

## 基于叶片纹理特征的棉花蚜害诊断模型研究

许敬诚, 吕新\*, 林皎, 张泽, 姚秋双, 范向龙, 洪延宏

(石河子大学农学院 / 新疆兵团绿洲生态农业重点实验室, 新疆 石河子 832003)

**摘要:**【目的】为探究棉花叶片蚜害快速识别的可行性,本研究对健康和受棉蚜为害的棉花叶片的高光谱图像进行识别。【方法】以新陆早 45 号为研究对象,通过获取健康和受棉蚜为害棉花叶片的高光谱图像,提取不同处理下棉花叶片的感兴趣区域光谱图像信息,并采用 3 种降维手段获取高光谱特征。利用灰度共生矩阵(Gray-Level Co-occurrence Matrix, GLCM)提取图像的纹理特征,构建棉蚜为害诊断模型。【结果】采用全纹理特征数据结合随机蛙跳-偏最小二乘线性判别分析模型(RF-PLS-LDA)建模得到的预测集识别率为 91.49%;以能量(Energy)作为输入,建立主成分载荷-偏最小二乘线性判别分析模型(PCA-Loading-PLS-LDA),对预测集识别率达到 92.55%。【结论】以灰度共生矩阵二阶统计量能量建模可有效地简化模型,减少计算量,提高预测的稳定性。基于纹理特征向量能有效地实现蚜害棉花叶片的识别,为虫情的快速识别提供技术支持。

**关键词:**棉花;棉蚜;高光谱成像;纹理特征;灰度共生矩阵

## The Diagnostic Model of Cotton Aphids Based on Leaf Textural Features

Xu Jingcheng, Lü Xin\*, Lin Jiao, Zhang Ze, Yao Qiushuang, Fan Xianglong, Hong Yanhong

(Agriculture College of Shihezi University/The Key Laboratory of Oasis Eco-Agriculture, Xinjiang Production and Construction Group, Shihezi, Xinjiang 832003, China)

**Abstract:** [Objective] The aim of this research is to explore the possibility of rapid identification of aphid damage to cotton leaves by identifying the hyperspectral image about healthy and aphid damage cotton leaves. [Method] Taking Xinluzao 45 as material, hyperspectral images of healthy and damaged cotton leaves by aphids were obtained, and hyperspectral image information was extracted from interested regions of different cotton samples. Then three descending dimension methods were used to extract hyperspectral feature, and Gray-Level Co-occurrence Matrix to extract image texture feature. Finally a cotton aphids damage diagnostic model was built up. [Result] The prediction accuracy based on all textural samples input Random frog-partial least-square-linear discriminant Function (RF-PLS-LDA) model was 91.49%. The prediction accuracy based on energy data input principal component analysis-loading partial least-square-linear discriminant function (PCA-Loading-PLS-LDA) model was 92.55%. [Conclusion] The second-order statistics (energy) of gray co-occurrence matrix can be used to simplify the model, reduce the computation and improve the stability of prediction. Based on the texture feature vector, the identification of aphid cotton leaves can be realized effectively, which provides a method for the rapid identification of insect pests.

**Keywords:** cotton; cotton aphid; hyperspectral imaging; textural features; Gray-Level Co-occurrence Matrix

随着转苏云金芽孢杆菌(*Bacillus thuringiensis*, Bt)基因抗虫棉的种植,棉铃虫种群增长被有效控制,但棉蚜成为新疆棉田主要害虫之一<sup>[1-2]</sup>。棉蚜多群居于棉花叶背,受到棉蚜危害的棉叶片失水变形,并出现褶皱,严重时向下卷曲<sup>[3]</sup>;叶背部棉蚜代谢的糖分滴落在叶表面阻碍植株的光合作用和呼吸作用,影响植株的生长发育。传统

棉蚜信息监测需要人工观察受害棉花叶片变化,统计棉蚜数量,费时费力,并且由于统计的延时性造成虫情信息滞后,不利于棉蚜危害精准防控。农田传感器的普及使棉蚜信息快速获取成为可能。机器视觉技术因其识别速度快而被广泛应用于虫情监测,Ghyar 等利用灰度共生矩阵和颜色矩阵分离叶片病变区域,实现水稻的病虫害的

收稿日期:2019-11-18 第一作者简介:许敬诚(1995—),男,硕士,xjchhh@sina.com。\*通信作者:lxshz@126.com

基金项目:新疆生产建设兵团重大科技项目(2018AA004);中央引导地方科技发展专项资金(2019BT0826)

机器识别<sup>[4]</sup>。基于 LAB 颜色空间, Madhuri 等通过提取目标对象的纹理特征构建支持向量机模型完成田间害虫的分类<sup>[5]</sup>。高光谱成像技术同时集成了机器视觉技术和高光谱技术特点,能同时反映被测样品外观形态与内部结构等特性,作为一种快速、无损的识别方法在植物虫害的症状检测方面具有很大优势。李震等通过基于叶绿素敏感波长建立叶绿素含量预测模型检测柑橘红蜘蛛为害<sup>[6]</sup>。Wu 等基于偏最小二乘判别分析法和反向传播神经网络对菜心叶片高光谱成像影像上幼虫部分进行识别,在近红外波段可以有效提取出幼虫所在区域<sup>[7]</sup>。Cao 等通过主成分分析法选择 505、659 和 955 nm 特征波长,以反向传播神经网络模型实现粮仓害虫快速的识别<sup>[8]</sup>。Huang 等采用连续投影法和竞争性自适应重加权抽样法选择特征波长,建立偏最小二乘判别分析和最小二乘支持向量机模型完成桑螟幼虫为害等级的快速诊断<sup>[9]</sup>。所以,基于高光谱成像技术,可以多角度对虫害进行检测,指导农作物的植保工作。

