

考虑直径影响的苹果霉心病透射光谱修正及检测

张海辉^{1,2} 田世杰^{1,2} 马敏娟^{1,3} 赵娟^{1,3} 张军华^{1,2} 张佐经^{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: 针对苹果霉心病近红外透射光谱信息受果实直径影响的难题, 提出了一种能够修正果实直径对透射光谱影响的方法。基于透射光谱采集平台获取 327 个红富士苹果的可见/近红外光谱(350 ~ 1 100 nm)信息, 采用电子游标卡尺获取其直径(光程)信息。以直径为 80 mm 健康苹果的平均光谱作为参考光谱, 将 327 个苹果的光谱与参考光谱进行比较, 结合直径信息利用公式求得透射光在果实内的衰减系数 P , 用衰减系数 P 进行透射光谱的修正。修正后光谱建立支持向量机(SVM)模型、误差反向传播神经网络(BP-ANN)模型, 并与修正前原始光谱建立模型进行对比。实验结果表明, 应用此光谱修正方法能够显著提高模型判别精度, 其中应用 SVM 算法对修正后的光谱建立模型效果最好, 对训练集和测试集的判别准确率分别为 99.34% 和 90.20%, 相对于原始光谱建立的模型判别准确率分别提高了 7.84 和 5.89 个百分点。基于此方法修正果实直径对于透射光谱的影响是可行的, 构建的模型能够实现苹果霉心病的准确判别。

关键词: 苹果; 霉心病; 近红外光谱; 光谱修正

中图分类号: O657.33; S436.611 **文献标识码:** A **文章编号:** 1000-1298(2019)01-0313-08

Detection Method of Moldy Core in Apples Using Modified Transmission Spectrum Based on Size of Fruit

ZHANG Haihui^{1,2} TIAN Shijie^{1,2} MA Minjuan^{1,3} ZHAO Juan^{1,3} ZHANG Junhua^{1,2} ZHANG Zuoqing^{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: Currently, the near infrared transmission spectrum of moldy core in apples was seriously affected by the size of fruit. In order to solve the problem, a transmission spectrum correction method based on size of fruit was proposed. A spectrum acquisition platform was constructed to acquire the transmission spectra (350 ~ 1 100 nm) of 327 Fuji apples and their diameters were measured with a vernier caliper. The spectrum of healthy apples with diameter of 80 mm was used as reference. Comparing the spectrum of 327 apples with the reference spectrum, a formula was built. The attenuation index of transmitted light in the fruit can be easily found by using the formula and diameters. Then the transmission spectrum was modified with the help of attenuation index. Error back propagation artificial neural networks (BP-ANN) and support vector machine (SVM) measurement model were established based on corrected spectrum and original spectrum. The results showed that the accuracy of the models based on corrected spectrum was much higher than those of the others, and its recognition accuracy rate reached 99.34% for the training set and 90.20% for the test set. The recognition rate of the model was 7.84 and 5.89 percentage points higher than that of the original spectrum. The results showed that the effect of the size on transmission spectra can be corrected by this method, and the method had high identification accuracy. Meanwhile, the results would provide theoretical basis for the development of on-line detection of internal quality in apples and provide a new idea for the study of internal disease detection models for different agricultural products.

Key words: apples; moldy core; near infrared spectrum; spectral correction

收稿日期: 2018-07-20 修回日期: 2018-08-29

基金项目: 国家自然科学基金项目(31701664)、陕西省重点研发计划项目(2017ZDXM-NY-017)和陕西省科技统筹创新工程项目(2016KTCQ02-14)

作者简介: 张海辉(1977—),男,教授,博士生导师,主要从事精准农业和农业物联网研究, E-mail: zhanghh@nwsuaf.edu.cn

0 引言

苹果霉心病是苹果果实的一种严重内部病害,严重影响果实的产量和品质,且在世界范围内广泛分布^[1]。苹果霉心病又被称为霉腐病、心腐病,由10余种病菌混合共生。病菌自花期开始至果实生长期都可侵染,其中以花期侵染率最高,发病症状主要表现在果实心室霉变(霉心型)、果实心室褐变(褐变型)、果心腐烂(心腐型)^[2]。近年来,苹果霉心病发病率普遍较高,一般为21%左右,但是其外部特征与正常苹果并无差异^[3]。随着市场对于高端果实的需求越来越高,霉心病苹果的检测识别成为苹果采后分级的一大难题。

国内外相关学者尝试采用生物阻抗特性、CT成像、低频磁共振和机器视觉等方法进行果品内部病害的检测^[4-8],但是近红外透射光谱无损检测技术凭借其分析速度快、低成本等优点在苹果内部病害检测方面得到广泛应用并获得认可。MCGLONE等^[9]对比两种不同的光谱扫描系统在线检测苹果的棕色核心,证明了近红外透射光谱检测苹果内部病害的可行性。CLARK等^[10]利用近红外透射光谱的方法检测布瑞本苹果的棕色核心,探究不同放置位置时的PLS模型效果,表明果轴水平放置判别效果较好。SHENDEREY等^[11]搭建了苹果霉心病在线检测系统,建立了PLS判别模型,判别精度达到了90.1%,效果较好。文献[12-14]同样利用近红外光谱的方法检测苹果霉心病,分别建立了各种不同的判别模型,识别效果均达到85%以上,并且找到了霉心病相关的特征波段。上述研究均在苹果大小均匀的样本理想状态下进行建模与验证,忽略了果实大小对透射光谱的影响,检测模型的实用性与先进性有待提高。

