

基于高光谱病害特征提取的温室黄瓜霜霉病早期检测

秦立峰^{1,2} 张 熹^{1,2} 张晓茜^{1,3}

(1. 西北农林科技大学机械与电子工程学院, 陕西杨凌 712100; 2. 农业农村部农业物联网重点实验室, 陕西杨凌 712100;
3. 陕西省农业信息感知与智能服务重点实验室, 陕西杨凌 712100)

摘要: 针对温室黄瓜早期霜霉病高光谱图像田间采集环境光照的影响及有效病害特征难以提取的问题,提出融合病害差异信息改进的竞争性自适应重加权算法(Competitive adaptive reweighted sampling, CARS)和连续投影算法(Successive projections algorithm, SPA)相结合的特征波段提取方法,并建立了黄瓜霜霉病早期检测模型。首先,采集黄瓜健康叶片和染病 12 d 内每天的高光谱图像,按病程分为 7 类;提取感兴趣区域,并计算平均光谱作为光谱数据;采用包络线消除法确定霜霉病害差异波段,基于病害差异波段采用 CARS 对 7 个不同阶段的光谱数据分别提取特征波段,再利用 SPA 进行二次降维寻优;最后,将各特征波段组合,得到 47 个特征波段数据,据此建立最小二乘-支持向量机(Least square support vector machines, LSSVM)模型,用于病害检测。在 94 个叶片样本组成的测试集上进行了病害检测实验,结果表明,融合病害差异信息的 Dis-CARS-SPA-LSSVM 对染病 2 d 到发病 12 d 均能取得 100% 的检测识别率;对染病 1 d 的测试集检测识别率达到 95.83%,其中染病样本的召回率达到 100%,相较于未融合病害差异信息的 CARS-SPA 特征提取方法识别率高 4.16 个百分点。说明所提出的 Dis-CARS-SPA-LSSVM 模型能够有效实现温室黄瓜霜霉病害的早期检测。

关键词: 温室; 黄瓜霜霉病; 早期检测; 高光谱成像; 病害差异波段; 特征波段

中图分类号: S436.5; TP391.4 **文献标识码:** A **文章编号:** 1000-1298(2020)11-0212-09 **OSID:** 

Early Detection of Cucumber Downy Mildew in Greenhouse by Hyperspectral Disease Differential Feature Extraction

QIN Lifeng^{1,2} ZHANG Xi^{1,2} ZHANG Xiaoqian^{1,3}

(1. College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China
2. Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture and Rural Affairs, Yangling, Shaanxi 712100, China
3. Shaanxi Key Laboratory of Agricultural Information Perception and Intelligent Service, Yangling, Shaanxi 712100, China)

Abstract: For the early hyperspectral images of cucumber downy mildew in greenhouses collected in field, it is influenced by environmental illumination and difficult to extract effective features from them. To solve these problems, a novel method of extracting feature bands based on disease difference information was proposed, which improved competitive adaptive reweighted sampling (CARS) and successive projections algorithm (SPA). Besides, an early detection model was built for cucumber downy mildew. Firstly, hyperspectral images were collected for leaves of healthy cucumber and leaves after infection in 12 consecutive days, which were divided into seven categories based on the degree of infection. Then, spectral data was calculated as the average spectrum of region of interest, the difference bands of downy mildew disease were determined by envelope elimination method and feature bands were extracted via CARS for seven different stages of it. SPA was used to perform secondary dimensionality reduction and optimization. Finally, all feature bands were combined to obtain 47 feature bands data. Based on this, a least square support vector machine (LSSVM) was established for disease detection. The disease detection test was performed on a test set of 94 leaf samples. The results showed that Dis-CARS-SPA-LSSVM fused disease difference information can obtain 100% detection rate after 2 ~ 12 days infection of disease. The detection rate of the test set infected with disease for 1 day reached

95.83% , the recall rate of infected samples reached 100% , and it avoided the randomness of CARS - SPA feature extraction method which did not fuse the disease difference information due to the interference bands of the non-downy mildew disease feature bands , and the recognition rate was 4.16 percentage points higher than that of CARS - SPA feature extraction model . The experiment results demonstrated that the proposed Dis - CARS - SPA - LSSVM model can effectively achieve early detection of downy mildew disease in greenhouse with a higher accuracy rate .

Key words: greenhouse ; cucumber downy mildew disease ; early detection ; hyperspectral imaging ; disease difference bands ; feature band

0 引言

温室黄瓜因处于大棚高湿环境而极易发生病害。其中,黄瓜霜霉病发病快、传播迅猛,是温室黄瓜的主要病害之一^[1]。霜霉病病原体侵染黄瓜植株的过程分为接触期、侵入期、潜育期和发病期。在侵入期和潜育期,无法通过肉眼发现病原菌在寄主体内的扩展蔓延。一旦进入发病期,病害便迅速传播扩散,其防治难度大、成本高^[2]。因此,早期进行黄瓜霜霉病害的检测识别对于病害防治具有十分重要的意义。

目前,在基于图像和光谱技术的病害早期检测中,基于可见/近红外光谱分析的定量分析精度受参考理化分析精度的影响,误差较大^[3-5];热红外技术受环境温湿度影响较大,对高温高湿的温室环境适应性差。高光谱成像技术具有光谱分辨率高、波段数多和图谱合一等优点^[6-8],可以反映对象内部的细微变化,适于病害早期植物体内部不明显的病变信息提取,且受温室环境影响很小,受到了研究者的广泛关注。文献[9]借助高光谱成像技术,在 400 ~ 1 000 nm 范围内针对玉米产毒真菌病害进行了早期检测研究;文献[10]基于高光谱成像技术实现了大豆花叶病早期检测与分级,提出了一种基于 CNN 模型的大豆花叶病早期检测方法;文献[11]研究了基于高光谱成像技术的大麦真菌病害早期检测,实现了 48 h 内光谱分析病害检测及 72 h 内微小病斑的识别。

研究表明,利用高光谱成像技术进行植物病害早期检测是可行的,实验室检测大多先采摘叶片、再采集信息,使叶片受到永久性损伤,而田间环境下叶片受光照、自然辐射、风速等影响较大。目前,对田间采集的病害早期特征提取的研究较少,且病害早期有效特征波段选择方法很难兼顾病害信息全面和数据量少的要求。本文以温室黄瓜的霜霉病为对象,对田间黄瓜叶片从健康至发病连续采集图像,利用高光谱技术分析病害发病早期与健康叶片的光谱差异,提出一种融合病害差异信息改进的竞争性自适应重加权算法(CARS)与连续投影法(SPA)结合