现阶段纹理特征获取多基于 RGB 图像<sup>[10-12]</sup>。高光谱成像技术因其光谱分辨率高、图像信息更为丰富,应用于提取目标特征是虫害监测的新方向之一<sup>[13-14]</sup>。本研究以受棉蚜为害的棉花叶片为研究对象,采集健康与受棉蚜为害的棉花叶片正面高光谱图像,通过不同光谱信息降维法提取特征波长,采用灰度共生矩阵方法提取特征波长下灰度图像的纹理特征,以纹理特征向量为输入建立受棉蚜为害棉花叶片的判别模型,以期为棉花蚜害快速监测与精准农业靶向施药提供技术支持。

## 1 材料与方法

### 1.1 材料

本试验于 2019 年 7 月在新疆维吾尔自治区石河子市石河子大学农学院试验场进行。试验田种植的棉花品种为新陆早 45 号,依照新疆传统高密度膜下滴灌栽培模式种植,采用常规水肥管理模式,不同处理的棉花植株分别于 2 顶孔径 0.075 mm(200 目)封闭防虫网帐篷中生长。

### 1.2 棉叶样本采集

在棉花盛花期采集叶片,上午以活体形式采

集,用果枝剪沿叶柄基部剪下叶片,除去寄生在叶背的棉蚜,并轻拭棉花叶片正面擦除杂质,随后装入单独的自封袋中编号,迅速放入装有生物冰袋的冰盒中冷藏保存,带回室内获取高光谱影像数据。本研究共采集 142 片棉花上部叶片,其中健康叶片 71 片,蚜虫为害叶片 71 片。

### 1.3 高光谱图像数据采集

高光谱影像数据采集采用 SOC710VP 可见光 - 近红外地物高光谱成像仪 (Surface Optics Corporation, 美国),光谱波段范围 351~1 044 nm,光谱分辨率可设置。本研究中,光谱分辨率设定为 1.3 nm,光谱波段数 128。采集暗室内置带刻度升降平台控制样本与相机之间的距离,光源为 2 盏 75 W 卤素灯(OSRAM, 德国),光线直射暗室内壁漫反射于待测物体上。在高光谱成像信息采集前通过对相机曝光度、焦距、光圈、扫描速率的调整,确保采集到的图像纹理清晰、不失真、无形变,仪器和样品不会被卤素灯高温损伤。通过尝试确定高光谱成像仪的参数分别设置如下:样品距离镜头 64 cm,光圈 1.4,曝光时间 19 ms,扫描速率 150~200 帧·s<sup>-1</sup>,图像分辨率为 696×520 像素,光源焦点在镜头与拍摄叶片二分之一处,如图 1 所示。

将新鲜棉花叶片依次置于底面垫有低反射率黑色背景板的高光谱成像暗室中,同时在距叶片边缘 2 cm 处放置标准灰板辅助黑白校正。SOC710VP 可见光 - 近红外地物高光谱成像仪获取的光谱图像原始数据是像元亮度值(Digital number, DN),通过自带软件 SRAnal710 进行光谱标定、空间辐射与光谱辐射标定可以将像元亮度值转化成带光谱反射率。全波段光谱首尾存在噪声,且光谱影像有椒盐噪声。去除噪声明显的光谱波段,对 400~840 nm 光谱采用多项式平滑法(Savitzky-Golay smoothing, SG)去噪。避开棉花叶片主叶脉所在区域,于单片棉花叶片上提取 2 处面积为 50×50 像素纹理清晰的感兴趣区域(Region of interest, ROI)作为样本,共得 284 个样本。随机划分 67%的样本作为建模集,其余部分作为预测集。

### 1.4 特征波长提取方法

全波段光谱数据的模型存在大量冗余数据

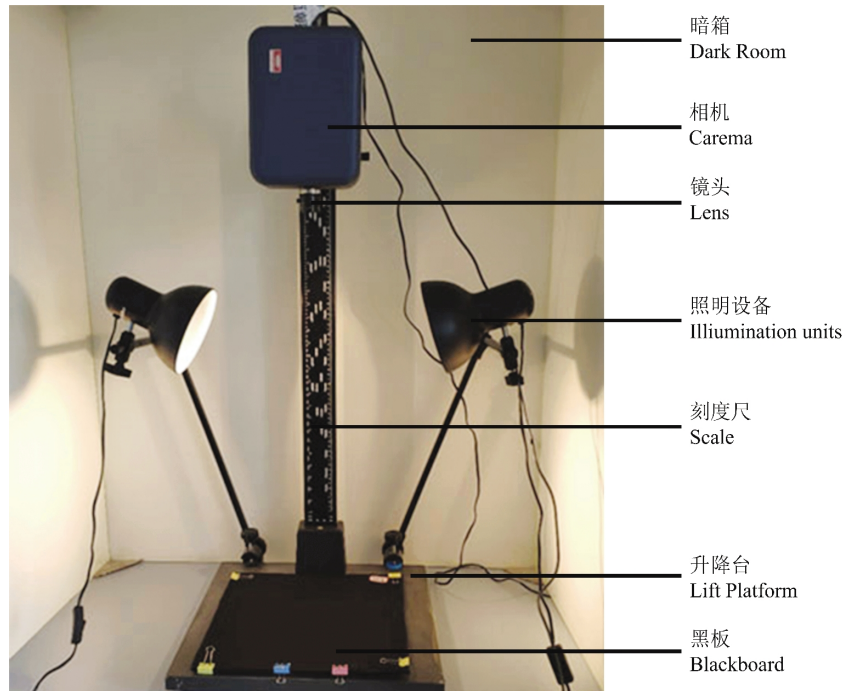


图 1 高光谱图像采集平台

Fig. 1 Hyperspectral image acquisition platform

和同质性数据影响判别精度,因此为了压缩数据量,降低模型复杂度,本研究采用主成分载荷法(Principal component analysis-Loading, PCA-Loading)<sup>[15]</sup>、随机蛙跳算法(Random frog, RF)<sup>[16]</sup>和连续投影法(Successive projections algorithm, SPA)<sup>[17]</sup>选择与受棉蚜为害后棉花叶片变化相关性更强的特征波长。

