

# 苹果霉心病光谱在线检测的摆放姿态及建模方法优化研究

秦 楷<sup>1</sup>, 陈 刚<sup>2</sup>, 张剑一<sup>1,2</sup>, 傅霞萍<sup>1\*</sup>

1. 浙江理工大学机械与自动控制学院, 浙江 杭州 310018

2. 德菲洛(杭州)科技有限公司, 浙江 杭州 310014

**摘 要** 苹果营养丰富、口味酸甜,是深受大众喜爱的一种水果。苹果霉心病是一种真菌侵染果实病害,隐蔽性极强,一般在近成熟期果实内部发生霉变,肉眼从外观观察难以分辨,市面上大多数品种的苹果都受其影响。霉心病病果重量变轻、口感变差,严重的甚至不能食用,对经济效益的影响巨大。采用可见近红外光谱分析技术,使用微型光谱仪在线无损检测苹果霉心病,针对4种苹果在线输送时摆放姿态(竖放柄朝上、竖放柄朝下、横放柄朝输送方向和横放柄垂直输送方向)的判别效果进行了优化分析。首先使用主成分分析对600~900 nm波段的透射光谱提取主成分后分别建立线性判别分析(LDA)、马氏距离(MD)和K近邻法(KNN)模型并对校正集和预测集的判别准确率进行对比;其次对600~900 nm波段中心化预处理后建立偏最小二乘判别分析(PLS-DA)模型并给出4种摆放姿态的判别效果;最后使用两种机器学习算法极限学习机(ELM)和支持向量机(SVM)分别建立霉心病判别模型进行预测。对比上述所有6种判别模型,通过观察4种摆放姿态整体的判别效果得到最佳的建模方法为PLS-DA,其中竖放柄朝上和竖放柄朝下摆放的判别准确率都为93.75%,其他2种摆放姿态的判别准确率也都超过85%,再根据PLS-DA模型波段变量投影重要性指标得分值分布提取特征波段690~720 nm重新建立模型,对比4种摆放姿态效果最好的是竖放柄朝上摆放,其预测集的判别准确率达到93.75%,并且对病果的判别效果最佳。研究结果表明PLS-DA可以作为判别苹果霉心病一种有效方法,竖放柄朝上摆放可以作为苹果霉心病在线检测时一种有效姿态。

**关键词** 苹果;霉心病;光谱;摆放姿态;建模方法

**中图分类号:** S123 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2021)11-3405-06

## 引 言

苹果富含矿物质、维生素、多酚及黄酮类等营养物质,营养价值高,口味酸甜,是很受消费者喜爱的一种水果。苹果种植过程中会出现很多病症,霉心病(又称心腐病)就是其中一种,该病由多种病原菌引起,病菌主要通过花萼传播危害果实健康,在发病初期会在苹果心室附近出现褐色斑点,恶化后会变成白色或粉色霉状物,较为严重的会导致心室腐烂并逐步向外扩散<sup>[1]</sup>。感染霉心病的苹果一般是不能食用的,因此在销售苹果前需要剔除病果。传统的霉心病判别方法往往结合经验进行破坏性抽样,费时费力且准确率低不适合当前水果生产销售的实际。寻找一种水果内部病变的无损检测方法以及建立水果内部品质快速、准确的智能分选线符

合目前市场需求,近年来行业内越来越多的企业已经陆续开始做多品类、大规模、针对内在品质的水果分级。

可见近红外光谱分析技术是一项高效、低成本、无污染的无损检测技术,已经广泛应用于农业、食品业、医药业及矿业等各行各业。近年来,国内外不少学者也利用可见近红外光谱分析技术开展水果病害的快速无损检测研究,其中涉及内部病变检测的如:Khawiwada等<sup>[2]</sup>利用该技术对在大气中储藏的苹果内部褐变病症进行了预测,采用LDA,PLS-DA,SVM和逻辑回归4种判别方法,结果显示PLS-DA效果最佳;Shenderay等<sup>[3]</sup>在现有的微型光谱分析仪的基础上,研制了一套适合于在线近红外光谱测量的仪器,采用偏最小二乘回归(PLSR)方法在线检测苹果霉心病;Sun等<sup>[4]</sup>利用可见近红外光谱在600~904 nm范围内,通过相关分析选取两个特征波段用于判别褐心梨;Mogollonca等<sup>[5]</sup>采用偏最小

收稿日期:2020-10-20,修订日期:2021-02-22

基金项目:国家自然科学基金项目(32071904),浙江省自然科学基金项目(LY20C130008),浙江理工大学科研启动基金项目(ZSTU 16022177-Y)资助