的优化特征提取方法,建立黄瓜霜霉病早期检测的偏最小二乘-支持向量机(LS-SVM)模型,实现对温室黄瓜霜霉病早期的高精度检测。

1 材料与方法

1.1 材料

实验采用盆栽培育黄瓜幼苗 20 盆,共计 40 株。黄瓜品种为北方大部分地区普遍种植的津优绿福 168,以陕西霖科生态工程有限公司的种植基质进行培植。室外大田搭设小型温室大棚模拟环境,棚内放置温湿度计和小兴看看 Q720 型无线高清智能摄像机,实时监控并记录棚内温湿度变化及黄瓜植株生长状况。午间高温时段敞棚通风降温 and 加湿器补充水分,夜间低温时段闭棚保温保湿,根据实时监控温湿度变化调控环境,保证黄瓜健康生长、避免自然感染病害,待生长至结果期、真叶完全伸展至成年人手掌大小时,进行霜霉病病菌接种。于西北农林科技大学北校区园艺场温室黄瓜大棚中采集感染霜霉病的黄瓜活体叶片,进行病菌孢子分离,制成悬浮液,对实验黄瓜叶片进行喷洒接种,接种维持棚内相对湿度 75% 以上,确保病菌成功侵染叶片。

1.2 仪器与设备

采用四川双利合谱科技有限公司 GaiaSky - mini 型推扫式高光谱成像系统便携式高光谱相机,相机及内置控制器质量 1.5 kg,光谱范围为 400 ~ 1 000 nm,光谱分辨率 3.5 nm,共 256 个光谱通道,单幅拍摄速度 10 s,全幅图像像素分辨率为 1 920 (空间维) × 1 440 (光谱维)。

1.3 高光谱图像采集与预处理

1.3.1 高光谱图像采集

对侵染霜霉病菌后的叶片进行高光谱图像信息采集。根据植保专家建议,将接病操作后的第 3 天作为黄瓜叶片成功接染霜霉病第 1 天,按成功染病后天数和染病程度将高光谱图像分为染病 1 d、2 d、3 d、4 d (肉眼不可见)、5 ~ 6 d (肉眼可见颜色变化)、7 ~ 8 d (肉眼可见小病斑)、9 d 后 (病害严重) 共 7 类。

相机采用日光光源,考虑到霜霉病从黄瓜中底

部叶片开始发病,为方便拍摄,采用如图 1 所示的图像采集方式。用支架固定高光谱相机,设置高光谱镜头斜向下 45°与叶片垂直距离 30 cm 架设,并在镜头旁绑定标尺以保证镜头与叶片距离固定,手托扶叶片进行拉展和垂直镜头采集图像。图像采集时间分别为每天 10:00—11:00 光照良好时段和 16:00—17:00 光照较弱时段,以提高后期所建立识别模型的不同光照条件适应度。每次采集图像前先对白板 and 黑板采集记录,从 20 盆黄瓜中选出生长状态良好、叶片平展宽大的 10 盆共 32 片叶片。首先采集染病前的健康叶片;再对接种黄瓜霜霉病菌后的叶片逐天连续采集 12 d,得到共 384 幅高光谱图像。将图像在 ENVI 5.1 软件中读取,剔除肉眼可见因设备或叶片抖动导致图像模糊的异常样本,如图 2a 所示,对比每幅图像的叶片区域光谱曲线,剔除光谱曲线明显差异于黄瓜叶片光谱曲线的异常样本,如图 2b 红线所示,对剩余 286 幅样本图像进行实验。

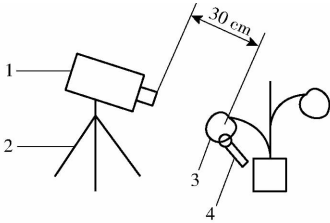


图 1 图像采集示意图
Fig. 1 Image acquisition diagram

1. 高光谱相机 2. 固定支架 3. 样本叶片 4. 操作人员手持托扶架

1.3.2 高光谱图像预处理

原始高光谱图像包含大量的噪声和背景干扰,需进行预处理以消减噪声,并去除背景,提取感兴趣区域。先用 SpecView 软件对叶片高光谱图像进行黑白校正以消减光照影响,再通过 ENVI 5.1 软件读取高光谱图像并分离背景后,对整片叶部分提取感兴趣区域。

图 3a 为叶片样本在某一波段的高光谱图像,对比目标叶片区域与背景的光谱曲线(图 3b)可知,在 670~750 nm 之间黄瓜叶片的光谱曲线出现一个较大跃变,而背景区域的光谱全波段内变化较小。因此,对第 120 个波长 671.0 nm 与第 150 个波长 744.6 nm 的光谱求取波段比,根据波段比差异可分离叶片与背景区域;通过腐蚀去除背景部分;膨胀后再腐蚀得到整体目标叶片的二值掩模,如图 4a 所示;将掩模图像与原图像的 400~922 nm 共 217 个波段的图像相乘得到目标整体叶片的感兴趣区域,如图 4b 所示。感兴趣区域的平均光谱如图 4c 所示,将它作为该样本的光谱数据。