文献[15]进行霉心病识别时考虑了果实直径对透射光谱的影响,将直径作为变量和透射光谱一起建模。文献[16]将每个果实的光谱换算成同一直径果实的光谱,作为消除果实大小影响的方法。基于以上研究,本文提出一种考虑果实直径对霉心病透射光谱影响的方法,基于此方法进行光谱修正,建立模型与原始光谱模型进行对比,为不同大小苹果内部病害工业化识别提供一种新方法。

1 材料与方法

1.1 实验材料

于2017年10月在陕西省白水某果园随机选取不同大小、无机械损伤、无外部缺陷的红富士苹果,运回西北农林科技大学机械与电子工程学院。

从中选择外表完好的327个苹果于第2天进行实验,并采用标签对样本进行逐一编号。实验过程中保持室内温度稳定(20℃左右),相对湿度为30%~40%。

1.2 实验设备与光谱采集方法

基于实验室自主搭建的透射光谱数据采集平台进行数据采集(图1)。光源为QR111型单灯珠卤素灯(德国Osram公司),额定功率为50W,光谱范围250~3000nm,发光角度6°,光通量20000lm,由检测暗箱外部的光源控制器进行控制。光谱仪采用OFS-1100型地物光谱仪(美国Ocean Optics公司),采集光谱范围350~1100nm,光谱分辨率1.3nm。光谱仪由计算机进行控制,并通过SpectraSuite软件进行数据采集和分析工作。图1中检测暗箱长为600mm,宽为400mm,高为1000mm,为了消除暗箱内部反射光对检测结果的影响,用黑色的吸光布将其内侧全部覆盖。检测支架放于暗箱中间,其从上至下依次装有卤素光源、移动标尺、检测托台,准直镜与接收光纤。检测支架放于暗箱中间,其从上至下依次装有卤素光源、移动标尺、检测托台、准直镜与接收光纤。

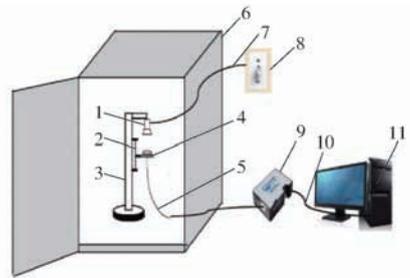


图1 光谱采集装置结构图

Fig. 1 Structure diagram of spectral collection device

1. 卤素光源 2. 移动标尺 3. 检测支架 4. 苹果检测托台 5. 准直镜与接收光纤 6. 检测暗箱 7. 光源控制器连接线 8. 光源控制器 9. 光谱仪 10. 数据连接线 11. 计算机

为保证光源稳定,每次采集光谱数据前15min打开光源进行预热工作。经前期预备实验可知,光源距离载物台表面130mm,SpectraSuite软件设置积分时间100ms,平均扫描次数为10,平滑度为5,去除暗噪声,添加非线性校正、杂散光校正时采集的光谱数据最优。采集数据时,保持暗箱门处于关闭状态,将苹果样本放置于苹果检测托台上,果柄轴向方向与光源照射方向垂直。每个样本于赤道面处选取均匀分布且无缺陷的3个点采集光谱信息,各点之间呈120°,最终取3次数据的平均值作为样本的光谱数据。

1.3 果实大小与霉心病病害信息的采集方法

苹果的透射光强和果实大小是霉心病检测最直接相关的参数,而果实大小主要由透射光程(果实横径)来体现。用G101-102-101型电子游标卡尺

于样本光谱采集点平行于赤道面方向测量果实横径,取 3 个点测量的平均值作为此样本的透射光程。

样本光谱信息与光程信息采集后,沿径轴横截面将样本切开,判断样本是否为霉心病果(图 2)。苹果发生霉心病病变后,无论发病程度如何均无法食用,故根据实际需求,进行苹果霉心病判定时仅划分为健康果与霉心病果。



(a) 健康果 (b) 轻度霉心病果 (c) 严重霉心病果

图 2 不同发病程度样本

Fig. 2 Extent of disease in different samples

1.4 光谱修正方法

光在苹果内部进行传输时,除了霉心病造成光谱衰减之外,光程(果实横径)造成的衰减也是十分严重的。文献[16]将不同直径果实的光谱换算到同一直径下,进而修正果实大小对蜜橘透射光谱的影响。本文在上述思想的基础上,提出了一种新的方法。将健康果肉看作均匀介质,忽略果皮对光散射的影响^[17],当平行光穿透果肉介质时,果肉对穿过其中的光有吸收和散射作用,分别表示为吸收系数和散射系数,而沿此方向上光的衰减系数 P 定义为吸收系数和散射系数之和^[18]。可以运用数学统计的方法近似求得光在健康果肉中透射时的衰减系数 P ,利用衰减系数 P 对不同大小果实的光谱信息进行修正,进而将不同直径果实的光谱换算到参考果实的直径下。

