

doi:10.6041/j.issn.1000-1298.2017.05.045

柑橘黑斑病反射光谱特性与染病果实检测方法研究

赵川源^{1,2} 何东健^{1,2} LEE Won Suk³

(1. 西北农林科技大学机械与电子工程学院, 陕西杨凌 712100; 2. 农业部农业物联网重点实验室, 陕西杨凌 712100;
3. 佛罗里达大学农业与生物工程学院, 盖恩斯维尔 FL 32611)

摘要: 通过对光谱仪采集的 340 ~ 1 030 nm 柑橘健康与感染黑斑区域光谱进行分析,在探明健康和黑斑病不同症状光谱特性的基础上,提出主成分分析结合特征排序的方法,选择出可识别染病与健康样本的最优波长(525 nm)建立 SMO 分类模型;基于序列浮动前向选择方法优选出 4 个特征波长(678、740、794、879 nm),建立 C4.5 算法识别柑橘黑斑病 3 种症状的方法。试验结果表明,用 525 nm 波长建立的 SMO 分类模型对健康和染病果样本的识别率达 99.37%,硬斑型、破裂型和黑斑型症状的识别率分别为 81.85%、71.88% 和 67.57%,3 种症状的平均识别率为 73.77%,比前人方法提高了 12.77 个百分点。

关键词: 柑橘黑斑病; 光谱特性; 特征选择; 识别模型; 主成分分析

中图分类号: S123 **文献标识码:** A **文章编号:** 1000-1298(2017)05-0356-07

Reflectance Spectral Characteristics of Black Spot Disease and Disease Detection Method for Citrus

ZHAO Chuanyuan^{1,2} HE Dongjian^{1,2} LEE Won Suk³

(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, Yangling, Shaanxi 712100, China
3. Department of Agricultural & Biological Engineering, University of Florida, Gainesville FL 32611, USA)

Abstract: Citrus black spot (CBS), which is one of the most common fungal diseases of citrus, causes lesions on the rind and early fruit drop before its mature stage. This disease can significantly reduce crop yield, making blemished fruit unsuitable for market. The objective of this research was to study the reflectance spectral characteristics of healthy and infected citrus fruits to identify diseased fruit from healthy ones. A portable USB2000 + spectrometer was used to acquire spectral reflectance of citrus fruit in the laboratory with wavelength ranged from 340 nm to 1 030 nm. However, the spectra contained thousands of wavelengths, and many of them would be considered as redundant, which may even decrease the classification accuracy. To reduce the data dimensionality and select the useful bands for further application, principal components analysis (PCA) and four band ranking methods, i.e., *T*-test, Kullback – Leibler distance, Chernoff bound and receiver operating characteristic (ROC) were applied. One important wavelength (525 nm) was selected and used to classify healthy and CBS infected fruits. Sequential minimal optimization (SMO), radical basis function network (RBF), and C4.5 classification methods were used to evaluate the performance of the selected band, and SMO achieved the highest accuracy of 99.37%. In order to compare the performance of classification accuracies according to optimal wavelengths selected by using different methods, two other methods, i.e., sequential floating forward selection (SFFS) and mutual information (MI), were applied. Wavelengths of 527 nm and 917 nm were selected based on SFFS, while the MI method selected 513 ~ 531 nm as the optimal wavelength range, and the highest recognition accuracy was 99.06%, which was lower than that of using 525 nm. Then SFFS was applied to find the optimal wavelengths for further distinguishing three CBS symptoms. C4.5 method was used to evaluate the performance of distinguishing CBS infected and healthy fruits based on selected wavelengths, and the highest overall classification accuracy was 73.77%.

Key words: citrus black spot disease; spectral characteristics; feature selection; identification model; principal components analysis

收稿日期: 2017-02-13 修回日期: 2017-03-17

基金项目: 国家高技术研究发展计划(863 计划)项目(2013AA102304)和国家自然科学基金项目(61473237)

作者简介: 赵川源(1988—),女,博士生,主要从事计算机视觉、图像处理和模式识别研究,E-mail: zhaochuanyuan@hotmail.com

通信作者: 何东健(1957—),男,教授,博士生导师,主要从事生物图像分析、智能化检测与控制研究,E-mail: hdj168@nwsuaf.edu.cn

引言

柑橘是美国佛罗里达州的重要果类,占全美柑橘总产量的 65%^[1]。在柑橘生产中,由柑橘球座菌 (*Guignardia citricarp*) 引起的柑橘黑斑病 (*Citrus black spot*, CBS) 是常见病害之一,几乎所有柑橘品种都会受到黑斑病的感染。虽然目前无报告显示黑斑病会对柑橘内部果实品质造成影响,但其影响果实的外观品质,从而影响果实在市场上的销售价格。黑斑病也会造成果实在成熟期前落果,严重影响柑橘产量,还会导致储藏期果实腐烂,对新鲜柑橘市场造成潜在影响^[2-4]。若能在果园中有效地控制黑斑病,及时移除被感染果实,则可减少柑橘生产损失。因此,寻求一种能够快速、无损检测黑斑病的方法具有实际意义。