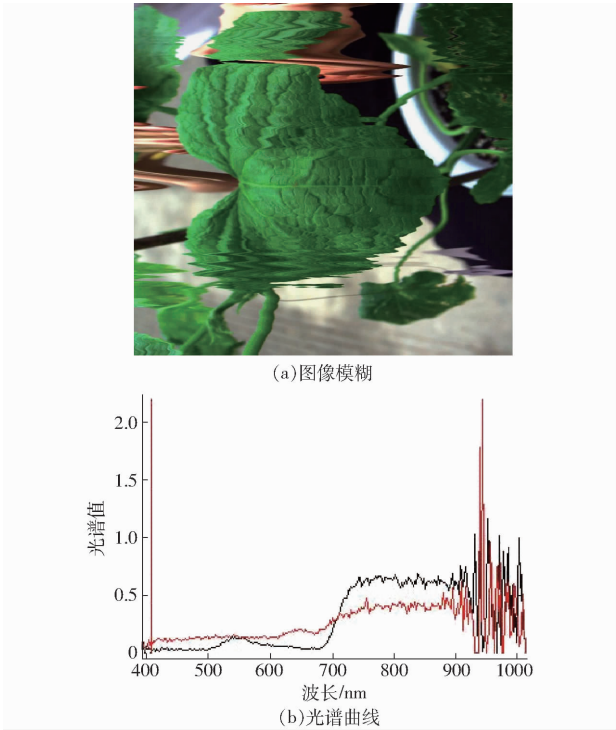


图 2 异常样本
Fig. 2 Abnormal sample

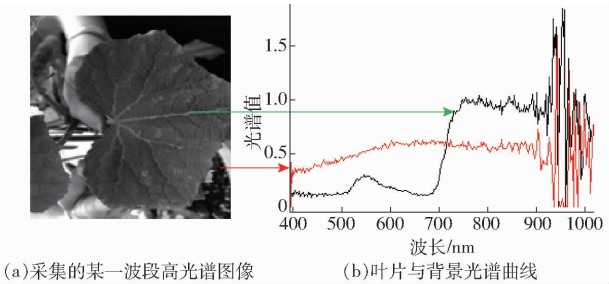


图 3 目标叶片区域与背景分割

Fig. 3 Target leaf area and background segmentation

1.4 特征波段提取算法

从高光谱图像所有波段中选择可分性好的波段子集,保留有用信息,不仅能降低数据维度,还能提高分类识别精度^[12-13]。

在常用特征提取算法中,CARS 结合蒙特卡洛采样及偏最小二乘(Partial least squares, PLS)回归系数筛选的特征波段优选方法^[14-15],以达尔文进化论的“适者生存”法则为指导理论^[16],在校正集中利用蒙特卡洛采样法选取样本,对样本进行偏最小二乘回归(PLSR)分析,利用自适应重加权算法(Adaptive reweighted sampling, ARS)选择波段变量回归系数绝对值大的波段点,对于绝对值小、权重较小的波段点,采用衰减指数法(Exponentially decreasing function, EDF)确定个数并去除,对选取的权重较大的波段变量建立 PLS 模型,利用交互验证选出均方根误差(Root mean square error of cross validation, RMSECV)最小的模型,其所对应的波段

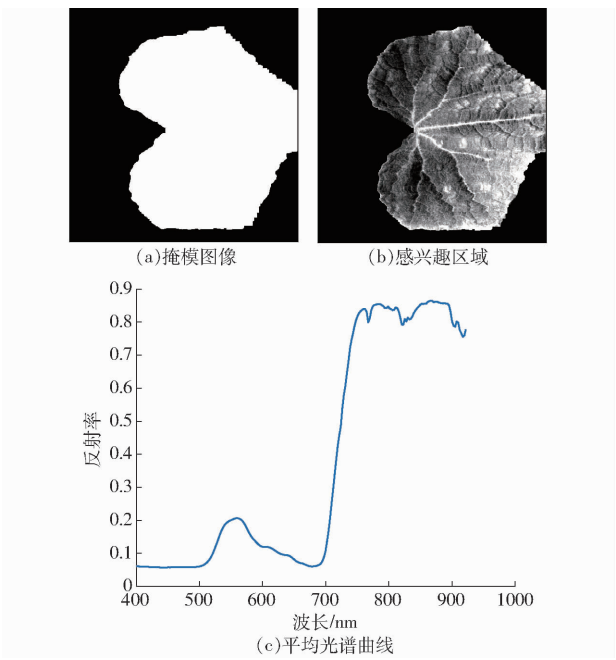


图 4 感兴趣区域提取与平均光谱计算

Fig. 4 Region of interest extraction and average spectral calculation

变量即为所选特征波段^[17-20]。

CARS 算法可用于提取感染霜霉病害的样本与健康样本相比变量权重变大的特征波段,但 CARS 算法的随机采样使该算法本身存在不确定性,每次提取的特征波段数量及位置随机性很大,所提取的特征存在漏选及冗余问题,无法作为病害特征进行建模。为解决该问题,在 CARS 算法中加入变量稳定性计算的循环迭代,通过多次执行 CARS 提取特征波段,将结果合并以解决漏选问题,同时比较并删减各次提取结果中位置相似的特征波段,以解决冗余问题。多次执行加入变量稳定性计算的算法后强制筛选出特征波段作为 CARS 算法提取的特征波段。

但 CARS 算法提取特征波段同时会保留因噪声引起的变量权重变大的波段。SPA 是一种稳定的前向变量选择算法^[21-22],从原始光谱信息中筛选寻找出冗余信息最少的波段变量组来概括代表大多数样本的光谱信息,使变量的共线性最小,明确样本特征^[23-25]。SPA 可以消除光谱冗余信息,进一步简化筛选 CARS 算法提取出的特征波段,便于建模,提高模型检测精度。

1.5 融合病害差异信息的特征提取方法改进

1.5.1 Dis-CARS-SPA 特征提取方法

由于图像通过便携式高光谱相机于户外棚间采集,光照强度变化导致采集到的图像光谱数据存在光谱反射率的差异和较多波动,而 CARS 结合 SPA 的特征提取算法仅对随机提取的波段样本中权重较

大的进行选取与筛选,因此采用传统的 CARS-SPA 算法对全波段光谱提取特征波段,易将反射率的波动差异作为权重较大的波段变量提取为病害特征,影响真正病害特征的选取。

光谱差异波段的分析与提取主要用于植被种类识别、土壤有机质与含水率的检测和水质分析等领域,文献[26]通过包络线变换前后的光谱曲线提取鄱阳湖湿地植被的光谱差异波段,利用马氏距离法检验了植被识别效果;文献[27]通过去包络线方法提取反射光谱特征差异,建立了土壤水分含量高光谱预测模型;文献[28]探讨相同树种叶片 SPAD 变化时的光谱差异和相同 SPAD 不同树种叶片的光谱特征,分析植物健康状况。本文借鉴光谱差异分析思想,在 CARS-SPA 特征提取算法前加入病害差异波段的提取,提出 Dis-CARS-SPA 特征提取方法,使 CARS-SPA 算法融合病害差异信息,在准确表征病害的波段范围中进行 CARS-SPA 的特征提取,有效避免因采集环境、不同光照强度或其他病害所导致的光谱出现差异的影响,提高模型准确性。