文献[19-20]提出光在果实内部会发生散射,果实越大,散射现象越严重。光在果实内部并不是线性衰减的,而是呈现随着果实横径增大逐渐加剧衰减的现象。用双曲正弦函数($\sinh(x)$ 函数)来模拟这种衰减现象,函数的 y 值随着 x 的增大离 $y=x$ 曲线越来越远(图 3),可以近似地模拟样本横径与参考果实横径差值越大导致光谱衰减差距越严重这一现象。首先,选定一类相同横径健康果的光谱作为参考光谱(以横径 80 mm 的苹果光谱作为参考光谱),其横径表示为参考横径。然后,将每一个健康样本的横径与参考横径作差,并用 $\sinh(x)$ 函数对得到的差值进行变换。同时,由于光在果肉中的散射程度是不可知的,所以引入函数系数 k ,通过对 k 进行寻优操作找到最能模拟其衰减规律的函数。接着,将每一个健康样本的光谱与参考光谱作差,并除以函数变换后的横径差值。最后,将每一个样本得到的衰减系数取平均值确定平均衰减系数 P 。当进

行不同大小苹果的霉心病判别时,进行上述过程的逆运算,从而利用衰减系数 P 进行光谱的修正,将不同大小果实的透射光谱转换为同一光程基准下的光谱,提高霉心病判别的准确率。相应的计算公式为

$$P = \frac{\sum_{i=1}^n \left(\frac{s_r^* - s_i}{\sinh \frac{d_r^* - d_i}{k}} \right)}{n} \quad (1)$$

$$S_i = \sinh \frac{d_r^* - d_i}{k} P + s_i \quad (2)$$

式中 S_i ——样本 i 修正后光谱强度

s_r^* ——参考光谱强度

d_r^* ——参考果实横径

s_i ——样本 i 光谱强度

d_i ——样本 i 果实横径

n ——样本总数

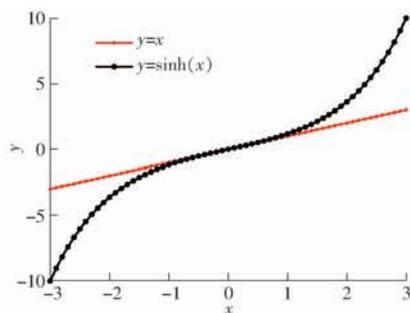


图 3 $\sinh(x)$ 曲线

Fig. 3 Diagram of $\sinh(x)$ curve

1.5 样本异常值检验与样本划分

通过数学分析软件 Matlab 2016a 进行采集光谱的分析工作。由于样本中健康果数量远多于霉心病果的数量,为了避免样本不均衡造成的模型过拟合现象,随机选择与病果数量相当的健康果进行模型的分析。数据预处理之前,首先对数据进行异常值的检验。基于主成分分析与马氏距离结合的方法可以很好地剔除掉样本中的异常值^[21]。用 Kennard - Stone (KS) 算法^[22-23]将光谱数据按照 3:1 的比例划分为建模集与验证集,能够保证训练集中样本按照空间距离均匀分布,使得建立的模型能够更好地反映样本信息。

1.6 近红外透射光谱的预处理与降维方法

由光谱仪光纤探头接收到的光谱中包含各种各样的噪声,因此在数据分析之前,应当根据实际情况,对原始光谱数据进行适当的预处理工作,以提高模型的收敛性能^[24-25]。通常用 Savitzky - Golay 卷积平滑法消除数据中的高频随机误差,多元散射校正(MSC)、变量标准化校正(SNV)、数据中心化校

正 (Mean centering) 对光谱数据进行位移校正以及多元散射校正。通过比较原始光谱和不同预处理方式下的模型判别准确率,找到适用于样本分析的较好的预处理方法。为了进一步提高模型收敛速度与模型精度,预处理环节还对数据进行了 $[0, 1]$ 区间归一化处理。

采用主成分分析 (Principal component analysis, PCA) 的方法对高维光谱数据进行降维操作。主成分分析^[26-27]是多元统计分析中用来分析数据的一种方法,它是用一种较少数量的特征对样本进行描述以达到降低特征空间维数的方法。主成分分析后得到的主成分信息能够反映原始变量的大部分信息。这些主成分通常表示为原始变量的线性组合,且各主成分之间互不相关。

1.7 建模方法

本文主要用两种方法来建立霉心病判别模型,分别是误差反向传播神经网络 (Error back propagation artificial neural networks, BP-ANN) 和支持向量机 (Support vectors machine, SVM)。误差反向传播神经网络 (BP-ANN) 是应用最广泛的一种人工神经网络。它基于仿生物学原理建立,具备良好的非线性映射能力、自组织自学习能力以及完备的联想能力^[28]。支持向量机是由 CORTES 等^[29]根据统计学习理论提出的原理性方法,可用于解决模式分类与非线性映射问题。其中 SVC (Support vectors classification) 模型可用于解决分类问题,其决策函数表示为

$$f(x) = \text{sgn} \left(\sum_{i=1}^l \alpha_i^* y_i K(x, x_i) + b^* \right) \quad (3)$$

