二三期 A	500	0	0	0	0	0	0
雨优十二三期 B	498	0	2	0	0	0	0.133
雨优十二三期 C	499	0	1	0	0	0	0.067
雨优十二三期 D	498	4	3	0	0	0	0.286
雨优十二四期 A	469	16	7	5	2	1	1.889
雨优十二四期 B	482	10	2	4	2	0	1.111
雨优十二四期 C	474	16	6	2	2	0	1.289
雨优十二四期 D	469	16	7	5	2	1	1.889
雨优十二五期 A	481	10	1	6	2	0	1.267
雨优十二五期 B	433	32	20	14	1	3	4.330
雨优十二五期 C	423	24	19	15	8	6	5.971
雨优十二五期 D	363	59	27	24	13	14	10.6
雨优十二六期 A	560	19	9	8	5	1	2.399
雨优十二六期 B	515	19	16	9	2	1	2.669
雨优十二六期 C	513	12	10	8	4	0	2.234
雨优十二六期 D	191	10	2	10	4	0	4.813
雨优十二七期 A	431	69	23	24	2	4	6.188
雨优十二七期 B	373	59	33	40	11	12	11.43
雨优十二七期 C	429	45	9	15	2	1	3.770
雨优十二七期 D	129	41	8	13	4	9	13.017

注：A, B 和 C 组采用不同农药抑制病虫害，D 组试验田作为对照组自然生长，未作任何病害抑制处理。

Note: Groups A, B, and C used different pesticides to suppress diseases and insect pests, and the test field in group D grew naturally as a control group without any disease suppression treatment.

5 和 7 四个等级，分别对应于病情指数的条件为 $DI < 1$, $1 < DI < 3$, $3 < DI < 5$ 和 $DI > 5$ 。本研究中等级的划分主要根据病情指数的分布梯度。

2.2 基于秩和的稻曲病显著性检验

对不同试验区的 NDVI 值进行秩和检验，验证各组样本中是否有显著差异，如果有差异再两两比较，确定差异性来源于何组样本。结果输出如表 3 所示，由表 3 的数据可知， $p < 0.01$ ，该值极为显著。

表 3 等级秩和检验输出表

Table 3 Output of rank sum test

方差来源	平方和	自由度	均方差	卡方统计量	p 值
组间	1.25948×10^9	1	1.25948×10^9	37.78	7.90871×10^{-10}
组内	6.65407×10^{11}	19 998	3.32737×10^7	—	—
总计	6.66667×10^{11}	19 999	—	—	—

表 3 的结果仅表明在这几组数据之间存在显著性差异，需要对数据进行两两检验，确定各组之间的显著性。多组数据秩和检验后两两比较就要把各组混合排秩，对秩次进行方差分析，得出两两比较的结果^[14-15]。

为了显示各组之间检验的异常值，本研究引入了箱型图(图 2)，箱型图最大的优点是不受异常值的影响，能够准确地描绘出数据的离散分布情况，同时也利于数据的清洗。鉴于数据组合形式较多，本研究选择等级 3 和 5 作为代表组。从 5 的箱型图中可看出两组数据都呈现对称分布，等级 3 的数据有 5 个异常点，等级 5 的数据有 3 个异常点，在对异常点定位后就需在秩和检验分析时过滤掉异常值。

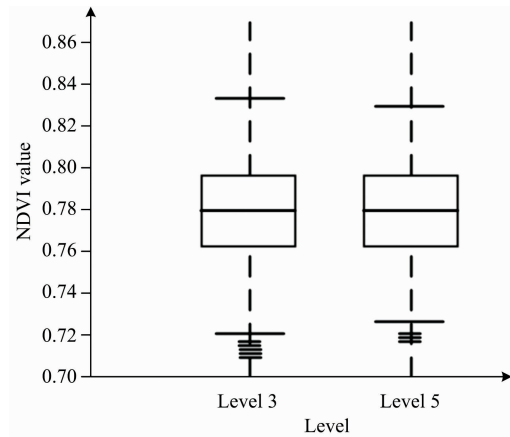


图 5 箱型图及注释

注：等级 3、等级 5 从上到下的横线分别代表样本上界、上 1/4 分位线、中位线、下 1/4 分位线、样本下界。

Fig. 5 Box drawing and notes

Note: The horizontal lines from level 3 to level 5 from top to bottom represent the sample upper bound, upper 1/4 quintile, median line, lower 1/4 quintile, and sample lower bound.