1.5.2 病害差异波段提取

同一叶片随病害侵染的严重程度的不同,高光谱曲线呈现不同程度的差异。黄瓜霜霉病叶片与健康叶片出现光谱差异的波段范围,即病害差异波段。

图 5 为叶片从健康到染病 9 d 后病害严重的逐日各阶段光谱曲线。

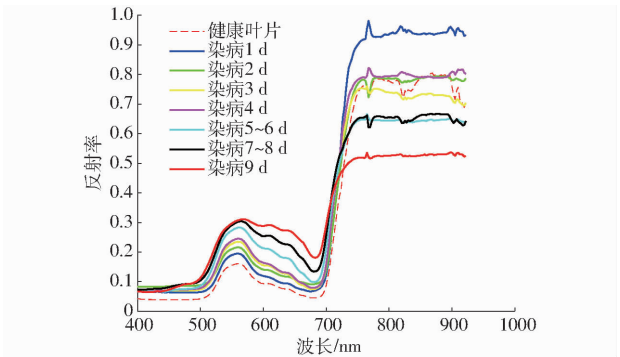


图 5 叶片健康至染病严重逐日光谱曲线

Fig. 5 Spectral curves from leaf health to serious disease day by day

由图 5 可知,在 450 ~ 720 nm 的波段范围内光谱曲线存在较明显差异变化,其中 550 ~ 650 nm 的差异变化尤为显著,正是由于该范围对应可见光谱的黄色及黄绿色,霜霉病的侵染导致叶片的叶绿体死亡减少,逐渐枯黄,因此在该波段范围内反射率增大;而 750 nm 后出现了光谱曲线较多波动及反射率的不同,是由于每天光照强度不同导致最终趋于平稳的光谱反射率有高有低;存在的较多波动是由于拍摄环境及设备自身的噪声影响。这两种影响采用

一般的光谱预处理方法无法消除,反而易消减病害带来的光谱差异,因此本文通过对病害差异信息的提取来改进 CARS-SPA 特征提取方法,以提高特征提取的针对性。

在光谱曲线相似的情况下,直接从中提取光谱特征不便于计算。为了更加精准地确定病害的差异波段范围,避免肉眼观察的主观性和不准确性,通过包络线消除法^[29]确定光谱吸收曲线。

包络线消除法可以有效突出光谱曲线的吸收、反射和发射特征,并将其归一到一致的光谱背景上,有利于和其他光谱曲线进行特征数值的比较。其算法步骤如下:通过求导得到光谱曲线上所有极大值点。以最大极大值点作为包络线的一个端点,计算该点与长波和短波方向各个极大值连线的斜率,以斜率最大点作为上包络线下一个端点,斜率最小点作为下包络线下一个端点,再分别以此点为起点循环,直至最后一点。连接所有端点,可形成曲线的包络线,用实际光谱反射率去除包络线上相应波段的反射率值,可得到包络线消除法归一化后的值。

1.6 检测模型

将病害早期检测作为一个分类识别问题,利用融合病害差异信息的 Dis-CARS-SPA 特征提取方法选取的特征波段训练 LSSVM 作为病害检测模型。LSSVM 通过优化目标中的线性约束,只求解一个线性方程组来代替 SVM 中的二次规划(QP)问题,通过解决等式约束以及最小二乘问题,简化求解过程。模型参数包括惩罚因子 gamma 和核参数 sig2,通过粒子群优化算法^[30-31](Particle swarm optimization, PSO)进行寻优选取。最后,用测试集的检测识别率 P 和模型混淆矩阵计算得到的召回率 R 作为模型评价标准,识别率计算为

$$P = \frac{N_{\text{Mod}}}{N_{\text{Act}}} \times 100\%$$

(1)

式中 N_{Mod} ——模型检测正确的健康与染病样本数
 N_{Act} ——实际输入的样本数

召回率计算公式为

$$R = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}} \times 100\%$$

(2)

式中 N_{TP} ——健康或染病样本模型正确识别的样本数

N_{FN} ——健康或染病样本模型错误识别的样本数

2 结果分析

2.1 病害差异波段提取

通过包络线消除法选择病害的差异波段,包络

线包括上包络线与下包络线,通过上包络线与实际光谱曲线相减得到光谱吸收曲线对比寻找病害差异波段。首先绘制健康与染病叶片的上包络线如图 6 所示,为凸显染病叶片与健康叶片的光谱差异,方便更好地识别并选择病害差异波段,染病叶片选择染病 9 d 后发病严重的样本图像。

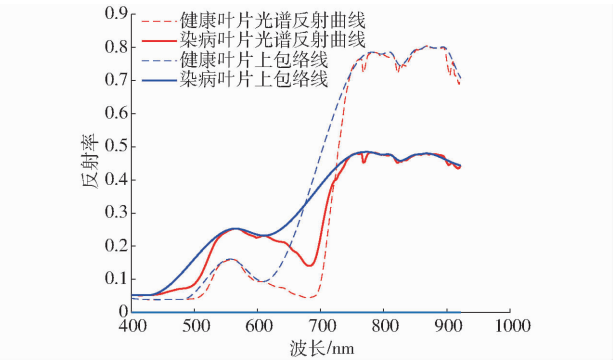


图 6 健康与染病叶片光谱曲线上包络线

Fig. 6 Envelopes on spectral curves of healthy leaf and diseased leaf

由图 6 可见健康与染病叶片存在较大差异,将上包络线与实际光谱曲线相减得到光谱的吸收曲线如图 7 所示,健康叶片吸收光谱与染病叶片吸收光谱作差,得到吸收光谱差异曲线如图 7 所示。

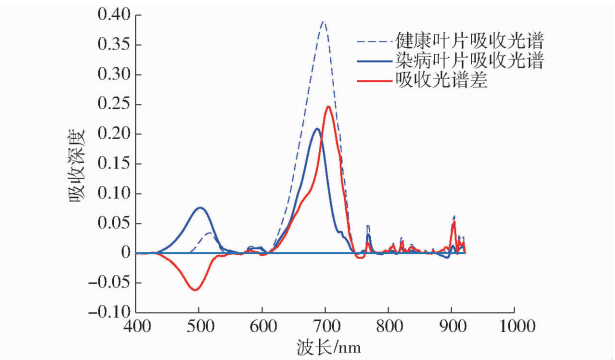


图 7 健康与染病叶片吸收光谱及差异曲线

Fig. 7 Spectral absorption and difference curves of healthy leaf and diseased leaf

根据健康与染病叶片的光谱吸收差异曲线与零轴的交点确定霜霉病的病害差异波段,并且去除 774 nm 后的不相关干扰波动波段,得到病害差异波段为 431 ~ 552 nm 及 571 ~ 774 nm 范围,即图 7 中吸收光谱差波动较大的部分。