式中 $K(x, x_i)$ ——核函数,本文采用 RBF 核函数对样本数据进行高维映射

l ——样本总数量

α_i^* ——Langrange 系数

y_i ——支持向量机的类别标签

b^* ——阈值

基于 LIBSVM 工具箱进行 SVM 核函数和相关参数 (主要是惩罚参数 c 和核函数参数 g) 的寻优,主要用网格寻优算法进行交叉验证选择最佳参数 c 与 g ,并采用最佳参数 c 与 g 对整个训练集进行训练获取支持向量机模型,利用获取的模型进行测试与预测。本文以判别准确率作为模型精度的评判标准。

2 结果与讨论

2.1 样本分布与光谱数据分析

样品分布均衡,果实横径在 75 ~ 85 mm 内的苹

果略多 (表 1)。对光谱仪采集到的光谱信号进行分析,发现样品的可见/近红外透射光谱在低于 550 nm 和高于 950 nm 的波段信号微弱且含有大量噪声,因此选择 550 ~ 950 nm 为样本光谱的有效范围 (图 4)。横坐标表示波长,纵坐标表示相应波长下光穿透苹果最终由光电传感器接收到的光强度。光谱在 643、714、804 nm 附近呈现明显波峰,尤其在 714 nm 附近光谱波动范围较大。这与 MCGLONE 等^[9]、雷雨等^[13]的研究一致,即 714 nm 周围波段与霉心病的相关性较高。

表 1 苹果在果径与病害方面的数量分布

Tab. 1 Number distribution of apples in terms of diameter and disease

样品类别	样本果径 d/mm					总计
	70 ~ 75	75 ~ 80	80 ~ 85	85 ~ 90	90 ~ 95	
健康果	23	74	93	39	5	234
霉心病果	6	44	30	13	0	93
总计	29	118	123	52	5	327

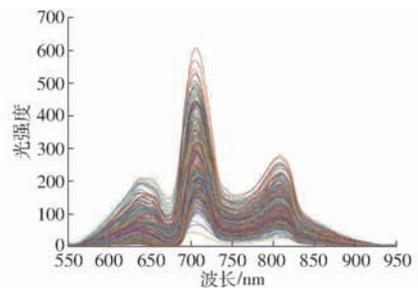


图 4 原始光谱数据

Fig. 4 Original spectral data

不同大小的苹果透射光谱波峰呈现明显的差异 (图 5)。将 70 mm 健康果与霉心病果 (70 mm 果表示果实横径在 70 ~ 75 mm 范围内)、75 mm 健康果与霉心病果 (果实横径在 75 ~ 80 mm 范围内)、80 mm 健康果与霉心病果 (果实横径在 80 ~ 85 mm 范围内)、85 mm 健康果与霉心病果 (果实横径在 85 ~ 90 mm 范围内) 的光谱信息进行对比,发现对于同一果径的苹果,714、804 nm 波峰处很容易将霉心病果与健康果进行区分。对于不同果径的健康果而言,其接收光谱强度随着直径的增大逐渐降低,这主要是光程对接收光谱的影响造成的。因此,仅考虑相同大小的苹果利用近红外透射光谱进行霉心病的判别,较为容易实现。但是将不同果径的苹果混在一起,则很难得到较好的霉心病判别模型,如果修正光程对于透射光谱的影响,则有望获得较好的霉心病判别模型。

2.2 光谱数据修正

为了提高模型精度,修正果实大小对于近红外透射光谱判别苹果霉心病的影响,用式 (1) 求光谱

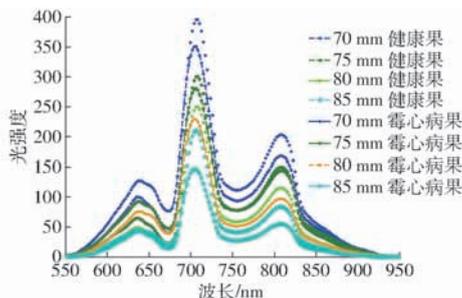


图 5 不同直径苹果的透射光谱图

Fig. 5 Transmission spectra of apples with different diameters

穿透果实的衰减系数 P 。以果径在 80 ~ 81 mm 的 20 个健康果光谱的平均值作为果径为 80 mm 健康果的参考光谱。将共 234 个健康果光谱代入式(1),对 k 进行寻优操作。首先,以 5 为步长进行 k 值初步寻优,即 k 取值 5、10、15、20,分别求取衰减系数 P ,并用此衰减系数对 70、75、80、85 mm 健康果进行光谱修正,寻找合适 k 值区间使得不同大小健康果修正后光谱图与参考光谱图差异最小,结果发现 k 在 5 ~ 10 区间时效果较好。然后,在该区间内以 0.5 为步长对 k 赋值,分别求取衰减系数 P 进行所有样品果光谱修正,并对修正后的光谱建立支持向量机模型,通过对比不同模型霉心病果判别准确率确定最优 k 值,结果表明 $k = 6.5$ 时效果最优。由图 6 可以发现,714 nm 和 804 nm 处光的衰减比较严重,该波段不仅与霉心病变化相关,而且也与光程变化相关,这也说明了仅考虑透射光谱建立模型来判别霉心病的精度是受到限制的。

