

范泽华, 张楠楠, 喻彩丽, 等. 基于近红外光谱的骏枣病虫害与风沙损伤检测[J]. 江苏农业科学, 2018, 46(2): 154–157.
doi:10.15889/j.issn.1002-1302.2018.02.041

基于近红外光谱的骏枣病虫害与风沙损伤检测

范泽华^{1,2,3}, 张楠楠³, 喻彩丽³, 白铁成³

(1. 塔里木大学信息工程学院, 新疆阿拉尔 843300; 2. 塔里木大学现代农业工程重点实验室, 新疆阿拉尔 843300;
3. 新疆南疆农业信息化研究中心, 新疆阿拉尔 843300)

摘要:为新疆南疆地区的骏枣种植设计一种基于近红外光谱分析的无损病虫害与风沙损伤检测技术。首先, 采集大量骏枣的近红外光谱, 采用遗传算法从全部近红外光谱中选择一部分最优的光谱特征; 然后, 使用二次判别分析模型对选择的光谱特征进行训练; 最后, 通过支持向量机对骏枣进行病虫害分析; 基于新疆南疆第一师阿拉尔市十四团-数字枣园示范基地的骏枣测试本技术的性能。结果表明, 本技术对骏枣的病虫害与风沙损伤检测均具有较高的检测准确率, 并具有较高的实用价值。

关键词:骏枣种植; 近红外光谱; 无损检测; 南疆风沙; 遗传算法

中图分类号: S126 **文献标志码:** A **文章编号:** 1002-1302(2018)02-0154-03

骏枣是新疆大枣领域中一个极为重要的农产品, 为新疆带来了巨大的经济效益^[1]。新疆南疆地区的风沙较大, 对农业种植具有巨大的影响, 骏枣生长过程中也会受到风沙的影响, 导致骏枣果实受到损伤^[2]。

黑铁头病与溃疡病是骏枣主要的 2 种病害^[3], 这 2 种病害会导致骏枣果皮上产生直径 2 mm 左右的病斑, 这种病害的病斑极小, 通过人工检测或者图像识别技术难以保证较高的检测准确率。桃小食心虫是骏枣的一种虫害^[4], 会导致骏枣表面产生 1 个直径 1~2 mm 的重口, 颜色为暗色, 病虫害具有极强的传播性, 所以必须尽早准确地检测出虫害, 避免造成重大的虫害灾情与经济损失, 虫害的虫口极小, 通过人工检测或者图像识别技术难以保证较高的检测准确率。

随着计算机技术的广泛应用, 许多研究人员将计算机技术引入农业工程技术中^[5], 其中在病虫害检测方面应用较为广泛且有效的技术为图像处理与机器视觉方案^[6]。例如, 结合原彩色图像与补偿模糊神经网络的脐橙病虫害识别技术^[7], 基于图像规则推理的玉米病虫害诊断系统^[8], 基于图像分析的玉米病虫害智能化识别方法^[9]。基于图像处理的技术可以减少参与的人工, 提高病虫害检测的效率, 但是对于病虫害症状不明显的情况检测准确率较差。骏枣的病虫害症状一般仅有 2 mm 左右的病症, 南疆风沙对骏枣的损坏一般也是颗粒形的损伤, 所以, 使用图像技术来检测骏枣的病虫害与风沙损伤难以获得理想的效果。

此外, 许多研究人员采用光谱技术检测果实的病虫害。例如, 基于高光谱成像的苹果虫害检测^[10], 基于太赫兹光谱

技术的山核桃内部虫害检测^[11]。受益于诸多研究人员的成果, 光谱技术是果实无损病虫害检测问题的理想解决方案。病虫害对骏枣种植的影响巨大, 应当尽早地识别并摘除。本研究将近红外光谱应用于南疆地区骏枣的病虫害与风沙损伤检测。首先, 采集大量骏枣的近红外光谱, 采用遗传算法从全部近红外光谱中选择一部分最优的光谱特征; 然后, 使用二次判别分析模型对选择的光谱特征进行训练; 最终, 通过支持向量机对骏枣进行病虫害分析。