2.2 CARS 算法提取特征波段

使用 CARS 算法在选取的 431 ~ 552 nm 及 571 ~ 774 nm 的病害差异波段内提取特征波段,将染病 1 d 至严重 7 个类别的样本定为染病叶片,与健康叶片进行特征波段的提取,得到共 18 个特征波段, RMSECV 最小值为 1.132 0。由于不同染病阶段之间也存在较大的光谱差异,提取特征波段均

方根误差较大,识别效果不佳。因此采取对染病的 7 个不同阶段分别提取特征波段再组合的方法。

以染病 1 d 叶片的特征波段提取为例,设置蒙特卡洛采样次数为 50,为解决 CARS 算法的随机性问题,加入变量稳定性计算的循环迭代,即对 CARS 算法每次随机采样提取的特征波段位置进行判定选取,使选取的波段位置稳定。本文通过预实验,设置循环提取次数为 3,3 次 CARS 提取的特征波段位置如图 8 所示,提取的特征波段数量分别为 31、22、23;由图 8 可见,3 次选取的特征波段位置相近,满足稳定性要求。将这 3 次的特征波段合并并去除重复项,得到共 48 个特征波段;以提取的每个特征波段为基准,对比其前后位置提取的特征波段序号(高光谱相机共 256 个波段,

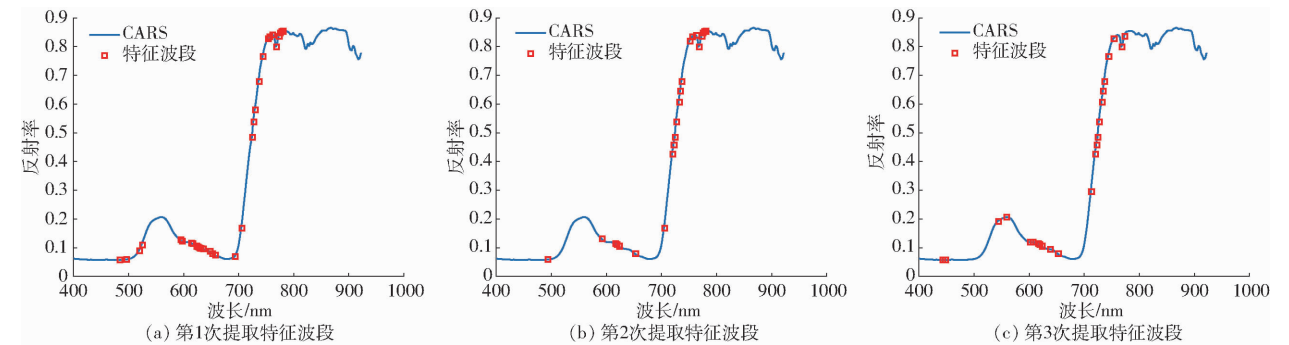


图 8 算法 3 次提取的特征波段位置
Fig. 8 Position of feature bands extracted by CARS for three times

强制变量筛选是选取 3 次加入变量稳定性计算后 CARS 算法均提取到的相同的特征波段,即从上述得到的 30、29、31 个特征波段中寻找相同特征波段,结果得到 28 个特征波段,如图 9 所示,即为本文算法选取的染病 1 d 的特征波段。

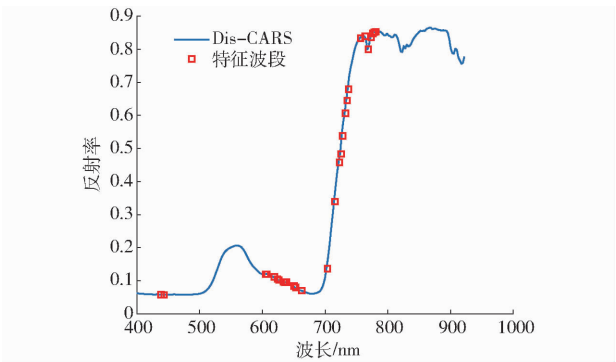


图 9 染病 1 d 融合病害差异信息的 CARS 算法提取的特征波段
Fig. 9 Feature extraction by CARS based on fusion of disease difference information in one day after infection

CARS 计算过程如图 10 所示,包括样本变量个数(Number of sampled variables, NSV)、交叉验证均方根误差(RMSECV)和回归系数(Regression

裁剪后剩余 217 个波段,序号即提取的 217 个波段中第几个波段),若该波段序号与前后提取的波段序号为相邻序号,则取 3 个波段的平均即中间波段,舍去前后提取的 2 个波段(如提取波段为第 88、89、90 个波段,则保留第 89 个波段,舍去第 88、90 个波段);若该波段序号仅与前后提取的波段之一的序号为相邻序号,则取相邻波段中较小一个,舍去较大一个(如提取波段为第 88、89、91 波段,则舍去第 89 个波段,保留第 88、91 个波段)。通过该方法对相似位置波段进行筛选,得到共 30 个特征波段。按该方法再重复进行 2 次特征波段的选取,分别得到 29、31 个特征波段,可见加入变量稳定性计算后 CARS 算法提取得到的特征段稳定在 30 个左右,降低了随机性,且提取到的特征波段几乎相同。

coefficients)随蒙特卡洛采样次数的变化趋势。在 RMSECV 最小的点取最优变量个数 26(图 10c 中的星号竖线),将回归系数绝对值大的波段保留,即为提取的特征波段。将染病 1 d 到严重 7 个阶段分别提取的波段组合作为 CARS 算法所提取特征波段。

2.3 SPA 算法提取特征波段

使用 SPA 算法对 7 个阶段的 CARS 特征波段分

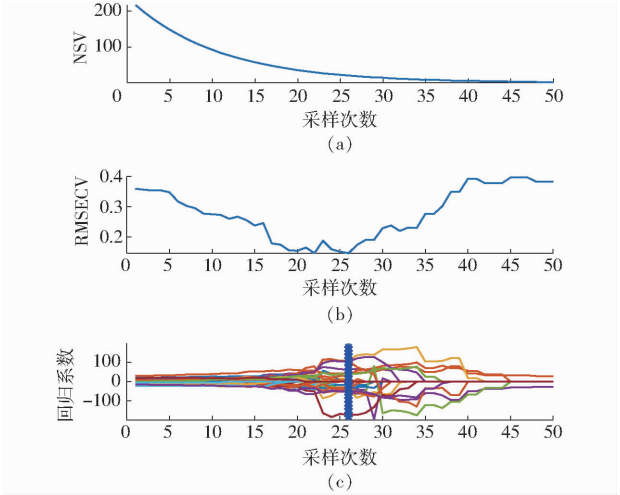


图 10 CARS 计算过程
Fig. 10 CARS calculation process

别进行降维再提取,将各波段组合后共得到 47 个特征波段,即为基于全波段的 CARS-SPA 提取的特征波段。如图 11 所示,特征波段都集中于霜霉病害的病害差异波段范围内,更准确代表霜霉病害的特征,消除了光谱末段噪声等干扰。