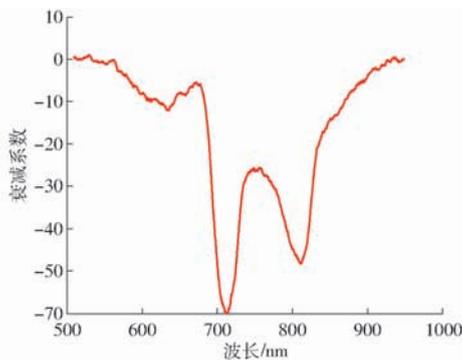


图 6 衰减系数曲线

Fig. 6 Curve of attenuation coefficient

通过式(2)用衰减系数 P 将 70、75、80、85 mm 健康果与 70、75、80、85 mm 霉心病果的光谱进行修正(图 7)。由图 7a 可以看到,不同果径的健康果经过修正后,其光谱与参考光谱大致相同,这样就把不同果径苹果的光谱转换到了同一基准下。而图 7b 修正后的霉心病果光谱波峰在波长 714 nm 附近显著低于参考光谱,这样把不同果径霉心病果的光谱转换到了同一基准下,建立霉心病判别模型可以得到较好的效果。

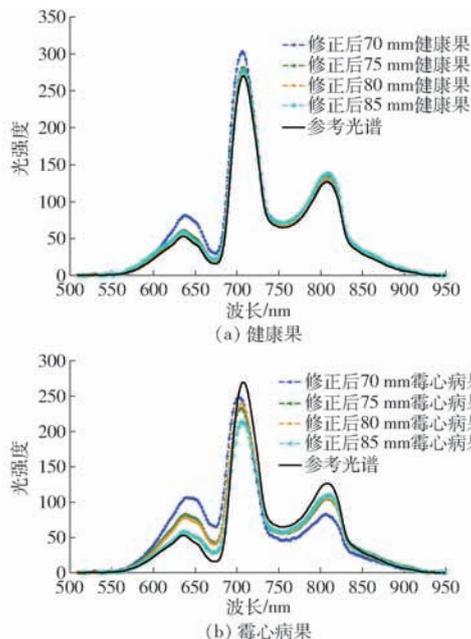


图 7 光谱修正图

Fig. 7 Spectral correction of apples

2.3 光谱数据预处理和样本划分

由于样本中健康果的数量远多于霉心病果的数量,为了避免模型分隔超平面过度偏向于健康果类导致泛化能力的降低,随机选择与霉心病果数量相当的健康果(115 个)进行数据分析,对选择的 115 个健康果进行马氏距离异常值检验,剔除掉 4 个异常样本,最终选定 204 个样本(111 个健康果、93 个霉心病果)进行接下来的数据分析。首先应用衰减系数对样本光谱进行修正,随后基于 Kennard - Stone 算法将样本按 3:1 划分为训练集与测试集。训练集 153 个样本(包括 75 个健康果与 78 个霉心病果),验证集 51 个样本(包括 36 个健康果与 15 个霉心病果)。

光谱经过 5 种不同预处理后,作为 SVM 模型的输入变量,建立霉心病预测模型(表 2)。建立模型之前,需要对预处理后的高维光谱数据进行主成分分析,以保证模型较好的判别性能与较小的计算量。一般为了提升分类器的性能,会选择增加一些主成分特征。但是当主成分特征达到一定数量后,增加特征不仅不能提高分类精度反而会导致分类精度的下降。所谓维灾难就是随着特征维度的增加计算成本急剧增加,而分类精度下降(或者聚类质量下降)的现象。为了探明不同预处理方式下的最优主成分数,分别选择累计贡献率达到 90%、93%、95%、98%、99% 的主成分数作为模型输入的特征变量,建立 SVM 判别模型进行对比,以期找到最优的预处理方式与最佳的主成分数。表 2 中可以看到,经过修正后的光谱进行归一化处理,选择累计贡献率达到

90%的前8个主成分建立模型效果最好,其对于训练集和测试集的苹果病害判别准确率分别为99.34%和90.20%。

表2 不同预处理方式下的模型效果

Tab.2 Model effect under different preprocessing methods

数据处理方式	最佳主成分数	SVM 模型判别准确率/%	
		训练集	测试集
修正后光谱-归一化	8	99.34	90.20
修正后光谱-多元散射校正(MSC)-归一化	10	88.89	84.31
修正后光谱-变量标准化校正(SNV)-归一化	8	88.89	80.39
修正后光谱-数据中心化校正-归一化	8	96.73	90.20
修正后光谱 Savitzky - Golay 卷积平滑-归一化	7	92.16	86.27