机器视觉技术的应用对食品和农产品质量评价和无损检测提供了巨大的潜力^[5-10]。PYDIPATI 等^[11]提出在实验室环境下,利用基于 HSI 颜色空间的共生矩阵纹理特征识别不同柑橘叶片病害的方法,识别准确率最高达到 100%。BLASCO 等^[12]研究基于近红外、紫外线等机器视觉技术检测柑橘果实常见损伤,结合不同光谱信息对不同损伤类型果实进行分类,结果表明,与彩色图像检测结果相比,用近红外光谱图像可将缺陷的识别精度从 86% 提高到 95%。

黑斑病在果实表皮上有多种不同症状,使得对黑斑病的检测造成一定困难。硬斑型黑斑病是黑斑病中最常见且最典型的一种症状,它多呈凹陷的小圆形,周边呈红棕色而中间为褐色。破裂型黑斑病在果皮表面表现为大而平的深棕色病斑,有些会有凸起的裂痕。黑斑型黑斑病通常是多种病斑的融合,并且在果实表面占据很大部分^[13]。黑霉病症状类型的识别,对研究黑斑病菌诱因与黑斑病形态之间关系具有重要意义。近年来,基于光谱技术来检测食品和农产品内部或外部品质取得了不少成果^[14-19]。QIN 等^[20]获取 450 ~ 930 nm 的柑橘高光谱图像,用基于光谱信息散度 (SID) 分类方法对柑橘溃疡病进行检测,识别率达到 96.2%。BULANON 等^[21]研究了基于高光谱图像的柑橘黑斑病检测,根据波长比与波长差特性,选取了 493、629、713、781 nm 作为最佳波长,通过人工神经网络和线性判别分析检测黑斑病,检测准确率达到 92%。为了进一步探索不同症状柑橘黑斑病的检测,ZHAO 等^[22]应用高光谱图像对柑橘黑斑病不同表现症状进行检测,对 5 种症状在 396 ~ 1 010 nm 进行光谱特性分析后,通过特征选择选取最佳波长,黑

斑病的识别准确率可达到 98%,但对每种症状的识别率较低,只有 61% 左右。

高光谱图像技术可以较好地识别柑橘黑斑病,但各种症状的识别率较低,且高光谱设备价格昂贵,性能受复杂环境因素影响,使用条件要求严格使得高光谱系统难以在田间实际应用和推广,迫切需要研究能用于柑橘黑斑病田间实时检测的技术和方法。

本文以开发果园中实际应用的柑橘黑斑病实时检测多光谱技术为目标,用便携式光谱仪采集橘柑果实光谱数据,探明健康和黑斑病不同症状的光谱特性,研究有效识别健康与染病柑橘的模型与方法,以及 3 种常见黑斑病症状的识别方法,为果园快速检测柑橘黑斑病提供依据。

1 材料与方法

1.1 供试柑橘样本

供试柑橘于 2015 年 4 月在美国佛罗里达州 Immokalee 附近的果园采摘。135 个柑橘样本中,黑斑病感染的柑橘果实 105 个,未染病的健康柑橘果实 30 个。采摘的柑橘样本被送往佛罗里达大学精准农业实验室用于试验。

数据采集之前,用水对柑橘样本进行清洗,并将表面擦拭干净。为了后续研究方便,在每个染病柑橘的表面,随机选取最多 5 个病斑,并以病斑为中心,用黑色记号笔标记出直径约为 10 mm 的圆形。为了与染病样本进行对比,在每个健康柑橘的表面也随机标记出 5 个直径约 10 mm 的圆形。本文供试数据为标记的 150 个健康样本区域以及 487 个黑斑病区域,其中,487 个黑斑病样本中,硬斑型、破裂型和黑斑型样本分别为 270、32 和 185 个。标记后的健康果实与黑斑病中常见症状病斑样本图像如图 1 所示。

1.2 光谱数据采集

本研究采用 USB2000 + 微型光纤光谱仪 (Ocean Optics 公司) 获取柑橘的光谱反射值,光谱仪系统连接如图 2 所示。

USB2000 + 微型光纤光谱仪的光谱范围为 340 ~ 1 030 nm,光谱分辨率为 0.3 nm,光学分辨率为 0.1 nm。光谱仪由 USB 2.0 的接口与计算机连接,通过内径为 600 μm 的 QR600-7-SR-125F 型光纤反射探头 (Ocean Optics 公司) 与光谱仪连接,手持光纤反射探头与待检测物体接触,测得接触点 340 ~ 1 030 nm 的光谱信息,并通过计算机进行保存。

为了保证探头在接触物体表面时保持 45° 不

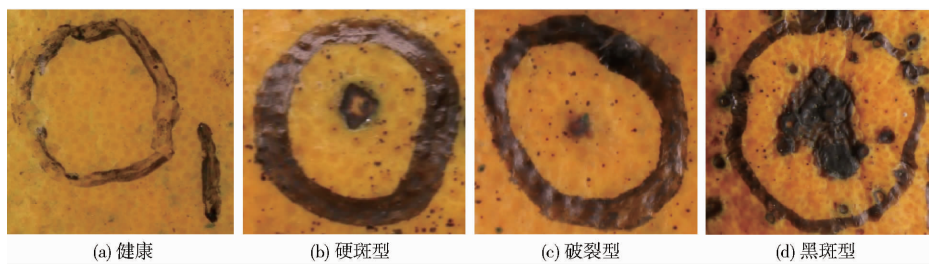


图 1 标记的健康果实和 3 种黑斑病症状示例

Fig. 1 Representative images for marked healthy fruit and fruits with three CBS symptoms