主成分载荷法根据主成分分析的结果,计算出每个波长的载荷,反映主成分对每个变量的相关性。较大的正或负载荷意味着该变量对于样本主成分具有重要的影响,对应的变量相对较为重要。

随机蛙跳算法是衡量变量重要性的一种方法,通过建立变量的正态分布空间并预选初始变量集,将初始变量集中的变量通过不断选择选入候选子集,候选子集中的变量随着迭代而不断变化,统计每次迭代各变量出现的频率,子集被选频率越高说明其越重要,依据样本频率排名选择变量序号对应的波段为特征波长。

连续投影法主要解决共线性问题,能从信息变量中选择冗余信息最少的变量,提取共线性最小的变量,将各个波长组合分别建立子集,每个子集逐一采用多元线性回归法计算均方根误差

(Root mean square error, RMSE)值,选择平稳且最小的 RMSE 值对应子集作为特征波长子集。

### 1.5 纹理特征提取

图像视觉特征提取方法主要有颜色、形状和纹理提取 3 类。健康棉花叶片光滑平展,受棉蚜为害的棉花叶片皱缩卷曲,叶表面纹理特征变化明显。本研究主要采用灰度共生矩阵(GLCM)<sup>[18]</sup>算法进行纹理特征提取,在提取纹理特征时,以 0°、45°、90°、135°作为常用方位角,通过计算 2 个像素之间的相近关系的概率来获得能体现物体特征的二阶统计量。灰度共生矩阵计算所获得的无量纲参数数量较大,不宜直接作为纹理特征,通常基于其构建统计量,结合方位角建立纹理特征向量。常用的 4 个不相关的纹理特征有能量(Energy)、熵(Entropy)、对比度(Contrast)、相关性(Correlation)<sup>[19]</sup>。能量反映图像灰度粗细分布的均匀程度,图像的灰度分布越均匀,能量值越大;熵反映图像灰度值散乱度,灰度值分布随机度越高熵值越大;对比度反映图像强度差异,图像灰度值差异大时,对比度大;相关性反映图像纹理在行或列上一致性,图像横向或纵向纹理均一时,相关性值增大。本研究中,基于特征波长选择方法选择的特征波长下的感兴趣区域灰度图



像进行纹理特征提取。

### 1.6 判别分析方法

本试验采用 2 种分类方法建立模型:线性判别分析算法偏最小二乘线性判别分析(Partial least-square-linear discriminant function, PLS-LDA)模型<sup>[20]</sup>和非线性判别分析算法支持向量机(Support vector machine, SVM)模型<sup>[21]</sup>。

PLS-LDA 模型是一种线性分类方法,它基于 PLS 回归模型预测每个样本的类别数,选取平方和预测误差最小值处的最小潜在变量数(Latent variables, LVs)。本研究采用去中心化法对数据进行归一化处理,通过 10 000 次蒙特卡洛验证获取建模集的最佳潜在变量值 LVs,通过多次寻优选择最佳阈值来为分类后的样本归类。

构建支持向量机(SVM)模型需要对 2 个参数寻优:核函数参数和模型的惩罚系数。由于选择合适的核函数没有系统的方法,而本试验的样本呈现非线性分布,径向基(Radial basis function, RBF)核函数能够更好处理非线性关系并降低训练过程计算的复杂程度。本试验选择径向基核函数作为 SVM 的核函数构建 RBF-SVM 模型,随后通过蒙特卡洛交叉验证对惩罚参数  $c$  和样本影响半径的倒数  $g$  寻优, $c$ 、 $g$  参数的变化范围设为  $-5 \sim 5$ ,基于多次交叉验证后的结果选择最优的  $c$  和  $g$  的组合。

## 2 结果与分析

### 2.1 棉花叶片的原始光谱曲线特征

试验获取的可见光-近红外光谱波长范围为  $351 \sim 1\,044$  nm,共 128 个波段。由于试验环境、仪器的影响以及暗电流干扰,造成获取光谱前后端有明显噪声,因此剔除噪声明显的波段,截取波长  $400 \sim 840$  nm 范围共 86 个波段光谱影像用于后续建模分析,健康棉花叶片和棉蚜为害棉花叶片的平均光谱如图 2 所示。

棉花叶片高光谱特征在  $415$  nm 有一个明显的吸收谷,绿光波段开始出现反射峰,在  $557$  nm 达到最高值;红光波段出现吸收谷, $677$  nm 为最低值。

### 2.2 主成分分析

主成分分析法可以将原可能相关的变量正

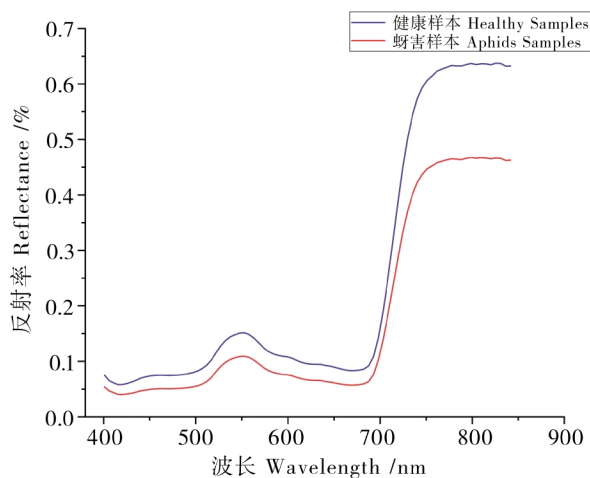


图 2 切除两端噪声后的平均光谱

Fig. 2 The mean spectra after removing both two ends noise

交变换为更多不相关的变量,显示数据的内部结构。利用主成分分析法对  $401 \sim 842$  nm 波段棉花叶片光谱数据进行分析,将 86 个波段的信号转化为若干个主成分(Principal component, PC),得到前 3 个主成分第一主成分 PC1(96.16%)、第二主成分 PC2(2.88%)、第三主成分 PC3(0.46%),累计贡献率达到 99.5%,能解释大部分变量。PC1、PC2 和 PC3 三维得分分布图如图 3 所示,健康样本和棉蚜为害样本之间存在非常明显的聚类和区分,同时可以发现,2 种样本在三维空间中存在一定的重合,需要进一步对光谱数据进行分析。