2.4 建模方法与建模结果的讨论

2.4.1 建模参数的选择与设定

用支持向量机(SVM)算法进行建模时,需考虑惩罚因子 c 和核函数参数 g 的选择。本文中选用RBF函数作为核函数,根据留一交互原则进行 c 和 g 的寻优。针对PCA选取的前8个主成分确定了SVM参数,惩罚因子 c 为4.00,核函数参数 g 为0.02。

用BP神经网络建模时,隐含层神经元传递函数选择正切对数“tansig”,输出层传递函数选择为线

性函数,训练函数选择LM训练法。学习率设置为0.05,目标误差设置为 10^{-6} ,最大训练次数1000次,每训练50次显示一次。由于隐含层节点神经元数目并没有明确的理论指导,本文结合经验公式并通过反复试验,最终确定了网络结构参数。输入层节点数设为8,隐含层节点数设为8,输出层节点数设为1。

2.4.2 建模结果

文献[15]将直径作为变量和透射光谱一起建模,作为修正果实大小对苹果霉心病透射光谱的方法,暂时把这种方法称为“添加直径”法。本文新方法暂时称为“衰减系数”法。为了进一步验证文中光谱修正方法的有效性,将参与建模的204个样本光谱用两种方法分别进行修正,经过预处理后建立SVM模型和BP-ANN模型与原始光谱建立相同的模型进行对比(表3)。由表3可知,在SVM模型中,“衰减系数”法修正光谱建立的模型对训练集和测试集的判别准确率分别为99.34%和90.20%,高于“添加直径”法修正光谱的91.50%和84.31%,同样高于原始光谱的91.50%和84.31%。在BP模型中,“衰减系数”法修正光谱建立的模型对训练集和测试集的判别准确率分别为92.16%和88.24%,高于“添加直径”法修正光谱的90.20%和84.31%,同样高于原始光谱的88.89%和84.31%。

表3 光谱修正前后模型对比

Tab.3 Model comparison before and after spectral correction

光谱类型与建模方法	训练集			测试集		
	误判数		判别准确率/%	误判数		判别准确率/%
	霉心病果	健康果		霉心病果	健康果	
原始光谱-归一化-PCA-SVM	8	5	91.50	8	0	84.31
“添加直径”法修正光谱-归一化-PCA-SVM	8	5	91.50	8	0	84.31
“衰减系数”法修正光谱-归一化-PCA-SVM	1	0	99.34	4	1	90.20
原始光谱-归一化-PCA-BP	11	6	88.89	5	3	84.31
“添加直径”法修正光谱-归一化-PCA-BP	10	5	90.20	5	3	84.31
“衰减系数”法修正光谱-归一化-PCA-BP	7	5	92.16	5	1	88.24

2.4.3 建模结果讨论

3种不同处理下的光谱基于两种算法建立的模型判别准确率均在84%以上。对比两种算法建立的模型可以发现,基于SVM算法建立的判别模型明显优于基于BP算法建立的模型,这也证明了SVM算法在处理非线性高维数据的分类问题时,具有更好的泛化能力。在SVM算法中,文中提出的“衰减系数”法修正光谱建立模型效果最好,相对于原始光谱建立的模型对训练集和测试集的判别准确率分别提高了7.84和5.89个百分点。

文中用“添加直径”法修正光谱在改善模型识

别精度方面的效果并不明显,主要是因为原始光谱包含的维度过大,达到了1001维,在光谱后添加1维的直径信息被高维光谱信息所覆盖。在便携式设备上,光谱维度较低时,这种方法在改善模型精度方面取得的效果更为显著。

通过对比不同算法下模型对于健康果和霉心病果的误判数发现,霉心病果的准确识别更为困难。“衰减系数”法修正光谱建立SVM模型相对于原始光谱建立的模型将霉心病果的误判数由8个减少为4个。但是,仍然存在4个霉心病果发生误判。其中2个霉心病果(图8),分别为158号果和141号

果。这 2 个苹果的发病情况较轻,141 号果为早期褐变,158 号果为早期霉心,模型判别失误是由于该果发病较轻,病菌尚未影响到果肉的生理特性,导致其透射光谱与正常苹果光谱差别较小,这种轻微发病或者发病早期的霉心病果判别较为困难。

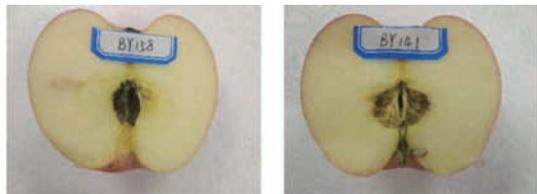


图 8 误判样本分析

Fig. 8 Analysis of misjudgement samples

本文在假设果肉均匀化的基础上,基于数学统计的方法求取衰减系数 P 进而修正透射光谱,尚未对光在苹果内的传输机理以及散射情况进行深入研究,后续研究将进一步对果肉不同部位分析并进一步修正透射光谱。这有助于解决轻微发病、早期发