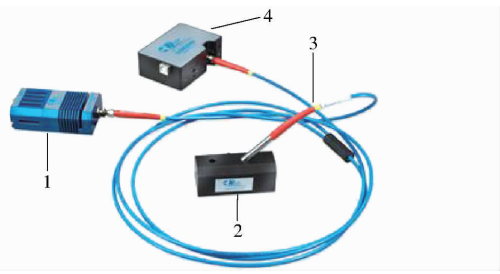


图 2 便携式 USB2000 + 微型光纤光谱仪

Fig. 2 Portable USB2000 + spectrometer

1. 光源 2. RPH-1 型反射探头架 3. 光纤反射探头 4. 光谱仪

变,使用 RPH-1 型反射探头架 (Ocean Optics 公司),将反射探头架与待检测物体表面相接触,再将光纤反射探头从 45°孔中深入,对物体进行光谱检测。在每次测量前,用反射率为 99% 的 SRS-99-020 标准参考板 (Labsphere 公司)对光谱仪进行校正。采集数据时,保持室内温度 25℃ 左右,相对湿度为 65% ~ 70%。为降低光谱检测中产生的噪声,在开始试验前 45 min,提前将光源 (LS-1 型卤钨光源, Ocean Optics 公司) 打开,以确保在试验过程中,光源处于平稳状态。

获取柑橘的光谱数据后,在 Matlab R2012a 和 怀卡托智能分析环境下,按照数据预处理、特征波段选择及分类器的构建等步骤,识别健康和染病果实,进一步选择特征波长并构建最优分类器以识别染病果实的不同症状。

1.3 数据预处理

由 USB2000 + 微型光纤光谱仪采集到的柑橘样本光谱数据,包含 340 ~ 1 030 nm 区间共 2 048 个波长。为了减少仪器及外界环境等因素对光谱数据的影响,舍去前 160 个噪声较大波长的光谱值,用剩余 400 ~ 1 030 nm 的 1888 个波长的光谱值进行分析。

为消除光谱数据中的噪声,需要对光谱数据进行平滑处理。考虑到 Savitzky - Golay 平滑滤波器^[23]可在消除光谱数据噪声的同时,保持光谱形状和相应的光谱特性,故本文采用 Savitzky - Golay 平滑滤波器对光谱数据进行平滑去噪。该方法基于低维线性最小二乘法多项式,实现相邻连续数据点的拟合。根据经验,应用四阶最小二乘多项式,选择平

滑点左、右各 20 个点。

1.4 特征波段选择

光谱数据包含丰富的信息含量,但是光谱各波段之间信息的关联性导致冗余信息也相应较多。为了准确识别健康与染病样本,需要特征波段选择等降维处理,以减少光谱数据的特征维数,剔除不相关或冗余的特征,降低计算时间,增强模型的泛化能力,提高不同染病症状的识别率。常用特征波段选择方法主要有主成分分析 (Principal component analysis, PCA)、序列浮动前向选择 (Sequential floating forward selection, SFFS) 及互信息 (Mutual information, MI) 等,考虑到各种方法在特征选择时的局限性,本文通过分析不同方法选择特征的识别效果,探明适合柑橘黑斑病识别的特征波段。

(1) 主成分分析

主成分分析 (PCA) 通过正交线性变换将原始相关变量转换成一组不相关的新变量,即主成分。将预处理后 637 × 1 888 的 Temp 数据分为健康与染病 2 类样本作 PCA 分析,分别得到特征向量矩阵 COEF、对原始数据解析生成的新的数据 Score 以及新生成的数据对应的贡献率 (协方差矩阵的特征值) LATENT。通过计算数据 Temp 与新生成的主成分数据 Score 之间的相关系数,得到各主成分的载荷,选取前几个表示了大部分信息的主成分载荷做进一步的分析。

(2) 序列浮动前向选择

序列浮动前向选择算法 (SFFS) 是在序列前向选择 (SFS) 基础上的改进^[24-25]。结合本研究所处理的高维数据,这里选择 K 最近邻分类器 (KNN) 作为选择特征的评价函数, K 取值设置为 5。从空集开始,每当选择 1 个特征波长加入特征子集时,均用评价函数 KNN 判断其识别率,若新的特征子集的准确率高于未加入此特征且此时的准确率最优,则确定将该特征加入特征子集,否则剔除该特征,继续评估下一个特征波长。

在初步确定特征子集后,考虑到该算法是一种贪心算法,即在特征选择过程中,只能加入特征而不

能剔除已选择的、多余的特征,所以在最终确定特征子集前加入 1 步回溯,即排除初步选择的特征子集中的 1 个特征,找到最没有用处的那个特征波长。重复该步骤,直到评估分数优于之前根据相同大小的特征子集得到的最高准确率,最优子集选择停止。