1 材料与方法

1.1 材料

试验地点为南疆第一师阿拉尔市十四团-数字枣园示范基地。骏枣生长过程中并未进行防虫害处理, 并且受到南疆地区的风沙影响, 有些骏枣易被风沙损伤。从骏枣树上采摘 2 kg 骏枣, 人工挑选出病虫害与风沙损坏的骏枣。

采摘的骏枣全部放置于实验室的热箱中, 将骏枣保持在 $(25 \pm 0.5)^\circ\text{C}$, 从而使得骏枣的温度与湿度平衡, 为近红外光谱数据的采集做准备。

1.2 近红外光谱采集

使用 Luminar 5030 型便携式近红外光谱仪采集骏枣的光谱, 光谱仪装备了反射后分光的光配置、1 个预校准的双光束灯与 1 个近红外探测器 (indium gallium arsenide, InGaAs) 阵列, 光谱采集的持续时间为 60 ms。对每个骏枣样本采集 3 次近红外光谱数据, 将平均值作为最终的试验数据, 进行进一步分析。通过文献[12]所述的方法采用近红外光谱仪自动测量参考光谱, 使用 SNAP! 2.03 光谱采集处理软件采集漫反射光谱。在采集近红外光谱之前, 人工选出受损骏枣与完好的骏枣, 最终将每个骏枣的光谱分类为有损光谱类与无损光谱类。

将每个骏枣的光谱数据建模为数据向量, 近红外光谱反射率作为数据向量的每个元素, 即骏枣的特征值。随机选择 50% 样本组成定标集, 剩下的 50% 样本分别组成预测集。从定标集的全部光谱提取光谱特征, 作为分类处理阶段的特征

收稿日期: 2017-04-29

基金项目: 国家自然科学基金 (编号: 61362026, 61501314); 塔里木大学-中国农业科学院农业信息研究所联合基金 (编号: NJZX201602)。

作者简介: 范泽华 (1982—), 男, 甘肃甘谷人, 硕士, 讲师, 研究方向为农业遥感、数据挖掘。E-mail: fanfanzehua@126.com。

通信作者: 白铁成, 硕士, 副教授, 研究方向为农业遥感。E-mail: 893209892@qq.com。

集合。图 1 所示为本研究基于近红外光谱分析的无损病虫害与风沙损伤检测技术的主要流程。

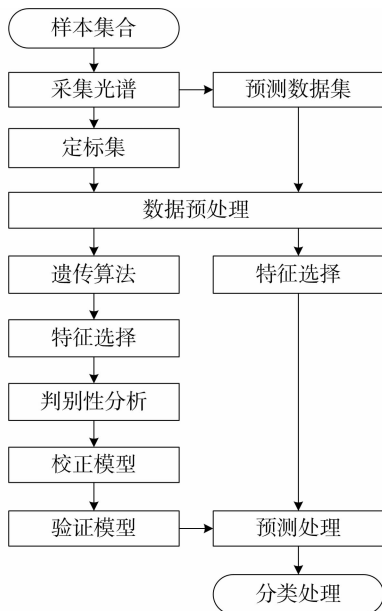


图1 基于近红外光谱分析的无损病虫害与风沙损伤检测技术的主要流程

1.3 光谱数据的分析处理

因为光谱数据之间存在高相关性,所以使用遗传算法来选择分类算法的输入特征集。遗传算法的优化目标是选择描述预测器变量与响应变量关联性的最优波段集合,即选出具有生物学意义的光谱波段。本研究使用遗传算法获得 n 个特征子集,根据文献[13] n 的范围应为 2~6,此时的特征子集是全部数据集的最优表征,为了最小化过度拟合问题,最终选择 6 个特征。

本研究考虑了 3 个分类器,即线性判别分析(LDA)、二次判别分析(QDA)、 k -近邻算法。线性判别分析与二次判别分析分类器均基于贝叶斯规则,而 k -近邻算法使用欧氏距离作为相似性度量。将所有情况的临界值设为标准值 0.5,因此,将预测结果的概率高于 0.5 的输出分类为一类,其他的输出分为另一类。线性判别分析的线性判别函数数量应当等于类别的数量减 1,因此对于 2 个类别的情况,应当生成 1 个线性判别函数。二次判别分析与线性判别分析的使用场景较为接近,一般用于类内协方差矩阵不理想的情况,对于大数据集,各类的协方差矩阵差异较大,二次判别分析的性能优于线性判别分析。线性判别分析与二次判别分析 2 个算法的时间复杂度与分类数量成比例关系,但是 k -近邻算法的时间复杂度对类别数量则不敏感。 k NN 分类器通过计算训练集中每个样本的距离,如果 1 个样本在特征空间中的 k 个最近邻样本中的大多数属于某一个类别,则该样本也属于这个类别,并具有这个类别上样本的特性。通过交叉检验方法选出最优的 k 个近邻,选出的最小 k 个近邻具有最小的平均误差。本研究最终选择二次判别分析分类器对遗传算法选出的最优特征进行分类处理。