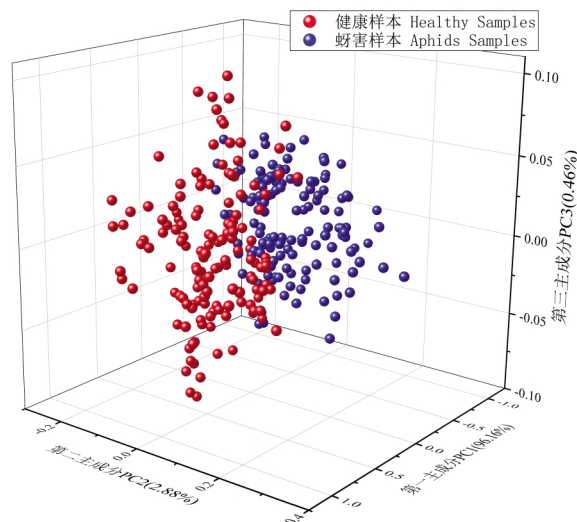


图 3 主成分聚类分布图

Fig. 3 Cluster plots based on the PCs

### 2.3 特征波长选择

选择特征波长对于去除高维数据中的冗余信息,优化校准模型,获得良好的结果具有重要意义。本试验利用主成分载荷法(PCA-Loading)、

随机蛙跳算法(RF)和连续投影法(SPA)提取特征波长(Characteristic wavelengths, CWs),3种算法分别筛选9、10、13个特征波长,如图4和表1表示。

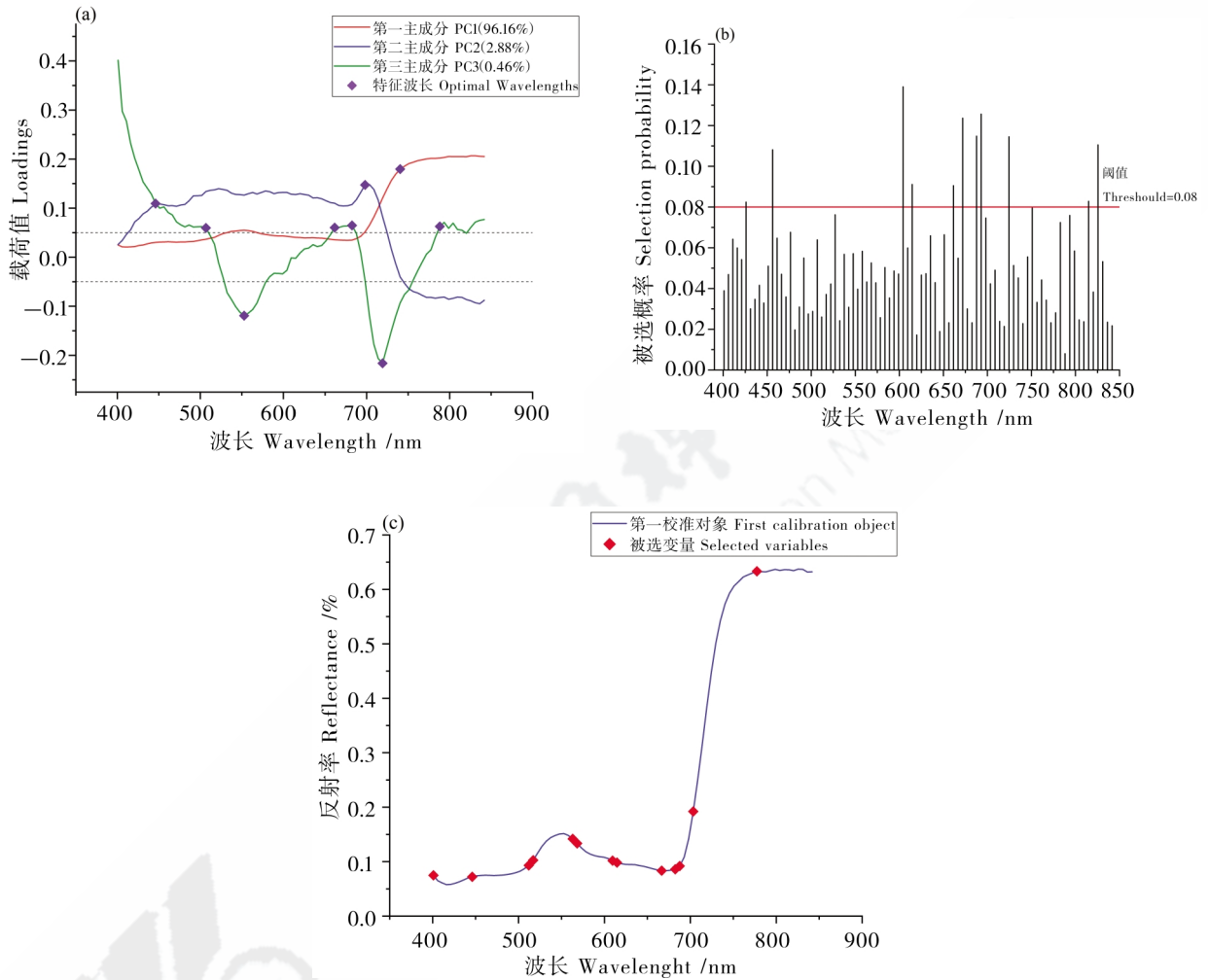


图4 PCA-Loading (a), RF (b), SPA (c)法提取特征波长示意

Fig. 4 The optimal wavelengths selection by PCA-loading (a), RF(b) and SPA (c)