(3) 互信息

互信息描述了 2 个系统之间的相关性,或互相包含信息的多少,是信息理论中一种衡量 2 个变量间相互关系的方法。互信息通过熵以及联合熵来反映对象之间信息的相互包容程度^[26]。

熵(Entropy)反映系统本身的总能量,其表达式为

$$H(X) = - \sum_{j=1}^N p(x_j) \lg p(x_j) \tag{1}$$

其中 $X = \{x_1, x_2, \cdots, x_N\}$

式中 X ——随机变量

x_j ——系统中变量中的元素

$p(x_j)$ ——变量出现的概率

联合熵反映了随机变量的相关性,设 2 个随机变量 X 和 Y ,则 X 和 Y 的联合信息熵 $H(X, Y)$ 表示为

$$H(X, Y) = - \sum p_{X, Y}(x, y) \lg p_{X, Y}(x, y) \tag{2}$$

式中 x, y ——随机变量 X, Y 各取值的概率

$p_{X, Y}$ ——随机变量出现的概率

对随机变量 T_a, T_e 来说,其互信息表示为

$$M_I(T_a, T_e) = H(T_a) + H(T_e) - H(T_a, T_e) \tag{3}$$

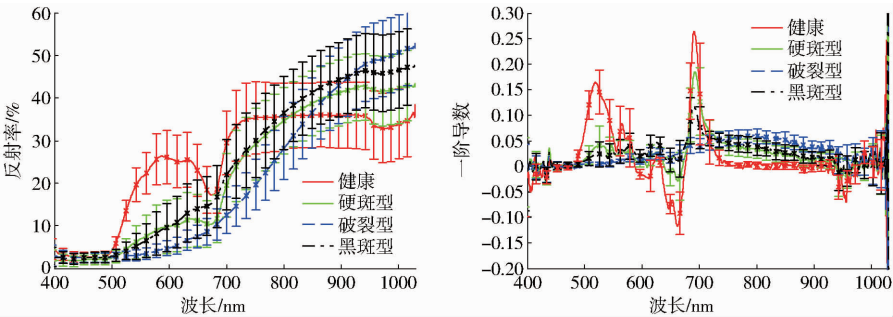


图 3 健康和黑斑病染病样本光谱特性分析

Fig.3 Spectral characteristics analysis of healthy fruit and fruits with three CBS symptoms

从图 3 中可以看出,在 400 ~ 800 nm 光谱范围内,健康样本表面的光谱反射率较高,而黑斑病染病样本的光谱反射率则较低。健康样本的反射率在 740 nm 后几乎保持不变,而在 800 nm 以后,黑斑病的光谱反射率高于健康样本。从标准差可以看出,在 500 ~ 930 nm 光谱范围内,健康和 CBS 染病样本的光谱反射率相差较大。从图 3 可以看出,在 500 ~ 710 nm 光谱范围内,健康和黑斑病光谱反射率的一阶导数差别较大;在 500 ~ 590 nm 和 690 ~

式中 $H(T_a), H(T_e)$ ——变量 T_a, T_e 的熵

互信息的特征选择方法是一种衡量特征值与对应类别的依赖关系,是 2 个特征之间距离的一种非线性度量方法。根据每个特征的概率分布,对特征进行量化。本研究中设特征量化等级为 15,计算量化后的特征与对应类别的互信息,并按照互信息的降序输出最终特征波长的选择结果。

1.5 分类器的构建

选择出区分健康和不同症状染病样本的特征波长后,需要构建分类器进行识别。构建分类器的方法很多,考虑到样本的维数较多,需要处理的数据量较大,而且是非线性可分问题,因此本文用序列最小优化算法 (Sequential minimal optimization, SMO)^[27]、径向基网络 (Radical basis function, RBF)^[28]、数据挖掘 C4.5 算法^[29-30] 构建分类器,对健康和染病样本进行分类,通过分类性能对比,优选出适合柑橘黑斑病检测的分类器。考虑到供试样本数量有限,故分类过程中采用 10 折交叉验证方法,以增强分类器的鲁棒性。

2 结果与分析

2.1 光谱特性分析

从 135 个供试柑橘中共获得 487 个黑斑病染病样本数据和 150 个健康样本数据,对样本数据进行平滑滤波预处理后,分别计算各波长下健康和染病样本的平均反射率,2 类样本在 400 ~ 1 030 nm 范围内的光谱特性如图 3 所示。

930 nm 范围内,一阶导数值为正,健康样本在 525 nm 和 694 nm 处一阶导数分别达到局部最大;在 590 ~ 690 nm 区域内,健康样本的一阶导数值为负;在 930 nm 之后,健康与黑斑病样本的一阶导数值曲线几乎重叠,表明在该光谱区域内健康和黑斑病的光谱反射率变化趋势相差不多,对应的反射率曲线,2 类样本在该区域的反射率均值重叠较多。

由图 3 可知,3 种黑斑病症状的光谱反射率曲线与健康样本的差别较大,破裂型症状的光谱特性

曲线与其他 2 类(硬斑型和黑斑型)有较为明显的区别,在整个光谱范围内,该症状的光谱反射率整体呈现上升趋势,且没有拐点。而硬斑型和黑斑型的光谱反射率曲线较为相似,结合图 1 中的彩色图像,可以看出这 2 类病状表现相似。