每个判别模型的分类结果定义为误报率、漏报率与总误差,使用接受者操作特征分析(receiver operating characteristics,简称 ROC)对每个模型进行分析,绘制每个模

型的 ROC 曲线。ROC 曲线包含误报率与漏报率:

$$\text{误报率} = FP / (FP + TN); \quad (1)$$

$$\text{漏报率} = FN / (TP + FN); \quad (2)$$

$$\text{总误差率} = (FP + FN) / (TP + FN); \quad (3)$$

$$\text{敏感性} = TP / (TP + FN)。 \quad (4)$$

式中: FP 表示被模型预测为正的负样本; TN 表示被模型预测为负的负样本; FN 表示被模型预测为负的正样本; TP 表示被模型预测为正的正样本。

2 结果与分析

2.1 数据集

图 2-a 所示是有损骏枣与完好骏枣的平均二阶导数反射光谱,图 2-b、2-c、2-d 是 95% 置信度定标集中 2 个类别的光谱变化曲线。光谱的变化情况说明,骏枣质量主要与 1 300~1 420 nm、1 650~2 300 nm 波段区域的光谱特征有关,这 2 个光谱范围内的光谱反射率一般与蛋白质、脂肪、水分相关,这些成分是骏枣的主要成分。

使用遗传算法主要在 1 320、1 460、1 650、1 920 nm 与 2 000~2 220 nm 附近选择特征,1 150~1 750 nm 范围的特征代表 O—H 键对与—CH(CH₃—、—CH₂—等)键对的近红外光谱泛频。上述频带反射率的变化一般与酚(~1 320 nm)、过氧化氢化物(~1 460 nm)有关,而酚与过氧化氢化物 2 个化学成分与骏枣的破损、虫害有关,一般认为骏枣具有愈伤作用,收到撞击以及害虫破坏等会使总氧化值提高,导致骏枣的酚含量下降。

骏枣有损区域中,与脂肪族酯、烷基、烯烃关联的近红外光谱为 1 100~1 375 nm,水和乙醇相关的近红外光谱为 ~1 460 nm,与脂肪酸与甘油三酸酯相关的近红外光谱为 (~1 600~1 850 nm),1 920 nm 与 2 000~2 220 nm 的波段与 O—H 键对、C—C/C—H 键对、C—H/C=O 键对、N—H/C=O 键对相关,这些键对则与水分、油脂、蛋白质相关联。

水分、酚、蛋白质、氨基酸之间近红外光谱反射率的差异则与水分流失、氧化反应、微生物腐蚀等有关。风沙与病虫害对骏枣的内部组织产生损坏,影响了骏枣的水分、密度与硬度。而这些影响可以通过近红外光谱观察出来。

2.2 分类结果分析

选择 2 组特征(近红外光谱的波长)进行 2 组独立的试验,结果见表 1。试验 1 的误报率、漏报率、总误差分别为 1.72%、5.97%、4.00%,试验 2 的误报率、漏报率、总误差分别为 5.17%、2.99%、4.00%。