PCA-Loading 法选取主成分载荷图曲线绝对值大于 0.05 的波峰和波谷作为特征波长点,共筛选出 9 个特征波长,占总波长数的 10.7%。RF 法设置算法迭代次数  $N$  为 10 000 次,提取数量 10 个,蛙跳初始种群数目  $Q$  为 2 个,以每个光谱被选择的可能性为筛选依据,运行结果为降序排列的被选择可能性,设定被选择可能性阈值为 0.829,共筛选出 10 个特征波长,占波长总数的 12.8%。SPA 法以全光谱范围的波长依次对应每个变量,设置被提取的最小变量个数为 5,最大变量个数为 30,以寻找最小共线性变量组,最终共

筛选出 13 个特征波长, RMSE 值为 0.001。3 种算法筛选出的特征波长主要集中在绿峰波段和红边波段。

### 2.4 图像纹理特征向量提取

健康的棉花叶片正面表面平展,由于棉蚜从棉花叶片背部吸食汁液,影响养分供给,导致受到棉蚜为害的棉花叶片表面与叶肉组织结构遭到破坏,出现皱缩,分支叶脉向下凹陷形成阴影,与向上突起的叶肉形成明显灰度差。在纹理特征感兴趣区域选择上,由于主叶脉灰度均一性较高,与叶肉有明显差异,且阴影面积大,不利于纹

表 1 对光谱样本提取的特征波长

Table 1 The effective wavelengths selection for hyperspectral sample

方法 Methods	特征波长数量 Number of characteristic wavelengths	特征波长 Characteristic wavelengths /nm
PCA-Loading	9	446, 507, 553, 661, 682, 698, 719, 740, 788
RF	10	456, 604, 615, 661, 672, 688, 693, 724, 815, 826
SPA	13	401, 431, 507, 517, 563, 568, 609, 615, 667, 682, 688, 703, 777

理特征的提取,因此对感兴趣区域的提取要避免叶片主叶脉位置。现从特征波长对应灰度图像中选取  $50 \times 50$  像素无主叶脉区域作为感兴趣区域。试验中将原始图片 256 等级的灰度值压缩到

16 级以减小计算量,以 4 个方向  $0^\circ$ 、 $45^\circ$ 、 $90^\circ$ 、 $135^\circ$ ,固定取样距离为 1 像素构建 4 个方向的灰度共生矩阵,分别统计各个矩阵能量、熵、对比度、相关性,如图 5 和表 2。

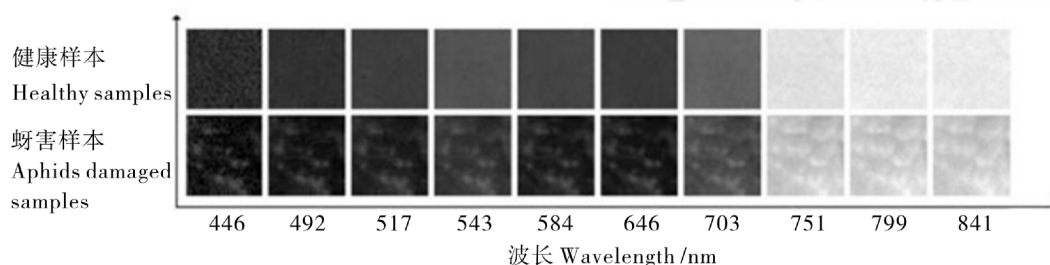


图 5 不同波长健康样本与为害样本对比

Fig. 5 The contrast of healthy and aphid cotton leaf's ROI in different wavelengths

表 2 不同样本纹理特征向量参数

Table 2 The average feature vector of GLCM for two types of leaves

指标 Index	样本类型 Sample type	方向角度 Direction				均值 Mean	标准差 Standard deviation
		$0^\circ$	$45^\circ$	$90^\circ$	$135^\circ$		
能量 Energy	健康样本 Healthy Sample	0.148	0.127	0.151	0.127	0.138	0.014
	为害样本 Aphid Sample	0.059	0.045	0.057	0.045	0.051	0.008
熵 Entropy	健康样本 Healthy Sample	2.876	3.091	2.889	3.090	2.987	0.126
	为害样本 Aphid Sample	3.591	3.855	3.649	3.854	3.737	0.145
对比度 Contrast	健康样本 Healthy Sample	8.166	12.977	8.563	12.951	10.664	2.747
	为害样本 Aphid Sample	6.490	10.246	6.918	10.201	8.464	2.139
相关性 Correlation	健康样本 Healthy Sample	0.042	0.036	0.042	0.036	0.039	0.060
	为害样本 Aphid Sample	0.041	0.036	0.040	0.036	0.038	0.064

根据不同波长下灰度图像可知,健康叶片表面平滑,分支叶脉的纹理在可见光波段不能辨认,在近红外波段能看到少量较粗的部分,单张图像中纹理分布均匀,灰度变化小。为害叶片可见光波段和近红外波段均能看见清晰的分支叶脉轮廓并由边缘始形成向心突起,纹理呈斜向分布,向心突起处有亮斑,与分支叶脉处形成明显

灰度差。

通过 4 个角度的灰度共生矩阵计算得到 4 个二阶统计量,包括能量、熵、对比度、相关性,共得到 16 组特征向量。分别计算不同处理样本的特征向量,得到 2 组样本基于这 4 个角度的纹理特征平均值和标准差。健康样本的能量、对比度均超过为害样本,熵低于为害样本,由于为害样

本纹理分布散乱,垂直方向和斜向灰度变化杂乱,导致相关性差异不显著。4个角度的特征向量及均值、标准差有显著差异,可以解释叶片受棉蚜为害后的变化。由于纹理特征标准差数据分布接近,因此采用特征波长图像的纹理特征标准差作为输入,建立了棉蚜识别模型。

## 2.5 判别分析模型建立

分别基于3种特征波长选择方法选择的特征波长下的灰度图像获取的纹理特征,建立PLS-LDA和RBF-SVM模型,其判别分析结果如表3所示。

表3 基于纹理特征的PLS-DA和RBF-SVM判别分析模型判别结果  
Table 3 The discriminant results of PLS-DA and RBF-SVM models using texture features