对健康和染病样本光谱反射特性的分析表明,用位于 500 ~ 930 nm 的光谱区间特征波长,可以对健康和黑斑病样本进行识别。

2.2 特征波长选择结果

分别用主成分分析、基于序列浮动前向选择和互信息等方法选择特征波长,基于序列最小优化(SMO)、径向基函数网络(RBF)和 C4.5 算法建立分类模型,以检测不同特征波长选择方法所选波长的识别效果,并采用 10 折交叉验证方法以增强分类器的鲁棒性。

2.2.1 健康与染病样本的特征波长选择结果

(1)基于 PCA 的特征波长选择结果

对健康和染病柑橘 2 类样本在 400 ~ 1 030 nm 的平均反射率进行 PCA 分析,其结果如图 4 所示。图 4a 为 487 个黑斑病样本和 150 个健康样本经过主成分分析后,在前 3 个主成分空间的分布情况,由图可以看出,用 3 个 PCA 分量即可将健康与染病样本分为 2 类。前 3 个主成分包含了超过全部数据 97.4% 的变量,故用前 3 个主成分代替全部数据。

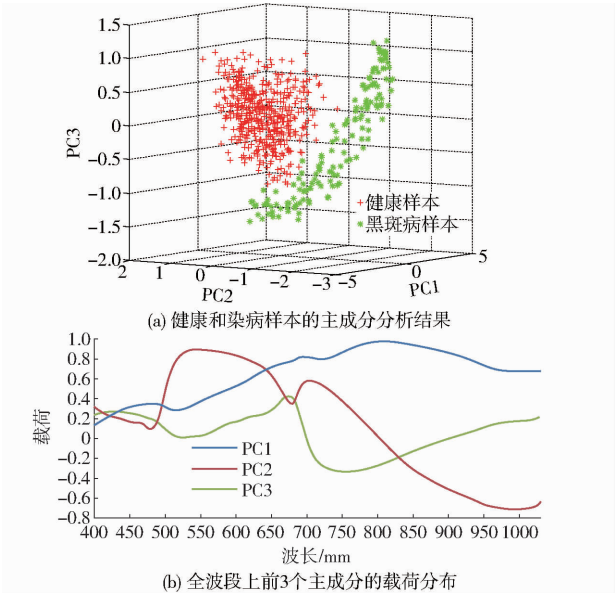


图 4 主成分分析结果

Fig.4 Results of PCA analysis

虽然前 4 个主成分包含了超过 97% 的数据,但本研究的最终目的是利用选择的若干个特征波长,构建价格较低的多光谱系统,以降低检测仪器的成本。而主成分分析仅将原始数据进行了压缩,故需要进一步从全部数据中选择少而优的特征波长。

图 4b 为前 3 个主成分在光谱范围内的载荷情况,载荷值反映了主成分在光谱范围内不同波长时的相关系数,图中呈现出波峰或波谷极值点对应的波长即为待选的特征波长。根据图 4b 共选出 525、675、700、740 nm 4 个特征波长。

为进一步优化选择出的最优波长,基于 *T* 检验、KL 散度、切诺夫界和 ROC 曲线共 4 种不同评价标准的特征排序方法,对选择出的特征波长子集进行分析。为便于对比将各方法得到的相关系数归一化为 0 ~ 1 范围,用衡量特征波长对分类重要性的相关系数进行排序,最终选择出最优的特征波长为 525 nm。

(2)基于 SFFS 的特征波长选择结果

通过 SFFS 特征选择方法,对健康和染病样本进行特征波长的选择。通过 10 次选择后,最终选择出 2 个最优特征波长,分别为 527 nm 和 917 nm。

(3)基于互信息的特征波长选择结果

对 2 类数据采用 MI 方法,得分前 10 的波长分布在 513 ~ 531 nm 区间内,因此选取 513 ~ 531 nm 光谱区域作为 MI 方法选择的特征波长区域。

2.2.2 黑斑病 3 种不同症状的特征波长选择结果

通过对黑斑病 3 种不同症状的光谱特性进行观察研究,破裂型与其他 2 种症状的光谱特性完全不同,故先将硬斑型和黑斑型看作同一类,因此问题变成了 2 类样本的分类识别。采用 SFFS 特征选择方法选择特征波长,并选择出 740、794、897 nm 3 个特征波长,用以区分破裂型和其他症状。

为对硬斑型和黑斑型进行分类,采用 SFFS 方法选择区分该 2 类病症的特征波长为 678 nm。因此,区分 3 种不同症状的特征波长为 678、740、794、897 nm 共 4 个特征波长。

2.3 识别结果及分析

2.3.1 健康与黑斑病样本识别结果

基于 PCA 和特征排序方法选择的特征波长 525 nm 和全波段的 3 种分类器,对健康和黑斑病染病样本的识别结果如表 1 所示。由表 1 可知,用 525 nm 特征波长建立的 SMO 模型分类准确率最高(99.37%),RBF 和 C4.5 模型分类准确率分别为 99.21% 和 99.06%,也取得了较好的识别效果。采用全波段建立的 SMO、RBF 和 C4.5 模型对健康和染病样本识别准确率分别为 99.84%、97.80% 和 98.59%,SMO 的识别准确率略高于单一波长,但采用单一波长与全波段相比,SMO 分类器处理时间从 0.2 s 减少到 0.01 s。而 RBF 和 C4.5 分类器基于单一波长的识别准确率均略高于基于全波段的识别率。