图 3 所示是二次判别分析模型的试验 1 分类结果,其中仅有 1 个有损骏枣与 4 个完好骏枣的分类发生了错误,总的分类准确率为 96%。

图 4 所示是二次判别分析模型的试验 2 分类结果,其中仅有 3 个有损骏枣与 2 个完好骏枣的分类发生了错误,总的分类准确率为 96%。

光谱数据中一般包含一些干扰因素、室内噪声以及化学物质的影响,从原数据集提取出最优的特征光谱可以降低分类模型的冗余信息,并提高分类的准确率。

3 结论与讨论

本研究探讨了近红外光谱(1 100~2 300 nm)检测病虫

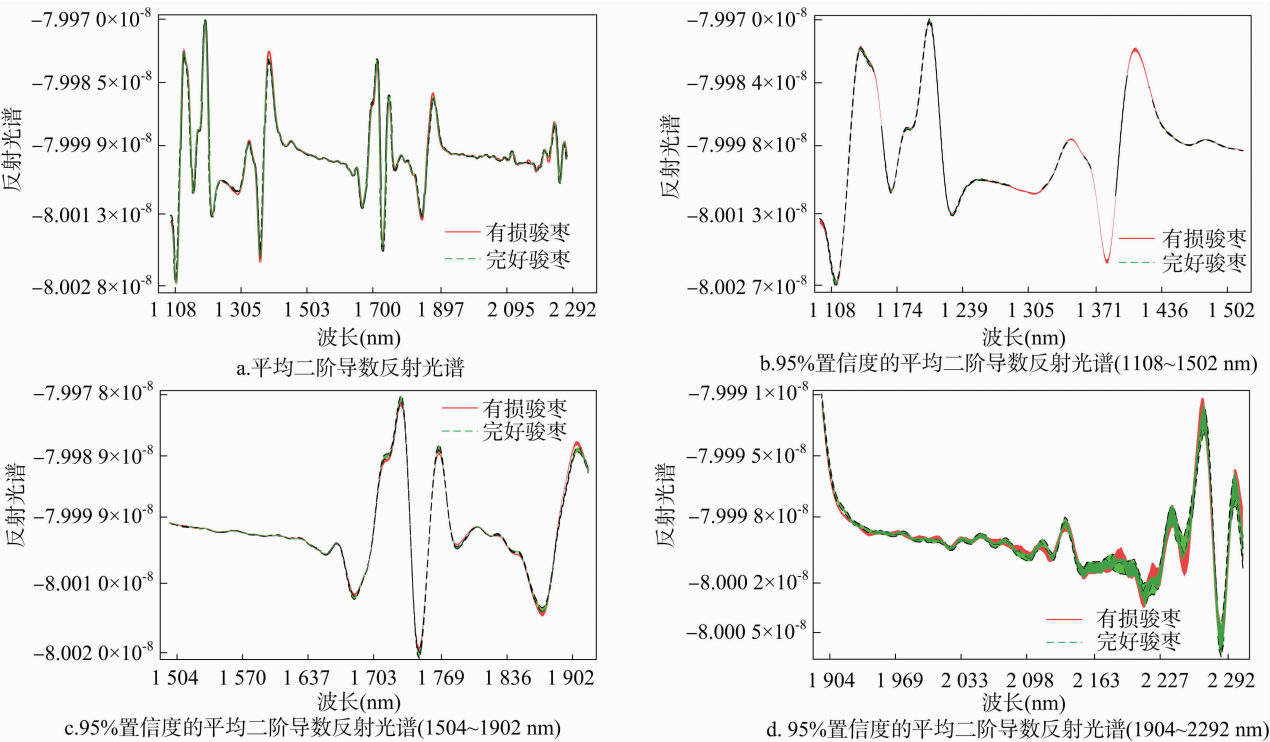


图2 不同波长的二阶导数反射光谱

表 1 二次判别分析模型对预定义光谱 5 个、6 个特征子集的判别性能

项目	试验 1	试验 2
特征波长 (nm)	1 452	1 320
	1 922	1 478
	2 088	1 648
	2 198	2 082
	2 222	2 198
	—	2 220
误差率 (%)	误报率	1.72
	漏报率	5.97
	总误差	4.00