除了考虑样本间的两两比较之外，还要考虑多种因素之间的相关性，在 3 组间进行秩和检验，以便结果更加可靠。

本试验中稻曲病的发病分为 4 级, 4 组样本按照组合的方式进行秩和检验, 所需要检测的组数为 $C_4^2 + C_4^3 + C_4^4 = 11$ 组, 除去总体秩和检验的 1 组外、2 组之间和 3 组之间的组合方式共有 10 种。本次秩和检验的显著性结果如表 4 所示。

表 4 显著性秩和检验结果

Table 4 Significance rank sum test results

参与检验的等级	p 值
1, 3	3.8×10^{-9}
1, 5	4.2×10^{-11}
1, 7	5.6×10^{-7}
3, 5	7.9×10^{-10}
3, 7	4.6×10^{-6}
5, 7	3.9×10^{-9}
1, 3, 5	5.2×10^{-8}
1, 3, 7	8.7×10^{-10}
1, 5, 7	6.9×10^{-7}
3, 5, 7	3.5×10^{-11}
1, 3, 5, 7	7.3×10^{-9}

与其他假设检验的条件相似, 秩和检验的显著性水平分 $\alpha=0.05$ 和 0.01 两种临界值, $0.05 < p < 0.01$ 表示差异显著, $p < 0.01$ 表示差异极为显著。本试验中原假设是样本集之间不存在显著性差异。

根据表 4 中所显示的结果, 不同样本间秩和检验的 p 值都远 < 0.01 , 说明不同组间的样本数据存在极为显著差异, 也反映出此方法用于分类的合理性。图 6 是试验田水稻发病等级划分结果图, 本结果和大田实际调查的结果一致。图 6

中深绿色表示发病少, 红色表示发病严重, 浅绿色和浅红色则介于二者之间。

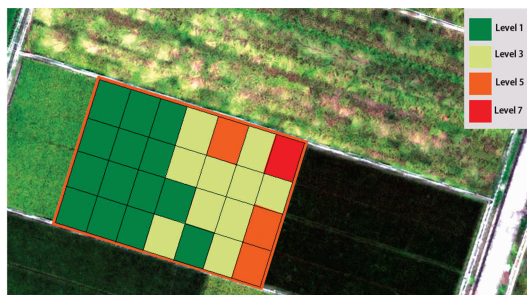


图 6 试验田稻曲病发病等级划分图

Fig. 6 Diagram of the classification of Rice False Smut in experimental field

3 结 论

尝试根据归一化植被指数确定光谱特征点, 并建立与稻曲病发病程度的相关性。首先, 实地调查水稻发病植株情况, 计算出发病指数并划分发病等级, 再用无人机搭载高光谱相机航拍试验区域。为了去除冗余数据, 将归一化植被指数 NDVI 引入稻曲病监测, 选取合理波段计算得到 NDVI 值, 并由此确定光谱中对应的特征点。最后, 对不同试验田的特征点进行秩和检验, 分析结果通过大田调查的结果验证了本研究方法的合理性。为通过高光谱和遥感技术对大面积稻曲病监测提供了一定的理论基础, 对稻曲病病情程度划分以及早期识别提供了新的思路。