病的病果识别问题。与国内外研究相比较,本文模型判别准确率较好。基于近红外透射光谱对不同大小的苹果进行分析建立霉心病判别模型,测试集判别准确率达到 90.20%,与文献[3,11]结果相差不多。但是与文献[13]还存在着一定的差距。这是由于其选取的是大小均匀的苹果,而且选择样品的生长年份不同、地区不同、发病程度不同,同样会导致模型的判别准确率存在差距。

3 结束语

提出了一种修正果实直径对透射光谱影响的新方法。为了将不同大小的果实光谱转换到同一果径基准下,用相关公式求取光在果实内部的衰减系数 P ,利用光在苹果内穿透时的衰减系数 P 修正透射光谱,进而利用修正后的光谱建立霉心病判别模型。该模型能够很好地识别出健康苹果与霉心病果,与未修正的模型进行对比,对训练集和测试集的判别准确率分别提高了 7.84 和 5.89 个百分点。

参 考 文 献

- [1] 刘会香. 苹果霉心病的研究现状及展望[J]. 水土保持研究, 2001, 8(3): 91-92.
LIU Huixiang. The studying advance and prospect of apple moldy core[J]. Research of Soil and Water Conservation, 2001, 8(3): 91-92. (in Chinese)
- [2] 李晓荣, 陈小飞, 李晓萍. 苹果霉心病发生原因及防治措施[J]. 果农之友, 2009(11): 26.
- [3] 苏东, 张海辉, 陈克涛, 等. 基于透射光谱的苹果霉心病多因子无损检测[J]. 食品科学, 2016, 37(8): 207-211.
SU Dong, ZHANG Haihui, CHEN Ketao, et al. Multiple-factor nondestructive detection of moldy core in apples based on transmission spectra[J]. Food Science, 2016, 37(8): 207-211. (in Chinese)
- [4] LAMMERTYN J, DRESSELAERS T, HECKEB P V, et al. MRI and X-ray CT study of spatial distribution of core breakdown in 'Conference' pears[J]. Magnetic Resonance Imaging, 2003, 21(7): 805-815.
- [5] LU Y, LU R. Non-destructive defect detection of apples by spectroscopic and imaging technologies: a review[J]. Transactions of the ASABE, 2017, 60(5): 1765-1790.
- [6] POREP J U, KAMMERER D R, CARLE R. On-line application of near infrared (NIR) spectroscopy in food production[J]. Trends in Food Science & Technology, 2015, 46(2): 211-230.
- [7] 李芳, 蔡驷, 马惠玲, 等. 基于生物阻抗特性分析的苹果霉心病无损检测[J]. 食品科学, 2013, 34(18): 197-202.
LI Fang, CAI Cheng, MA Huiling, et al. Nondestructive detection of apple mouldy core based on bioimpedance properties[J]. Food Science, 2013, 34(18): 197-202. (in Chinese)
- [8] 杨亮亮. 基于机器视觉和 X 射线的苹果霉心病检测方法研究[D]. 杨凌: 西北农林科技大学, 2009.
YANG Liangliang. Research of the detection of mould core apple based on machine vision and X-ray[D]. Yangling: Northwest A&F University, 2009. (in Chinese)
- [9] MCGLONE V A, MARTINSEN P J, CLARK C J, et al. On-line detection of brownheart in Braeburn apples using near infrared transmission measurements[J]. Postharvest Biology and Technology, 2005, 37(2): 142-151.
- [10] CLARK C J, MCGLONE V A, JORDAN R B. Detection of Brownheart in 'Braeburn' apple by transmission NIR spectroscopy[J]. Postharvest Biology and Technology, 2003, 28(1): 87-96.
- [11] SHENDEREY C, SHMULEVICH I, ALCHANATIS V, et al. NIRS detection of moldy core in apples[J]. Food and Bioprocess Technology, 2010, 3(1): 79-86.
- [12] 韩东海, 刘新鑫, 鲁超, 等. 苹果内部褐变的光学无损检测研究[J]. 农业机械学报, 2006, 37(6): 86-93.
HAN Donghai, LIU Xinxin, LU Chao, et al. Study on optical-nondestructive detection of breakdown apples[J]. Transactions of the Chinese Society for Agricultural Machinery, 2006, 37(6): 86-93. (in Chinese)
- [13] 雷雨, 何东键, 周兆永, 等. 苹果霉心病可见/近红外透射能量光谱识别方法[J/OL]. 农业机械学报, 2016, 47(4): 193-200. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20160426&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2016.04.026.
LEI Yu, HE Dongjian, ZHOU Zhaoyong, et al. Detection of moldy core of apples based on visible/near infrared transmission energy spectroscopy[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(4): 193-200. (in Chinese)