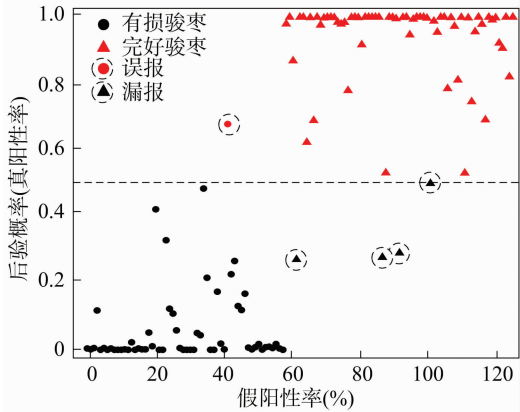


图3 二次判别分析模型的试验1分类结果

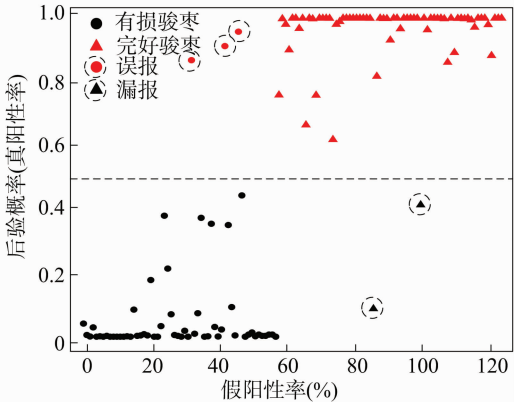


图4 二次判别分析模型的试验2分类结果

1 320、1 400~1 550、1 650、2 000~2 220 nm,这些光谱波段一般与酚、过氧化氢、烷基等关联,实时的骏枣检测系统应当从这些光谱波段选择特征。

本技术是一种自动化、高准确率、无损的骏枣病虫害损伤检测系统。未来将引入大数据集处理技术,增加近红外光谱光波特征的数量,从而提高检测的准确率与处理速度。

参考文献:

[1] 宋锋惠,哈德尔·依沙克,史彦江,等. 新疆塔里木盆地骏枣果实营养与土壤养分相关性分析[J]. 果树学报,2010,27(4):626-630.

[2] 郑素慧,葛清华,车风斌,等. 气调贮藏不同温度对新疆骏枣干枣贮藏品质及生理活性的影响[J]. 新疆农业科学,2014,51(4):620-626.

[3] 韩夫云,崔大鹏,葛喜宇. 朝阳地区大枣主要病虫害的发生特点及防治技术[J]. 现代农业科技,2015(8):144.

害与风沙导致的损坏南疆骏枣,通过 2 组近红外光谱特征子集的试验结果可看出,本模型的检测总误差小于 5%,具有较好的检测准确率。

病虫害与风沙受损骏枣检测的最优近红外光谱特征包含

沈启扬, 於 锋, 周学剑. 江苏省设施蔬菜关键生产环节机具的选型与配备研究[J]. 江苏农业科学, 2018, 46(2): 157–161.
doi:10.15889/j.issn.1002-1302.2018.02.042

江苏省设施蔬菜关键生产环节机具的选型与配备研究

沈启扬^{1,2}, 於 锋¹, 周学剑¹

(1. 江苏省农机具开发应用中心, 江苏南京 210017; 2. 南京农业大学, 江苏南京 210095)

摘要:通过在江苏多个设施农业园区进行调研,对设施蔬菜耕整地、施肥、种植、植物保护 4 个关键环节所用机具进行选型。结合专家和用户意见确定选型的主要指标,并给出重要指标的权重,通过模糊综合评判法对其进行排序,优选出关键生产环节机具。以江苏省典型设施农业园区常熟碧溪现代农业生态园作为试验对象,以作业成本最小为目标建立函数,对筛选出来的机具进行优化配备分析,以甘蓝种植工艺流程和时间表为约束,建立线性规划约束方程及目标方程,采用 MATLAB 软件进行求解得出结论,为江苏省其他设施蔬菜园区提供参考。