References

- [1] Tobias B Hank, Katja Berger, Heike Bach, et al. Springer Netherlands, 2019, 40(3): 515.
- [2] MA De-gui, SHAO Lu-shou, GE Jing, et al(马德贵, 邵陆寿, 葛 婧, 等). Chinese Agricultural Science Bulletin(中国农学通报), 2008, 24(9): 485.
- [3] WU Shu-wen, WANG Ren-chao, CHEN Xiao-bin, et al(吴曙文, 王人潮, 陈晓斌, 等). Journal of Shanghai Jiaotong University • Agricultural Science(上海交通大学学报 • 农业科学版), 2002, 20(1): 73, 84.
- [4] LI Yue, HE Hong-chang, WANG Xiao-fei, et al(李 月, 何宏昌, 王晓飞, 等). Bulletin of Surveying and Mapping(测绘通报), 2019, (9): 13.
- [5] WANG Jing-jing, BAI Xue, DENG Xiao-qu, et al(王晶晶, 白雪, 邓晓曲, 等). Geo-Information Science(地球信息科学), 2008, 10(6): 6808.
- [6] LIU Rong-jie, ZENG Fan-rong, GE Chao-nan, et al(刘荣杰, 曾凡荣, 葛超楠, 等). Crop Research(作物研究), 2016, 30(5): 507.
- [7] LIU Jian, WEI Ya-feng, YANG Mei-ying, et al(刘 建, 魏亚凤, 杨美英, 等). Jiangsu Agricultural Sciences(江苏农业科学), 2010, (4): 53.
- [8] HUANG Yan-hua, ZHENG Fang-lin(黄艳华, 郑芳琳). Journal of Qinzhou University(钦州学院学报), 2018, 33(3): 21.
- [9] SHI Chen-zi, GUO Yu-ren, LU Bao-li(施辰子, 郭玉人, 陆保理). Journal of Shanghai Jiaotong University(上海交通大学学报), 2003, 21(2): 152.
- [10] DUAN Xiao-bin(段小斌). Journal of Agricultural Mechanization Research(农机化研究), 2020, 42(7): 208.
- [11] WANG Fu-min, HUANG Jing-feng, WANG Xiu-zhen, et al(王福民, 黄敬峰, 王秀珍, 等). Journal of Remote Sensing(遥感学报), 2008, 12(4): 626.
- [12] ZHANG Jing-cheng, YUAN Lin, WANG Ji-hua, et al(张竞成, 袁 琳, 王纪华, 等). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2012, 28(20): 1.
- [13] LUO Hong-xia, KAN Ying-bo, WANG Ling-ling, et al(罗红霞, 阚应波, 王玲玲, 等). Guangdong Agricultural Sciences(广东农业科学), 2012, 39(18): 76.

[14] TIAN Bing(田 兵). Journal of Tonghua Teachers College(通化师范学院学报), 2013, 34(10): 13, 18.

[15] WANG Jun, WU Xi(王 俊, 吴 熙). Chinese Journal of Health Statistics(中国卫生统计), 2008, 25(1): 55, 58.

Detection of Rice False Smut Grade Degree Based on the Rank Sum Test of Spectral Feature Points

SANG Jia-mao, CHEN Feng-nong*

School of Automation, Hangzhou Dianzi University, Hangzhou 310018, China

Abstract Rice smut is known as the cancer of rice. The pathogen not only affects rice yield, but also causes health risks if it attaches to food and enters the body. In this study, the normalized vegetation index (NDVI) of hyperspectral images was used to obtain feature points, and the incidence of rice smut disease was detected by the method of rank sum test. Firstly, 28 adjacent rice test areas with the same area were selected from the research base of the China National Rice Research Institute. Four farmland management methods were adopted in the area, namely natural growth and spraying with 3 different pesticides. Each management method had 7 different planting dates. The sowing dates of the plots in the adjacent experimental areas differed by 1 week before and after the plots, successively decreasing, each area planted about 500 rice plants. In the peak period of rice smut disease, the incidence of rice was first investigated on the spot, and the incidence index was obtained according to the number of incidences of rice ears per unit area. Then use the UAV-borne hyperspectral camera to shoot the test field according to the corresponding trajectory. In order to facilitate the subsequent hyperspectral image stitching, it is necessary to ensure that the aerial photography path covers the test field. According to the aerial photo coordinates, elevation information and similarity, multiple hyperspectral samples are sorted, and each hyperspectral image is stitched with high quality by the corresponding algorithm. Finally a complete hyperspectral image covering the entire test area is obtained. The normalized vegetation index that best reflects the incidence of rice smut disease is extracted from the hyperspectral image, and the feature points in the corresponding spectrum are obtained according to the index to achieve the purpose of feature dimensionality reduction. The data is cleaned with box plots to remove the feature points. Then use the cleaned feature points to perform rank sum test on the disease feature values of different rice test areas. The rank sum test is divided into two steps. The first step is to perform a rank sum test on the total sample to verify whether there is a significant difference, determine which set of samples the difference comes from; the second step is to arrange and combine the 4 sets of samples to obtain a total of 11 sets of samples to be tested in different combinations, and perform rank sum tests on the 11 sets of sample data. The significance level obtained by each group is much less than 0.01, indicating that there are extremely significant differences in sample data between different groups, and it also reflects the rationality of this method for detecting the incidence of rice smut disease. In order to show the different incidence areas, the planting areas with different incidences of rice smut are marked with different colors. Finally, the field rice incidence index was used as the control group to compare with the rank sum test results. The results showed that the rank sum test was feasible to detect the incidence of rice smut disease.

Keywords Diseases; Hyperspectral image; Spectrum analysis; Rice; Rank sum test; Disease grade; Rice False Smut

(Received Sep. 10, 2020; accepted Jan. 5, 2021)

* Corresponding author