作者简介:秦 楷,1996年生,浙江理工大学机械与自动控制学院硕士研究生 e-mail:2018g0507014@mails.zstu.edu.cn

\* 通讯作者 e-mail:fuxp@zstu.edu.cn

二乘和支持向量机方法建立了定性和定量模型, 预测“Cripps Pink”苹果内部是否褐变及褐变面积; Zhou 等<sup>[6]</sup>提出了一种基于透光光谱的苹果霉变快速无损检测方法, 采用人工神经网络(ANN)和 SVM 两种分类算法建立模型, 利用遗传算法(GA)对模型参数进行优化, 判定苹果核是否霉变; 采用透射光谱区分霉心病健康果和病果时果径会对预测结果产生影响, 为解决这一问题, Tian 等<sup>[7]</sup>提出了一种基于果实大小的光谱校正方法, 通过计算透射光在苹果内部的消光系数对透射谱进行了修正, 从而得到更好的判别效果。Hu 等<sup>[8]</sup>通过连续投影算法(SPA)选择出区分健康苹果和霉心病苹果的最佳波长, 将提取的波长作为反向传播人工神经网络(BP-ANN)的输入建立判别模型。

在上述研究基础上, 本研究针对苹果在线检测中不同摆放姿态对光谱采集的影响, 开展了果实姿态对霉心病检测的影响分析, 通过采用不同算法建立多种判别模型进行对比分析, 实现苹果霉心病光谱在线检测的摆放姿态及建模方法优化。

## 1 实验部分

### 1.1 样品

样品是产自中国陕西省洛川县的红富士苹果, 在洛川当地选购后快速运往实验室。剔除有明显外部损伤的果实后, 最终选出 96 个苹果样品用于本次检测。测量所有样品外观参数后用标签标注, 放置在实验室内等待第二天的检测。

### 1.2 光谱采集

采用带暗箱的传送装置, 将苹果样本置于托盘上方, 由传送带输送至检测位置, 卤素光源(12 V, 100 W)布置在传送带两侧, 透射光由下方的检测探头收集, 与之连接的光纤将信号传输至微型可见近红外光谱仪(QE65Pro, Ocean Insight, FL, USA), 整个检测单元置于暗箱中以避免环境光影响, 暗箱两侧留有样品进出口。光谱仪通过 USB 数据线连接电脑, 使用 SpectraSuite 软件(Ocean Insight, FL, USA)查看和存储光谱。为考察苹果样品摆放姿态对检测结果的影响, 每个样品采集 4 次光谱, 分别对应 4 种摆放姿态: 竖放柄朝上、竖放柄朝下、横放柄朝输送方向和横放柄垂直输送方向。96 个样品对应 4 种摆放姿态共获得 384 条光谱。

### 1.3 数据分析

光谱采集完成后, 将苹果沿果柄蒂轴对半切开观察心室是否霉变(如图 1 所示), 将健康果归于 1 类、霉心病果归于 2 类。基于获取的光谱及有损检测获得的霉心病果判据, 开展苹果霉心病的判别分析, 取 2/3 样本作为校正集、1/3 样本作为预测集。采用线性判别分析<sup>[9]</sup>、马氏距离判别<sup>[10]</sup>、偏最小二乘判别<sup>[11-12]</sup>、K 近邻法<sup>[13]</sup>、极限学习机<sup>[14]</sup>和支持向量机<sup>[15]</sup>对 4 种摆放姿态分别建立判别模型并对比判别效果(以判别准确率为指标)。数据分析主要采用 MATLAB (R2016a, The MathWorks, Natick, MA, USA)软件开展。



图 1 健康果(左)和霉心病果(右)心室对比示意图

Fig. 1 Contrast of healthy fruit (left) and fruit with moldy core (right)

## 2 结果与讨论

### 2.1 果实外观参数及透射光谱

经统计, 96 个样品中包含 78 个健康果和 18 个霉心病果。样品外部参数指标如表 1 所示。

表 1 健康果和霉心病果的外观参数统计

Table 1 Appearance parameter statistics of healthy fruit and fruit with moldy core

样品分类	数量 / 个	平均最大果径 / mm	平均高度 / mm	平均重量 / g
健康果	78	87.0	71.5	240.3
病果	18	86.4	70.7	224.6
全部样品	96	86.9	71.4	237.4