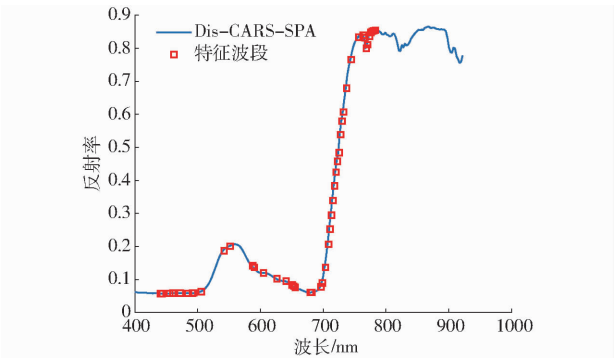


图 11 Dis-CARS-SPA 提取特征波段
Fig. 11 Dis-CARS-SPA extracted feature bands

2.4 早期检测模型建模与实验

将 286 个样本按 2:1 划分为建模集与测试集,健康叶片及染病 1 d 到发病严重 7 个阶段建模集样本数分别为 16、32、30、23、28、30、20、13;测试集样本数分别为 8、16、15、11、14、15、9、6。为验证模型从样本中识别出染病叶片的能力,将健康叶片样本作为正例,7 个不同染病阶段的叶片样本分别作为反例,建立 Dis-CARS-SPA-LSSVM 模型,同时建立未融合病害差异信息,基于全波段提取特征波段的 CARS-SPA-LSSVM 模型,作为对比分析检测效果。用 PSO 算法选择模型的惩罚因子 gamma 和核参数 sig2,不同染病阶段建模集与测试集样本数不同,PSO 选择的参数不同,结果如表 1 所示。

表 1 不同染病阶段模型参数选择
Tab.1 Selection of model parameters in different stages of infection

染病阶段	模型参数	
	惩罚因子	核参数
染病 1 d	1 000	200
染病 2 d	100	20
染病 3 d	100	20
染病 4 d	1 000	200
肉眼可见颜色变化	1 000	100
肉眼可见小病斑	100	20
染病严重	100	20

为验证模型稳定性,按 2:1 比例随机划分建模集和测试集,重复实验 3 次,将 3 次实验的识别率求平均,结果如表 2 所示。

未融合差异信息建立的 CARS-SPA-LSSVM 模型对各染病阶段的准确识别率虽均达到 90% 以

表 2 不同模型对各染病阶段识别结果比较
Tab.2 Comparison of recognition results of different models for each infection stage

输入变量	建立模型	测试集识别率 P/%
健康 + 染病 1 d	CARS-SPA-LSSVM	91.67
	Dis-CARS-SPA-LSSVM	95.83
健康 + 染病 2 d	CARS-SPA-LSSVM	95.65
	Dis-CARS-SPA-LSSVM	100
健康 + 染病 3 d	CARS-SPA-LSSVM	94.74
	Dis-CARS-SPA-LSSVM	100
健康 + 染病 4 d	CARS-SPA-LSSVM	95.46
	Dis-CARS-SPA-LSSVM	100
健康 + 染病可见 颜色变化	CARS-SPA-LSSVM	95.65
	Dis-CARS-SPA-LSSVM	100
健康 + 染病可见 小病斑	CARS-SPA-LSSVM	94.12
	Dis-CARS-SPA-LSSVM	100
健康 + 染病严重	CARS-SPA-LSSVM	92.86
	Dis-CARS-SPA-LSSVM	100

上,然而随染病程度加深,模型的识别效果并不稳定,由表 2 可见,该模型对染病 2 d 的测试集识别率达到 95.65%,而对染病 9 d 的测试集识别率只有 92.86%。这是由于未融合病害差异信息提取的特征波段有较多集中在 750 nm 之后,为光谱的随机性波动,并非霜霉病的病害特征,模型识别准确性受到影响,无法作为黄瓜霜霉病害的早期检测模型。

而本文提出的融合病害差异信息改进的 CARS-SPA 提取特征波段所建立的 Dis-CARS-SPA-LSSVM 模型对染病 2 d (不可见)到染病 9 d 后发病严重等 6 个阶段的识别率均达到 100%,相较于 CARS-SPA-LSSVM 模型均有提高,能够很好地实现对黄瓜叶片霜霉病害的早期及后期识别。表明融合病害差异信息提取特征波段能有效避免光谱干扰性波动带来的影响,针对病害差异波段区域提取特征,更准确地表征黄瓜霜霉病害。

提出的 Dis-CARS-SPA-LSSVM 模型对染病 1 d 的样本,测试集识别率达到了 95.83%,比 CARS-SPA-LSSVM 模型的识别率高 4.16 个百分点;其混淆矩阵如表 3 所示,计算健康样本的召回率为 87.50%,染病样本的召回率为 100%,表明能将染病叶片全部检测识别,仅会出现将健康叶片误判为染病的情况,而对于病害的早期检测及防治问题,染病样本的召回率更为关键,部分健康叶片误检测为染病是可以容忍的。该结果表明该模型对黄瓜霜霉病害具有很好的早期检测识别能力,可将黄瓜叶片霜霉病害的早期检测时间提前到染病后 1 d 的时间点。

表 3 Dis - CARS - SPA - LSSVM 模型染病 1 d 混淆矩阵

Tab.3 Dis - CARS - SPA - LSSVM model confusion matrix for 1 day after infection			
混淆矩阵	模型识别		召回率 $R/\%$
	健康	染病	
健康	7	1	87.50
染病	0	16	100

3 讨论

利用高光谱成像技术对温室黄瓜霜霉病进行早期检测,提出了融合病害差异信息改进的 CARS - SPA 特征提取方法。先用包络线消除法选出表征病害特征的病害差异波段,通过 CARS 算法提取特征波段后再用 SPA 算法降维,既保证了对病害特征信息的有效提取,又保证了对冗余波段的剔除降维。在提取的 431 ~ 552 nm 及 571 ~ 774 nm 的病害差异波段范围以 CARS - SPA 精确提取表征黄瓜霜霉病害的特征波段,通过最小二乘支持向量机建立的 Dis - CARS - SPA - LSSVM 病害识别模型对 7 个染