方法 Method	输入数据		建模集		预测集	
	Data Set(All data, energy, entropy, contrast, correlation)	最优模型参数 Best parameter	Calibrations set		Predictions set	
			识别数/总数 Counts/total	识别率 Accuracy /%	识别数/总数 Counts/total	识别率 Accuracy /%
PCA-Loading-PLS-LDA	全特征集 All data	Best LVs = 10	170/190	89.47	83/94	88.30
	能量 Energy	Best LVs = 6	173/190	91.05	86/94	91.49
	熵 Entropy	Best LVs = 2	154/190	81.05	71/94	75.53
	对比度 Contrast	Best LVs = 3	158/190	83.16	77/94	81.91
	相关性 Correlation	Best LVs = 2	145/190	76.32	75/94	79.79
RF-PLS-LDA	全特征集 All data	Best LVs = 10	180/190	94.74	86/94	91.49
	能量 Energy	Best LVs = 4	167/190	87.89	79/94	84.04
	熵 Entropy	Best LVs = 1	151/190	79.47	78/94	82.98
	对比度 Contrast	Best LVs = 2	155/190	81.58	84/94	89.36
	相关性 Correlation	Best LVs = 3	158/190	83.16	73/94	77.66
SPA-PLS-LDA	全特征集 All data	Best LVs = 12	172/190	90.53	83/94	88.30
	能量 Energy	Best LVs = 9	167/190	87.89	87/94	92.55
	熵 Entropy	Best LVs = 2	148/190	77.89	71/94	75.53
	对比度 Contrast	Best LVs = 3	156/190	82.11	79/94	84.04
	相关性 Correlation	Best LVs = 2	161/190	84.74	72/94	76.60
PCA-Loading-RBF-SVM	全特征集 All data	c = 3.249      g = 0.467	184/190	96.84	77/94	81.91
	能量 Energy	c = 5.66      g = 32	163/190	85.79	83/94	88.30
	熵 Entropy	c = 1.62      g = 14.93	156/190	82.11	78/94	82.98
	对比度 Contrast	c = 0.23      g = 0.76	171/190	90.00	78/94	82.98
	相关性 Correlation	c = 18.38      g = 32	147/190	77.37	77/94	81.91
RF-RBF-SVM	全特征集 All data	c = 21.112      g = 0.102	184/190	96.84	70/94	74.47
	能量 Energy	c = 32      g = 32	165/190	86.84	82/94	87.23
	熵 Entropy	c = 0.27      g = 29.86	156/190	82.11	76/94	80.85
	对比度 Contrast	c = 5.28      g = 0.04	169/190	88.95	78/94	82.98
	相关性 Correlation	c = 11.31      g = 25.99	160/190	84.21	75/94	79.79
SPA-RBF-SVM	全特征集 All data	c = 1.231      g = 0.616	183/190	96.32	78/94	82.98
	能量 Energy	c = 2.64      g = 32	157/190	82.63	85/94	90.43
	熵 Entropy	c = 3.03      g = 19.70	159/190	83.68	77/94	81.91
	对比度 Contrast	c = 14.93      g = 0.05	176/190	92.63	76/94	80.85
	相关性 Correlation	c = 10.56      g = 32	158/190	83.16	76/94	80.85



比较模型分析效果,使用全部纹理特征数据集进行分析,SPA法提取的特征波长组识别率最高,2种模型识别率在82.98%以上,RF-PLS-LDA模型效果最好,预测集识别率达到91.49%。结果显示,使用全部纹理特征数据集建模识别效果较好,但输入数据量较大,模型计算时间效率较低。逐一以纹理特征二阶统计量作为输入建立判别模型,2类模型预测集的识别精度均在74%以上,可以实现对棉花上棉蚜为害叶片的精确识别。以能量作为输入的模型对棉蚜为害植物的识别效果最佳,预测集平均识别率达到89%;其次是对比度作为输入的模型,预测集平均识别率达到84%。由于相关性纹理特征差异不显著,识别率最低。其中,最优模型为SPA-PLS-LDA模型,预测集识别率达到92.55%,其次是PCA-Loading-PLS-LDA-Energy模型(91.49%)和SPA-RBF-SVM-Energy模型(90.43%)。

### 3 讨论

基于高光谱技术的作物监测模型主要以作物反射率作为模型输入数据<sup>[22-23]</sup>,但作物在不同胁迫环境生长下形态、颜色等特征会发生变化,出现“同物异谱”或“异物同谱”现象。高光谱成像技术可以同时获取光谱信息和图像信息,避免此类现象发生。因此本研究使用高光谱成像仪进行健康、棉蚜为害叶片正面高光谱成像数据采集,通过不同方法提取特征图像。基于灰度共生矩阵提取图像纹理特征,依此建立蚜害判别模型,证明使用高光谱成像对受棉蚜为害棉花叶片进行分类预测的可行性,为基于多光谱成像的棉蚜监测装置提供技术支持。

棉花受到蚜虫为害后光谱特征上会出现变化<sup>[24]</sup>。本试验中受蚜虫为害的叶片蓝光、红光所在波段与近红外波段吸收率下降,而植株光谱红、蓝光波段反射率降低是由于叶绿素含量降低所引起,近红外的反射率下降的原因是植株细胞结构变化所产生,这也从光谱角度证实了蚜虫为害会使叶片叶绿素和相对含水量下降,细胞结构出现变化<sup>[25-28]</sup>。纹理特征上,蚜虫为害棉花叶片会导致其出现皱缩、卷曲,叶片局部形成凸起、向下卷

曲等症状。本研究利用灰度共生矩阵计算特征图像灰度值的变化,经由得出的二阶统计量可以直观地反映图像部分特征<sup>[29-30]</sup>。纹理特征参数中健康样本的能量对比度超过蚜害样本,熵低于蚜害样本。由于健康叶片表面平滑,图像灰度分布均匀,因此表面灰度值相近,反映图像元灰度平方和的能量较高;而蚜虫为害叶片表面有明显由褶皱、蚜蜜造成的阴影和亮斑,因此图像的灰度值差异性较大,反映图像无序性的熵较大。