通过 SFFS 和 MI 2 种特征选择方法选择特征波长后,分别用 SMO、RBF 和 C4.5 分类器构建识别模型,对健康和黑斑病样本分类,结果如表 2 所示。

表 1 525 nm 特征波长和全波段建立的 3 种分类器识别结果

Tab.1 Identification results of healthy and CBS infected fruits using three classifiers based on selected wavelength of 525 nm and whole range of spectral data

分类方法		525 nm			全波段		
		健康	染病	识别率/%	健康	染病	识别率/%
SMO	健康	150	0	99.37	150	0	99.84
	染病	4	483		1	486	
RBF	健康	148	2	99.21	143	7	97.80
	染病	3	484		7	480	
C4.5	健康	149	1	99.06	143	2	98.59
	染病	5	482		7	485	

表 2 SFFS 和 MI 选择的特征波段对健康和染病样本的分类结果

Tab.2 Classification results of healthy and CBS infected fruits using three classifiers based on selected wavelengths with SFFS and MI methods

分类方法		SFFS(527 nm, 917 nm)			MI(513 ~ 531 nm)		
		健康	染病	识别率/%	健康	染病	识别率/%
SMO	健康	148	2	98.27	143	7	97.02
	染病	9	478		12	475	
RBF	健康	148	2	99.06	144	6	97.17
	染病	4	483		12	475	
C4.5	健康	145	5	98.12	135	15	95.92
	染病	7	480		11	476	

由表 2 可以看出,由 SFFS 方法选择的 2 个特征波长 527 nm 和 917 nm,对健康和黑斑病的识别准确率均高于 MI 的特征子集,RBF 分类识别率最高达到 99.06%,最低的识别率为基于 MI 特征子集的 95.92%。

对比表 1 和表 2 可知,本文提出的 PCA 载荷结合特征排序方法选择的 525 nm 特征波长,对健康和黑斑病染病样本的识别准确率高,优于 SFFS 和 MI 方法所选特征子集的识别率,且数据维数更低,即仅用 1 个特征波长 525 nm 就可将健康和黑斑病样本区分出来。

2.3.2 黑斑病不同症状的分类结果

为了识别破裂型、硬斑型和黑斑型 3 种黑斑病症状,用 SFFS 方法依次选择出能区分破裂型、硬斑型和黑斑型的特征波长,结合健康和黑斑病特征波长 525 nm 与 678、740、794、897 nm,采用数据挖掘 C4.5 算法对健康与黑斑病 3 种症状识别结果的混淆矩阵如表 3 所示。

从表 3 可以看出,健康样本的识别率最高,达到 98.67%,150 个样本中有 2 个样本被误识别为硬斑型样本。对 3 种症状的分类识别中,硬斑型的识别率最高,达到 81.85%,破裂型和黑斑型的识别率均较低,分别为 71.88% 和 67.57%,3 种病害症状平均识别率为 73.77%,比文献[21]采用高光谱分析方法(61.00%)提高了 12.77 个百分点。3 种症状

表 3 基于 5 个特征波长对健康和黑斑病 3 种症状识别结果的混淆矩阵

Tab.3 Confusion matrix of accuracy rates using selected five optimal wavelengths for healthy fruit and fruits with three CBS symptoms classification

	分类结果				总数
	健康	破裂型	硬斑型	黑斑型	
健康	148 (98.67%)	0	2	0	150
破裂型	0	23 (71.88%)	3	6	32
硬斑型	2	11	221 (81.85%)	36	270
黑斑型	0	5	55	125 (67.57%)	185

分类识别率并不理想,这是因为硬斑型和黑斑型的光谱反射特性曲线(图 3)趋势相似,故难以高精度识别;破裂型样本数较少(32 个),虽然在分类过程中采用了 10 折交叉验证方法,但样本数仍太少,使得分类准确率较低。

3 结论

(1)提出主成分分析结合特征排序选择最优波长构建 SMO 分类器识别健康和染病样本的方法,试验结果表明,用 525 nm 波长 SMO 分类器对健康和染病样本识别率为 99.37%,仅比基于全波段的分

类识别率低 0.47 个百分点,但分类器处理时间从 0.2 s 减少到 0.01 s,大大提高了识别效率。

(2)用 SFFS 和 MI 特征选择方法分别选择出 2 个特征波长(527 nm 和 917 nm)和 513 ~ 531 nm 光谱区域,对健康和黑斑病的识别率均低于 PCA 结合特征排序方法选择的最优波长的识别率,但也达到

满意的识别准确率。

(3)用优选的特征波长建立 C4.5 算法分类器对健康果和黑斑病 3 种症状样本进行识别,健康样本识别率达 98.67%,黑斑病 3 种症状的平均识别率为 73.77%,比前人方法提高了 12.77 个百分点。