关键词:江苏设施蔬菜园区;设施蔬菜机具;关键生产环节;选型;配备模型;模糊综合评判法;线性规划约束方程

中图分类号: S233.74 **文献标志码:** A **文章编号:** 1002-1302(2018)02-0157-05

近年来,江苏省设施农业发展迅速,特别是设施蔬菜发展迅猛,截至 2014 年年底,全省建成永久性蔬菜基地 7.39 万 hm^2 ,设施农业面积达 75.8 万 hm^2 ,占耕地面积比重提高至 16.5%,占比居全国第一。高效设施农机化科技示范基地有 242 个,总面积达 3.7 万 hm^2 ;以省级示范基地为龙头,市县级示范基地为主体,占总基地数量的 75.2%。基地以蔬菜、经济林果(西兰花、甘蓝、草莓、葡萄、蜜桃、柿子、西瓜、金针菇、杏鲍菇、蕃茄、青椒、萝卜、甜瓜、马铃薯、药草、仙人掌、花卉、苗木)种植为主。设施蔬菜关键生产环节机具的选型与配备是设施农业当中的一个重要环节,但目前设施农业大部分生产环节仍为传统的人工劳动,工作强度大,作业效率低,人均作业面积仅为欧美国家的 1/4,且作业质量差,单位面积产量仅为欧美国家的 1/3。同时,伴随着我国劳动力日益短缺,我国对设施蔬菜机械的需求大量增长,特别是一些示范基地对设施蔬菜机械的需求增加,正确的选型与合理地进行配备以达到满足生产、节约成本、降低费用的要求成为当前急需解决的问题。目前对农机选型与配置有 2 种偏见:一

是片面追求数量,以农机装备的拥有量和总动力来衡量农机化发展水平,缺乏最佳配置研究;二是片面追求高档,以农机装备的“高、新、特”等来衡量农机化水平,缺乏农机经济研究。随着新型农机具更新换代步伐的加快,农机具的选型和配备问题在农机化技术推广领域中将占有越来越重要的地位,相关选型与配备问题的决策模型在理论和实践上都有着重大的意义。

1 设施蔬菜关键生产环节机具选型

1.1 选型方法及程序

设施蔬菜机具选型是后期机具优化配备的基础,本研究结合专家意见确定选型的主要指标,并给出重要指标的权重,通过模糊综合评判法对其进行排序,优选出耕整地机具 3 种进行优劣排序。根据对设施蔬菜定园区作业规模、作业模式等信息的调研,确定备选机型,运用调查评测法对备选机型进行综合评价^[1],再通过试验获取相关数据分析运算,最后选出综合性能最好的机具型号。选型程序如图 1 所示^[2]。

1.2 评价指标及权重

根据行业标准 NY/T 1931—2010《农业机械先进性评价一般方法》相关内容,评价机具的综合性能,通常是对其适用性、经济性、操作维护保养方便性和人机关系进行综合分析。本研究以作业质量、作业效率、维护保养方便性、操作方便性、

收稿日期:2017-04-05

基金项目:江苏省农业科技自主创新资金[编号:CX(15)1039]。

作者简介:沈启扬(1984—),男,江苏启东人,硕士,工程师,主要从事设施农业机械化试验、示范、推广研究。E-mail:13913990904@163.com。

[4]王燕平,王文平. 关于枣树病虫害防治技术的几点思考[J]. 黑龙江科技信息, 2016(34): 254.

[5]汪京京,张 武,刘连忠,等. 农作物病虫害图像识别技术的研究综述[J]. 计算机工程与科学, 2014, 36(7): 1363–1370.

[6]姚 青,张 超,王 正,等. 分布式移动农业病虫害图像采集与诊断系统设计与试验[J]. 农业工程学报, 2017, 33(增刊 1): 184–191.

[7]温芝元,曹乐平. 基于补偿模糊神经网络的脐橙不同病虫害图像识别[J]. 农业工程学报, 2012, 28(11): 152–157.

[8]刘同海,黄斌博,李少昆,等. 基于图像规则推理的玉米病虫害草害诊断系统的设计[J]. 中国农业大学学报, 2012, 17(4): 154–158.

[9]杨 青. 一种基于图像分析的玉米病虫害智能化识别方法[J].

江苏农业科学, 2014, 42(7): 128–130.

[10]田有文,程 怡,王小奇,等. 基于高光谱成像的苹果虫害检测特征向量的选取[J]. 农业工程学报, 2014, 30(12): 132–139.

[11]李 斌,张伟立,赵春江,等. 基于太赫兹光谱技术的山核桃内部虫害检测初步研究[J]. 光谱学与光谱分析, 2014, 34(5): 1196–1200.

[12]Cayuela J A, Weiland C. Intact orange quality prediction with two portable NIR spectrometers[J]. Postharvest Biology & Technology, 2010, 58(2): 113–120.

[13]Brusco M J, Singh R, Steinley D. Variable neighborhood search heuristics for selecting a subset of variables in principal component analysis[J]. Psychometrika, 2009, 74(4): 705–726.