病阶段的病害叶片进行检测,实验结果显示,融合病害差异信息改进的 CARS - SPA 特征提取方法进行特征波段提取,避免了非病害因素引起的特征波段光谱波动的影响;对染病仅 1 d 的叶片平均检测精度达到 95.83%,比单纯的 CARS - SPA 特征提取方法建立的 CARS - SPA - LSSVM 模型高 4.16 个百分点,染病样本的召回率达到 100%;对染病 2 d 之后的叶片检测精度均达到 100%。

4 结束语

本文提出的融合病害差异信息改进的 CARS - SPA 特征提取方法能有效利用病害差异信息,在精确表征病害特征的基础上有针对性地进行特征波段的提取。据此方法建立的 Dis - CARS - SPA - LSSVM 模型能有效提取温室黄瓜霜霉病害高光谱特征,将黄瓜霜霉病害早期检测的有效识别时间点提前到染病第 1 天。该方法适用于实际农业生产中的复杂环境,也可用于其他病害的田间早期检测。

参 考 文 献

[1] YU K, LEUFEN G, HUNSCHE M, et al. Investigation of leaf diseases and estimation of chlorophyll concentration in seven barley varieties using fluorescence and hyperspectral indices[J]. Remote Sensing, 2013, 6(1):64 - 86.

[2] 吴晓青,周方园,张新建.微生物组学对植物病害微生物防治研究的启示[J].微生物学报,2017,57(6):867 - 875. WU Xiaoqing, ZHOU Fangyuan, ZHANG Xinjian. Enlightenment from microbiome research towards biocontrol of plant disease [J]. Acta Microbiologica Sinica, 2017, 57(6): 867 - 875. (in Chinese)

[3] IKEOGU U N, DAVRIEUX F, DUFOUR D, et al. Rapid analyses of dry matter content and carotenoids in fresh cassava roots using a portable visible and near infrared spectrometer (vis/NIRS) [J]. PLOS ONE, 2017, 12(12):1 - 17.

[4] OMASA K, TAKAYAMA K. Simultaneous measurement of stomatal conductance, non-photochemical quenching, and photochemical yield of photosystem II in intact leaves by thermal and chlorophyll fluorescence imaging[J]. Plant and Cell Physiology, 2003, 44(12): 1290 - 1300.

[5] PRYTZ G, FUTSAETHER C M, JOHANSSON A. Thermography studies of the spatial and temporal variability in stomatal conductance of avena leaves during stable and oscillatory transpiration[J]. New Phytologist, 2003, 158(2): 249 - 258.

[6] 何勇,聂鹏程,刘飞.农业物联网与传感仪器研究进展[J/OL].农业机械学报,2013,44(10):216 - 226. HE Yong, NIE Pengcheng, LIU Fei. Advancement and trend of internet of things in agriculture and sensing instrument[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(10):216 - 226. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20131035&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2013.10.035. (in Chinese)

[7] KAMRUZZAMAN M, ELMASRY G, SUN D W, et al. Prediction of some quality attributes of lamb meat using near-infrared hyperspectral imaging and multivariate analysis[J]. Analytica Chimica Acta, 2012, 714:57 - 67.

[8] 谢传奇,方孝荣,邵咏妮,等.番茄叶片早疫病近红外高光谱成像检测技术[J/OL].农业机械学报,2015,46(3):315 - 319. XIE Chuanqi, FANG Xiaorong, SHAO Yongni, et al. Detection of early blight on tomato leaves using near-infrared hyperspectral imaging technique[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(3):315 - 319. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20150346&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2015.03.046. (in Chinese)

[9] FIORE D A, REVERBERI M, RICELLI A, et al. Early detection of toxigenic fungi on maize by hyperspectral imaging analysis [J]. International Journal of Food Microbiology, 2010, 144(1):64 - 71.

[10] 桂江生,吴子娴,李凯.基于卷积神经网络模型的大豆花叶病初期高光谱检测[J].浙江大学学报(农业与生命科学版),2019,45(2):256 - 262.

[11] 许凯雯.基于高光谱成像技术的大麦真菌病害早期检测的研究[D].杭州:浙江大学,2018.

[12] 金秀,齐海军,李绍稳.基于动态时间弯曲的马铃薯干腐病发病期时序高光谱诊断方法[J].食品科学,2018,39(19):233 - 240. JIN Xiu, QI Haijun, LI Shaowen. Diagnosis of symptom appearance of potato dry rot disease using time series hyperspectral imaging based on dynamic time warping[J]. Food Science, 2018, 39(19): 233 - 240. (in Chinese)

[13] 龙燕,连雅茹,马敏娟,等.基于高光谱技术和改进型区间随机蛙跳算法的番茄硬度检测[J].农业工程学报,2019,35(13):270 - 276. LONG Yan, LIAN Yaru, MA Minjuan, et al. Detection of tomato hardness based on hyperspectral technology and modified