参 考 文 献

- 1 Florida Department of Agriculture and Consumer Services. Florida citrus statistics 2009—2010[M]. Maitland, FL: USDA NASS Florida Field Office, 2011.
- 2 CHOI D, LEE W S, EHSANI R, et al. A machine vision system for quantification of citrus fruit dropped on the ground under the canopy[J]. Transactions of the ASABE, 2015, 58(4):933–946.
- 3 唐淬,周常勇. 柑橘黑斑病的发病特点及检测方法研究进展[J]. 中国南方果树, 2012, 41(3):55–60.
TANG Cui, ZHOU Changyong. Advances in occurrence characteristics and detection methods of citrus black spot[J]. South China Fruits, 2012, 41(3): 55–60. (in Chinese)
- 4 王兴红,陈国庆,李叶红. 中国柑橘黑斑病相关叶点霉属(Phyllosticta)真菌种类研究[C]//中国菌物学会第五届会员代表大会暨 2011 年学术年会论文摘要集, 2011.
- 5 彭辉,文友先,吴兰兰,等. 采用自适应 canny 算子的树上柑橘图像边缘检测[J]. 计算机工程与应用, 2011, 47(9):163–166.
PENG Hui, WEN Youxian, WU Lanlan, et al. Citrus edge detection based on self-adaptive canny operator[J]. Computer Engineering and Applications, 2011, 47(9):163–166. (in Chinese)
- 6 AL Ohali Y. Computer vision based date fruit grading system: design and implementation[J]. Journal of King Saud University—Computer and Information Sciences, 2011, 23(1): 29–36.
- 7 宋怀波,何东健,潘景朋. 基于凸壳理论的遮挡苹果目标识别与定位方法[J]. 农业工程学报, 2012, 28(22):174–180.
SONG Huaibo, HE Dongjian, PAN Jingpeng. Recognition and location methods of occluded apples based on convex hull theory[J]. Transactions of the CSAE, 2012, 28(22):174–180. (in Chinese)
- 8 MIZUSHIMA A, LU R. An image segmentation method for apple sorting and grading using support vector machine and Otsu's method[J]. Computers and Electronics in Agriculture, 2013, 94: 29–37.
- 9 LU J, SANG N. Detecting citrus fruits and occlusion recovery under natural illumination conditions[J]. Computers and Electronics in Agriculture, 2015, 110: 121–130.
- 10 刘燕德,肖怀春,孙旭东,等. 基于高光谱成像的柑橘黄龙病无损检测[J/OL]. 农业机械学报, 2016, 47(11):231–238.
http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20161132&flag=1. DOI:10.6041/j.issn.1000-1298.2016.11.032.
LIU Yande, XIAO Huaichun, SUN Xudong, et al. Non-destructive detection of citrus Huanglong disease using hyperspectral image technique[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(11): 231–238. (in Chinese)
- 11 PYDIPATI R, BURKS T F, LEE W S. Identification of citrus disease using color texture features and discriminant analysis[J]. Computers and Electronics in Agriculture, 2006, 52(1): 49–59.
- 12 BLASCO J, ALEIXOS N, GOMEZ J, et al. Citrus sorting by identification of the most common defects using multispectral computer vision[J]. Journal of Food Engineering, 2007, 83(3):384–393.
- 13 DEWDNEY M M, PERES N A, RITENOUR M, et al. Citrus black spot in Florida[J]. Citrus Industry, 2010, 91:18–21.
- 14 洪添胜,李震,吴春胤,等. 高光谱图像技术在水果品质无损检测中的应用[J]. 农业工程学报, 2007, 23(11): 280–285.
HONG Tiansheng, LI Zhen, WU Chunyin, et al. Review of hyperspectral image technology for non-destructive inspection of fruit quality[J]. Transactions of the CSAE, 2007, 23(11): 280–285. (in Chinese)
- 15 蔡健荣,王建黑,黄星奕,等. 高光谱图像技术检测柑橘果锈[J]. 光电工程, 2009, 36(6):26–30.
CAI Jianrong, WANG Jianhei, HUANG Xingyi, et al. Detection of rust in citrus with hyperspectral imaging technology [J]. Opto-Electronic Engineering, 2009, 36(6): 26–30. (in Chinese)
- 16 BENXUE M A, XIAO W, QU N, et al. Detection of fruits slight bruises based on hyperspectral imaging technology[C]//2011 International Conference of Electronic Control Engineering, 2011: 4445–4448.
- 17 LI H, LEE W S, WANG K, et al. “Extended spectral angle mapping (ESAM)” for citrus greening disease detection using airborne hyperspectral imaging[J]. Precision Agriculture, 2014, 15(2):162–183.
- 18 欧阳爱国,吴建,刘燕德. 高光谱成像在农产品无损检测中的应用[J]. 广东农业科学, 2015, 42(23): 164–171.
OUYANG A G, WU J, LIU Y D. Application of hyperspectral imaging in nondestructive detection of agricultural products[J]. Guangdong Agricultural Sciences, 2015, 42(23): 164–171. (in Chinese)
- 19 ZHANG C, GUO C, LIU F, et al. Hyperspectral imaging analysis for ripeness evaluation of strawberry with support vector machine[J]. Journal of Food Engineering, 2016, 179:11–18.