- [14] 李顺峰,张丽华,刘兴华,等.基于主成分分析的苹果霉心病近红外漫反射光谱判别[J].农业机械学报,2011,42(10):158-161.
LI Shunfeng, ZHANG Lihua, LIU Xinghua, et al. Discriminant analysis of apple moldy core using near infrared diffuse reflectance spectroscopy based on principal component analysis [J]. Transactions of the Chinese Society for Agricultural Machinery, 2011, 42(10):158-161. (in Chinese)
- [15] 张海辉,陈克涛,苏东,等.基于特征光谱的苹果霉心病无损检测设备设计[J].农业工程学报,2016,32(18):255-262.
ZHANG Haihui, CHEN Ketao, SU Dong, et al. Design of nondestructive detection device for moldy core in apples based on characteristic spectrum [J]. Transactions of the CSAE, 2016, 32(18):255-262. (in Chinese)
- [16] 河野澄夫.近赤外分光法による果実糖度の測定[J].食糧_その科学と技術,2005,43:69-86.
- [17] QIN J, LU R. Monte Carlo simulation for quantification of light transport features in apples [J]. Computers and Electronics in Agriculture, 2009, 68(1):44-51.
- [18] 刘燕德.水果糖度和酸度的近红外光谱无损检测研究[D].杭州:浙江大学,2006.
LIU Yande. Study on methods of nondestructive measurement of sugar content and acidity in fruits using near-infrared spectroscopy [D]. Hangzhou: Zhejiang University, 2006. (in Chinese)
- [19] CEN H, LU R, MENDOZA F, et al. Relationship of the optical absorption and scattering properties with mechanical and structural properties of apple tissue [J]. Postharvest Biology and Technology, 2013, 85:30-38.
- [20] BEERS R V, AERNOOTS B, WATTE R, et al. Effect of maturation on the bulk optical properties of apple skin and cortex in the 500~1850 nm wavelength range [J]. Journal of Food Engineering, 2017, 214:79-89.
- [21] 闵顺耕,李宇,张明祥.近红外光谱分析中异常值的判别与定量模型优化[J].光谱学与光谱分析,2004,24(10):1205-1209.
MIN Shungeng, LI Ning, ZHANG Mingxiang. Outlier diagnosis and calibration model optimization for near infrared spectroscopy analysis [J]. Spectroscopy and Spectral Analysis, 2004, 24(10):1205-1209. (in Chinese)
- [22] 陈奕云,赵瑞瑛,齐天赐,等.结合光谱变换和 Kennard-Stone 算法的水稻土全氮光谱估算模型校正集构建策略研究[J].光谱学与光谱分析,2017,37(7):2133-2139.
CHEN Yiyun, ZHAO Ruiying, QI Tianci, et al. Constructing representative calibration dataset based on spectral transformation and Kennard-Stone algorithm for VNIR modeling of soil nitrogen in paddy soil [J]. Spectroscopy and Spectral Analysis, 2017, 37(7):2133-2139. (in Chinese)
- [23] 展晓日,朱向荣,史新元,等.SPXY 样本划分法及蒙特卡罗交叉验证结合近红外光谱用于橘叶中橙皮苷的含量测定[J].光谱学与光谱分析,2009,29(4):964-968.
ZHAN Xiaori, ZHU Xiangrong, SHI Xinyuan, et al. Determination of hesperidin in tangerine leaf by near-infrared Spectroscopy with SPXY algorithm for sample subset partitioning and monte carlo cross validation [J]. Spectroscopy and Spectral Analysis, 2009, 29(4):964-968. (in Chinese)
- [24] KAYA-CELIK H, MALLIKARJUNAN P K, KAYA A. Mid-infrared spectroscopy for discrimination and classification of *Aspergillus spp.* contamination in peanuts [J]. Food Control, 2015, 52:103-111.
- [25] 李江波,赵春江,陈立平,等.基于可见/近红外光谱谱区有效波长的梨品种鉴别[J/OL].农业机械学报,2013,44(3):153-157,179. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20130328&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2013.03.028.
LI Jiangbo, ZHAO Chunjiang, CHEN Liping, et al. Variety identification of pears based on effective wavelengths in visible/near infrared region [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(3):153-157,179. (in Chinese)
- [26] 郭文川,刘大洋.猕猴桃膨大果的近红外漫反射光谱无损识别[J/OL].农业机械学报,2014,45(9):230-235. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20140937&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2014.09.037.
GUO Wenchuan, LIU Dayang. Identification of expanded kiwifruit by near-infrared diffused spectroscopy [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(9):230-235. (in Chinese)
- [27] 岳学军,全东平,洪添胜,等.柑橘叶片叶绿素含量高光谱无损检测模型[J].农业工程学报,2015,31(1):294-302.
YUE Xuejun, QUAN Dongping, HONG Tiansheng, et al. Non-destructive hyperspectral measurement model of chlorophyll content for citrus leaves [J]. Transactions of the CSAE, 2015, 31(1):294-302. (in Chinese)
- [28] 栾郭宏,贺凯迅,程辉,等.基于神经网络的近红外光谱辛烷值模型的研究及应用[J].计算机与应用化学,2014,31(1):63-68.
LUAN Guohong, HE Kaixun, CHENG Hui, et al. Octane model based on neural network by near-infrared spectroscopy and its application [J]. Computers and Applied Chemistry, 2014, 31(1):63-68. (in Chinese)
- [29] CORTES C, VAPNIK V. Support-vector networks [J]. Machine Learning, 1995, 20(3):273-297.