为精确蚜虫为害后光谱反射率变化的敏感区域,降低数据冗余度,使用主成分载荷法、随机蛙跳算法和连续投影法提取特征波长,从全波长信息中分别优选出9、10、13个特征波长。以特征波长高光谱图像作为分析对象,利用灰度共生矩阵提取图像的纹理信息,并取其二阶统计量的标准差作为模型的输入,最终通过偏最小二乘判别分析法和径向基支持向量机建立分析判别模型。通过对比试验发现,RF-PLS-LDA模型效果最好,预测集识别率达到91.49%。由于模型输入数据量较大,导致建模效率低,通过减少模型输入数据对模型进行优化。逐一以纹理特征二阶统计量作为输入建立判别模型,其中PCA-Loading-PLS-LDA-Energy模型性能最好,预测集识别率达到92.55%。该方法为棉花棉蚜虫情识别提供了新的思路,但其预测性能还须进一步改进。下一步的研究重点为找寻不同蚜害等级下纹理特征的变化特点。

### 4 结论

由于棉蚜为害导致棉花叶片叶绿素和相对含水量下降、细胞结构产生变化,因此棉蚜为害样本高光谱曲线蓝光、红光波段以及近红外波段的反射率较健康样本更低。通过对比试验发现,RF-PLS-LDA模型对棉蚜为害叶片的识别率最好,预测集识别率达到91.49%。由于模型输入数据量较大,导致建模效率低,通过减少模型输入数据对模型进行优化。逐一以纹理特征二阶统计量作为输入建立判别模型,其中PCA-Loading-PLS-LDA-Energy模型性能最好,预测集识别率达到92.55%。



## 参考文献:

- [1] 芦屹, 王佩玲, 刘冰, 等. 新疆棉花主栽品种的抗蚜性及其机制研究[J]. 棉花学报, 2009, 21(1): 57-63.  
Lu Yi, Wang Peiling, Liu Bing, et al. Resistance and relevant mechanism to *Aphis gossypii* Glover of main cotton varieties in Xinjiang[J]. Cotton Science, 2009, 21(1): 57-63.
- [2] 张帅, 吕丽敏, 王春义, 等. 不同棉花品种在棉蚜胁迫下防御相关基因应答反应[J]. 棉花学报, 2013, 25(5): 388-395.  
Zhang Shuai, Lü Limin, Wang Chunyi, et al. Response of three resistance genes to damage of *Aphis gossypii* Glover (Hemiptera: Aphididae) in cotton[J]. Cotton Science, 2013, 25(5): 388-395.
- [3] 雒珺瑜, 张帅, 任相亮, 等. 近十年我国棉花虫害研究进展[J]. 棉花学报, 2017, 29(S1): 100-112. <https://doi.org/10.11963/1002-7807.ljycj.20170825>  
Luo Junyu, Zhang Shuai, Ren Xiangliang, et al. Research progress of cotton insect pests in China in recent ten years[J]. Cotton Science, 2017, 29(S1): 100-112.
- [4] Ghyar B S, Birajdar G K. Computer vision based approach to detect rice leaf diseases using texture and color descriptors [C]// 2017 International Conference on Inventive Computing and Informatics (ICICI). New York: IEEE, 2017: 1074-1078. (2018-05-28) [2019-11-01]. <https://doi.org/10.1109/ICICI.2017.8365305>
- [5] Chodey M D, Tamkeen H. Crop pest detection and classification by K-means and EM clustering[J]. International Research Journal of Engineering and Technology (IRJET), 2019, 6(9): 2130-2135.
- [6] 李震, 洪添胜, 倪慧娜, 等. 用高光谱成像技术检测柑橘红蜘蛛为害叶片的色素含量[J]. 农业工程学报, 2014, 30(6): 124-130. <https://doi.org/10.3969/j.issn.1002-6819.2014.06.015>  
Li Zhen, Hong Tiansheng, Ni Huina, et al. Pigment content measurement for citrus red mite infected leaf using hyper-spectral imaging technology[J]. Transactions of the Chinese Society of Agricultural Engineering, 2014, 30(6): 124-130.
- [7] Wu X, Zhang W, Qiu Z, et al. A novel method for detection of *Pieris rapae* larvae on cabbage leaves using NIR hyperspectral imaging[J]. Applied Engineering in Agriculture, 2016, 32(4): 311-316. <https://doi.org/10.13031/aea.32.11444>
- [8] Cao F, Liu F, Guo H, et al. Fast detection of *Sclerotinia sclerotiorum* on oilseed rape leaves using low-altitude remote sensing technology[J]. Sensors, 2018, 18(12): 4464. <https://doi.org/10.3390/s18124464>
- [9] Huang L, Yang L, Meng L, et al. Potential of visible and near-infrared hyperspectral imaging for detection of *Diaphania pyloalis* larvae and damage on mulberry leaves[J]. Sensors, 2018, 18(7): 2077. <https://doi.org/10.3390/s18072077>
- [10] Rajan P, Radhakrishnan B. A survey on different image processing techniques for pest identification and plant disease detection[J]. International Journal of Computer Science and Network (IJCSN), 2016, 5(1): 137-141.
- [11] Javed M H, Noor M H, Khan B Y, et al. K-means based automatic pests detection and classification for pesticides spraying [J]. International Journal of Advanced Computer Science and Applications, 2017, 8(11): 236-240.
- [12] Dey A, Bhoumik D, Dey K N. Automatic detection of whitefly pest using statistical feature extraction and image classification-n methods [J]. International Research Journal of Engineering and Technology, 2016, 3(9): 950-959.
- [13] 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. <https://doi.org/10.3969/j.issn.1002-6819.2014.12.016>  
Tian Youwen, Cheng Yi, Wang Xiaoqi, et al. Feature vectors determination for pest detection on apples based on hyperspectral imaging[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2014, 30(12): 132-139.
- [14] Fan Y, Wang T, Qiu Z, et al. Fast detection of striped stem-borer (*Chilo suppressalis* Walker) infested rice seedling based on visible/near-infrared hyperspectral imaging system[J]. Sensors, 2017, 17(11): 2470. <https://doi.org/10.3390/s17112470>
- [15] Kong W, Zhang C, Cao F, et al. Detection of *Sclerotinia* stem rot on oilseed rape (*Brassica napus* L.) leaves using hyperspectral imaging[J]. Sensors, 2018, 18(6): 1764. <https://doi.org/10.3390/s18061764>
- [16] Li X, Wei Y, Xu J, et al. SSC and pH for sweet assessment and maturity classification of harvested cherry fruit based on NIR hyperspectral imaging technology[J]. Postharvest Biology and Technology, 2018, 143: 112-118. <https://doi.org/10.1016/j.postharvbio.2018.05.003>
- [17] Lei T, Lin X H, Sun D W. Rapid classification of commercial Cheddar cheeses from different brands using PLSDA, LDA and SPA-LDA models built by hyperspectral data[J]. Journal of Food Measurement and Characterization, 2019, 13 (4): 1-11. <https://doi.org/10.1007/s11694-019-00234-0>
- [18] Abbas Z, Rehman M, Najam S, et al. An efficient Gray-Level Co-occurrence Matrix (GLCM) based approach towards classification of skin lesion[C/OL]// 2019 Amity International Conference on Artificial Intelligence (AICAI). (2019-02-01) [2019-11-01]. New York: IEEE, 2019: 317-320. <https://doi.org/10.1109/AICAI.2019.8701374>
- [19] Setiawan A S, Wesley J, Purnama Y. Mammogram classification using law's texture energy measure and neural networks[J]. Procedia Computer Science, 2015, 59: 92-97. <https://doi.org/10.1016/j.procs.2015.07.341>
- [20] Bonifazi G, Capobianco G, Serranti S. Hyperspectral imaging