14 MEI H, LU X. The quantitative description between zeta potential and fluorescent particle adsorption on Cu surface [J]. International Journal of Dermatology, 2014, 46(1):56 – 60.

15 STENTIFORD G D, BATEMAN K S, FEIST S W, et al. Areospora rohanae n. gen. n. sp. (Microsporidia; Areosporiidae n. fam.) elicits multi-nucleate giant-cell formation in southern king crab (*Lithodes santolla*) [J]. Journal of Invertebrate Pathology, 2014, 118:1 – 11.

16 李青林, 毛罕平, 左志宇, 等. 基于叶片微观结构变化的番茄磷水平检测方法[J/OL]. 农业机械学报, 2014, 45(10): 282 – 287. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20141044&flag=1. DOI:10.6041/j.issn.1000-1298.2014.10.044.

LI Qinglin, MAO Hanping, ZUO Zhiyu, et al. A detection method of tomato phosphorus level based on micro-structure of leaf[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(10): 282 – 287. (in Chinese)

17 绍云, 姜丽娜, 李万昌, 等. 砷、铅胁迫对小麦幼苗毒害效应及叶片下表皮扫描电镜观察[J]. 西北农业学报, 2009, 18(1): 133 – 138.

SHAO Yun, JIANG Li'na, LI Wanchang, et al. Toxic effects of As and Pb to wheat seedling and scanning electron microscopic observation on nether epidermis of leaves [J]. Acta Agriculturae Boreali-occidentalis Sinica, 2009, 18(1):133 – 138. (in Chinese)

18 TAN K, CHAI Y, SONG W, et al. Identification of soybean seed varieties based on hyperspectral image[J]. Transactions of the CSAE, 2014, 30(9):235 – 242.

19 孙俊, 周鑫, 毛罕平, 等. 基于 PDWT 与高光谱的生菜叶片农药残留检测[J/OL]. 农业机械学报, 2016, 47(12):323 – 329. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20161240&flag=1. DOI:10.6041/j.issn.1000-1298.2016.12.040.

SUN Jun, ZHOU Xin, MAO Hanping, et al. Detection of pesticide residues on lettuce leaves based on piece-wise discrete wavelet transform and hyperspectral data[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(12):323 – 329. (in Chinese)

20 郭志明, 赵春江, 黄文倩, 等. 苹果糖度高光谱图像可视化预测的光强度校正方法[J/OL]. 农业机械学报, 2015, 46(7): 227 – 232. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20150733&flag=1. DOI:10.6041/j.issn.1000-1298.2015.07.033.

GUO Zhiming, ZHAO Chunjiang, HUANG Wenqian, et al. Intensity correction of visualized prediction for sugar content in apple using hyperspectral imaging[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(7):227 – 232. (in Chinese)

~~~~~

(上接第 362 页)

20 QIN J, BURKS T F, RITENOUR M A, et al. Detection of citrus canker using hyperspectral reflectance imaging with spectral information divergence[J]. Journal of Food Engineering, 2009, 93(2):183 – 191.

21 BULANON D M, BURKS T F, KIM D G, et al. Citrus black spot detection using hyperspectral image analysis[J]. Agricultural and Engineering International, 2013, 15(3):171 – 180.

22 ZHAO C, LEE W S, HE D. Citrus black spot detection based on selected wavelengths using hyperspectral images[C]. ASABE, 2015 Annual Meeting Papers 152181190, 2015.

23 SAVITZKY A, GOLAY M J E. Smoothing and differentiation of data by simplified least squares procedures[J]. Analytical Chemistry, 1964, 36(8):1627 – 1639.

24 SOMOL P, PUDIL P, NOVOTI J. Adaptive floating search methods in feature selection[J]. Pattern Recognition Letters, 1999, 20(11 – 13):1157 – 1163.

25 PUDIL P, FERRI F, NOVOTI J, et al. Floating search methods for feature selection with nonmonotonic criterion functions [C]//Proceedings of the 2nd IAPR International Conference on Pattern Recognition, 1994, 2: 279 – 283.

26 PENG H, LONG F, DING C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(8):1226 – 1238.

27 PLATT J C. Sequential minimal optimization: a fast algorithm for training support vector machines[M]//Advances in kernel methods—support vector learning. Cambridge: The MIT Press, 1998: 212 – 223.

28 MOODY J, DARKEN C. Fast learning in networks of locally-tuned processing units[J]. Neural Computation, 1989, 294(1): 281 – 294.

29 程龙, 蔡远文. 数据挖掘 C4.5 算法的编程设计与增量学习改进[J]. 计算技术与自动化, 2009, 28(4): 83 – 87.

CHENG Long, CAI Yuanwen. Program design and an improved incremental learning method of C4.5 algorithm[J]. Computing Technology and Automation, 2009, 28(4): 83 – 87. (in Chinese)

30 高阳, 廖家平, 吴伟. 基于决策树的 ID3 算法与 C4.5 算法[J]. 湖北工业大学学报, 2011, 26(2):54 – 56.

GAO Yang, LIAO Jiaping, WU Wei. ID3 algorithm and C4.5 algorithm based on decision tree[J]. Journal of Hubei University of Technology, 2011, 26(2): 54 – 56. (in Chinese)