- interval random frog algorithm[J]. Transactions of the CSAE, 2019, 35(13): 270 – 276. (in Chinese)
- [14] 王海龙, 杨国国, 张瑜, 等. 竞争性自适应重加权算法和相关系数法提取特征波长检测番茄叶片真菌病害[J]. 光谱学与光谱分析, 2017, 37(7): 2115 – 2119.
WANG Hailong, YANG Guoguo, ZHANG Yu, et al. Detection of fungal diseases on tomato leaves with competitive adaptive reweighting sampling and correlation analysis methods[J]. Spectroscopy and Spectral Analysis, 2017, 37(7): 2115 – 2119. (in Chinese)
- [15] 石岩, 孙冬梅, 熊婧, 等. 近红外光谱结合竞争性自适应重加权采样算法用于人工牛黄的质量分析研究[J]. 中国药理学杂志, 2018, 53(14): 1216 – 1221.
SHI Yan, SUN Dongmei, XIONG Jing, et al. Analysis of artificial Cow-Bezoar by near-infrared spectroscopy coupled with competitive adaptive reweighted sampling method[J]. Chinese Pharmaceutical Journal, 2018, 53(14): 1216 – 1221. (in Chinese)
- [16] 詹白勺, 倪君辉, 李军. 高光谱技术结合 CARS 算法库尔勒香梨可溶性固形物定量测定[J]. 光谱学与光谱分析, 2014, 34(10): 2752 – 2757.
ZHAN Baishao, NI Junhui, LI Jun. Hyperspectral technology combined with CARS algorithm to quantitatively determine the SSC in Korla Fragrant Pear[J]. Spectroscopy and Spectral Analysis, 2014, 34(10): 2752 – 2757. (in Chinese)
- [17] WU D, SUN D W. Potential of time series-hyperspectral imaging (TS – HSI) for non-invasive determination of microbial spoilage of salmon flesh[J]. Talanta, 2013, 111(13): 39 – 46.
- [18] 邱素君, 罗晓健, 张国松, 等. 近红外漫反射光谱无损预测片剂硬度研究[J]. 中国药理学杂志, 2016, 51(11): 904 – 909.
QIU Sujun, LUO Xiaojian, ZHANG Guosong, et al. Non-destructive prediction of tablet hardness by near infrared diffuse reflection spectroscopy[J]. Chinese Pharmaceutical Journal, 2016, 51(11): 904 – 909. (in Chinese)
- [19] 彭严芳, 史新元, 李洋, 等. 基于多变量检测限的模型变量筛选方法研究[J]. 世界科学技术-中医药现代化, 2014, 16(5): 960 – 965.
PENG Yanfang, SHI Xinyuan, LI Yang, et al. Optimization of near infrared variable selection method based on multivariate detection limit[J]. World Sci. Technol. Mod. Tradit. Chin. Med. Mater Med., 2014, 16(5): 960 – 965. (in Chinese)
- [20] 杜晨朝, 赵安邦, 吴志生, 等. 近红外光谱结合不同变量筛选方法用于金银花提取过程中绿原酸量的在线监测[J]. 中草药, 2017, 48(16): 3317 – 3321.
DU Chenzhao, ZHAO Anbang, WU Zhisheng, et al. Online control of chlorogenic acid in *Lonicerae Japonicae* Flos by near infrared spectroscopy combined with different variable selections[J]. Chinese Traditional and Herbal Drugs, 2017, 48(16): 3317 – 3321. (in Chinese)
- [21] GAO Junfeng, LI Xiaoli, ZHU Fengle, et al. Application of hyperspectral imaging technology to discriminate different geographical origins of *Jatropha curcas* L. seeds[J]. Computers and Electronics in Agriculture, 2013, 99:186 – 193.
- [22] 刘思伽, 田有文, 张芳, 等. 采用二次连续投影法和 BP 人工神经网络的寒富苹果病害高光谱图像无损检测[J]. 食品科学, 2017, 38(8): 277 – 282.
LIU Sijia, TIAN Youwen, ZHANG Fang, et al. Hyperspectral imaging for nondestructive detection of Hanfu apple diseases using successive projections algorithm and BP neural network[J]. Food Science, 2017, 38(8): 277 – 282. (in Chinese)
- [23] PAN Leiqing, SUN Ye, XIAO Hui, et al. Hyperspectral imaging with different illumination patterns for the hollowness classification of white radish[J]. Postharvest Biology and Technology, 2017, 126: 40 – 49.
- [24] CHENG Junhu, SUN Dawen. Rapid and non-invasive detection of fish microbial spoilage by visible and near infrared hyperspectral imaging and multivariate analysis[J]. LWT-Food Science and Technology, 2015, 62(2): 1060 – 1068.
- [25] 吴迪, 宁纪锋, 刘旭, 等. 基于高光谱成像技术和连续投影算法检测葡萄果皮花色苷含量[J]. 食品科学, 2014, 35(8): 57 – 61.
WU Di, NING Jifeng, LIU Xu, et al. Determination of anthocyanin content in grape skins using hyperspectral imaging technique and successive projections algorithm[J]. Food Science, 2014, 35(8): 57 – 61. (in Chinese)
- [26] 况润元, 曾帅, 赵哲, 等. 基于实测高光谱数据的鄱阳湖湿地植被光谱差异波段提取[J]. 湖泊科学, 2017, 29(6): 1485 – 1490.
KUANG Runyuan, ZENG Shuai, ZHAO Zhe, et al. Extraction of the discriminative bands of Lake Poyang wetland vegetation based on the measured hyperspectral data[J]. Lake Sci., 2017, 29(6): 1485 – 1490. (in Chinese)
- [27] 金慧凝, 张新乐, 刘焕军, 等. 基于光谱吸收特征的土壤含水量预测模型研究[J]. 土壤学报, 2016, 53(3): 627 – 635.
JIN Huining, ZHANG Xinle, LIU Huanjun, et al. Soil moisture predicting model based on spectral absorption characteristics of the soil[J]. Acta Pedologica Sinica, 2016, 53(3): 627 – 635. (in Chinese)
- [28] 彭建, 徐飞雄, 邓凯, 等. 琅琊山景区不同叶绿素条件下树种叶片光谱差异分析[J]. 光谱学与光谱分析, 2018, 38(6): 1839 – 1849.
PENG Jian, XU Feixiong, DENG Kai, et al. Spectral differences of tree leaves at different chlorophyll relative content in Langya Mountain[J]. Spectroscopy and Spectral Analysis, 2018, 38(6): 1839 – 1849. (in Chinese)
- [29] 贾方方, 洪权春, 宋唯一. 基于去包络线法的番茄叶霉病发病程度估测方法[J]. 中国生态农业学报, 2017, 25(6): 805 – 811.
JIA Fangfang, HONG Quanchun, SONG Weiyi. Continuum removal method for monitoring *Fulvia fulva* morbidity using hyperspectral data[J]. Chinese Journal of Eco-Agriculture, 2017, 25(6): 805 – 811. (in Chinese)
- [30] 刘环宇, 陈海涛, 闵诗尧, 等. 基于 PSO – SVR 的植物纤维地膜抗张强度预测研究[J/OL]. 农业机械学报, 2017, 48(4): 118 – 124.
LIU Huanyu, CHEN Haitao, MIN Shiyao, et al. Tensile strength prediction for plant fiber mulch based on PSO – SVR[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(4): 118 – 124. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20170415&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2017.04.015. (in Chinese)
- [31] WU Xiaojuan, HUANG Qi, ZHU Xinjian. Thermal modeling of a solid oxide fuel cell and micro gas turbine hybrid power system based on modified LS – SVM[J]. International Journal of Hydrogen Energy, 2011, 36(1): 885 – 892.