- and hierarchical PLS-DA applied to asbestos recognition in construction and demolition waste[J]. *Applied Sciences*, 2019, 9(21): 4587. <https://doi.org/10.3390/app9214587>
- [21] Zhao W, Li S, Li A, et al. Hyperspectral images classification with convolutional neural network and textural feature using limited training samples [J]. *Remote Sensing Letters*, 2019, 10(5): 449-458. <https://doi.org/10.1080/2150704X.2019.1569274>
- [22] 李翠玲, 姜凯, 马伟, 等. 基于高光谱的番茄叶片斑潜蝇虫害检测[J]. *光谱学与光谱分析*, 2018, 38(1): 253-257. [https://doi.org/10.3964/j.issn.1000-0593\(2018\)01-0253-05](https://doi.org/10.3964/j.issn.1000-0593(2018)01-0253-05)
- Li Cuiling, Jiangkai, Mawei, et al. Tomato leaf *Liriomyza Sativae* blanchard pest detection based on hyperspectral technology [J]. *Spectroscopy and Spectral Analysis*, 2018, 38(1): 253-257.
- [23] 王小龙, 邓继忠, 黄华盛, 等. 基于高光谱数据的棉田虫害鉴别研究[J]. *华南农业大学学报*, 2019, 40(3): 97-103. <https://doi.org/10.7671/j.issn.1001-411X.201807041>
- Wang Xiaolong, Deng Jizhong, Huang Huasheng, et al. Identification of pests in cotton field based on hyperspectral data[J]. *Journal of South China Agricultural University*, 2019, 40(3): 97-103.
- [24] Chen T, Zeng R, Guo W, et al. Detection of stress in cotton (*Gossypium hirsutum* L.) caused by aphids using leaf level hyperspectral measurements[J]. *Sensors*, 2018, 18(9): 2798. <https://doi.org/10.3390/s18092798>
- [25] 陈兵, 王克如, 李少昆, 等. 蚜虫胁迫下棉叶光谱特征及其遥感估测[J]. *光谱学与光谱分析*, 2010, 30(11): 3093-3097. [https://doi.org/10.3964/j.issn.1000-0593\(2010\)11-3093-05](https://doi.org/10.3964/j.issn.1000-0593(2010)11-3093-05)
- Chen Bing, Wang Keru, Li Shaokun, et al. Study on spectrum characteristics of cotton leaf and its estimating with remote sensing under aphid stress[J]. *Spectroscopy and Spectral Analysis*, 2010, 30(11): 3093-3097.
- [26] 罗菊花, 黄木易, 赵晋陵, 等. 冬小麦灌浆期蚜虫危害高光谱特征研究[J]. *农业工程学报*, 2011, 27(7): 215-219. <https://doi.org/10.3969/j.issn.1002-6819.2011.07.038>
- Luo Juhua, Huang Muyi, Zhao Jinling, et al. Spectrum characteristics of winter wheat infected by aphid in filling stage [J]. *Transactions of the CSAE*, 2011, 27(7): 215-219.
- [27] 刘林. 桃蚜诱导杏幼叶局部细胞超微结构退化[J]. *果树学报*, 2011, 28(6): 1050-1053, 1147. <https://doi.org/10.13925/j.cnki.gsxb.2011.06.003>
- Liu Lin. Ultrastructural degeneration of part cells in young leaves of apricot caused by green peach aphids [J]. *Journal of Fruit Science*. 2011, 28(6): 1050-1053, 1147.
- [28] Sankaran S, Mishra A, Ehsani R, et al. A review of advanced techniques for detecting plant diseases[J]. *Computers and Electronics in Agriculture*, 2010, 72(1): 1-13. <https://doi.org/10.1016/j.compag.2010.02.007>
- [29] Sai S M, Gopichand G, Reddy C V, et al. High accurate unhealthy leaf detection[Z]. arXiv. 2019. <https://arxiv.org/abs/1908.09003>
- [30] Zhao Y, Yu K, Feng C, et al. Early detection of aphid (*Myzus persicae*) infestation on Chinese cabbage by hyperspectral imaging and feature extraction[J]. *Transactions of the ASABE*, 2017, 60(4): 1045-1051. <https://doi.org/10.13031/trans.11886>