基于所采集的全部样品 4 种摆放姿态的透射光谱, 分别计算健康果和霉心病果 4 种摆放姿态的平均透射光谱, 结果如图 2 所示, 可以发现在 400~600 nm 间存在较高的噪声, 而 900 nm 之后样品的透射率几乎为零, 因此选取 600~900 nm 的透射光谱进行分析。4 种摆放姿态的健康果的平均透射光谱在 600~900 nm 波段内整体都高于病果的平均透射光谱, 尤其是在 710 nm 附近健康果的透射光谱峰值明显高于病果, 可能是因为病果果实内部霉变组织的吸收较大所致。

### 2.2 霉心病果判别结果

基于 600~900 nm 的透射光谱信息, 结合主成分分析提取的主成分(PCs)建立的 LDA, MD 和 KNN 三种判别模型, 其中 KNN 算法中  $k=7$ , 表示对于一个需要预测的输入向量  $x$ , 我们只需要在训练数据集中寻找 7 个与  $x$  最近的向量的集合, 然后把  $x$  的类别预测为这 7 个样本中类别数最多的那一类。表 2 显示三种判别模型提取的主成分数以及训练集和预测集的判别效果, 4 种摆放姿态的 LDA 判别模型预测集的判别准确率最高的是竖放柄朝上摆放为 90.63%, 有 3 个健康果被误判为病果, 整体来看 LDA 模型竖放柄朝上摆放和竖放柄朝下摆放的判别效果明显优于横放柄朝输送方向和横放柄垂直输送方向摆放时的效果; MD 模型 4 种摆放姿态的判别效果差别不大, 预测集判别准确率都达到 87.50%, 虽然 MD 模型对健康果没有产生误判但其对病果的识别效果较差; KNN 模型中预测集的判别准确率最高的是竖放柄朝下的 90.63%, 相对判别效果较差的横放柄朝输送方向摆

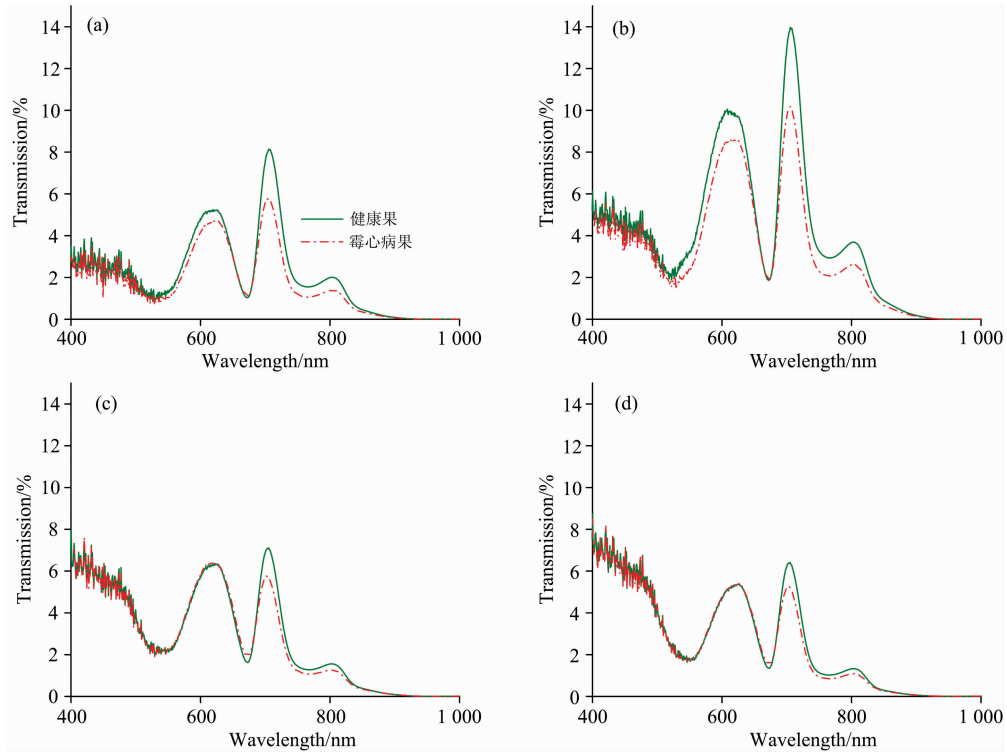


图 2 健康果和病果 4 种摆放姿态下的平均透射光谱

(a): 竖放柄朝上; (b): 竖放柄朝下; (c): 横放柄朝输送方向; (d): 横放柄垂直输送方向

Fig. 2 Average transmission spectra of healthy fruit and fruit with moldy core in four transportation postures

(a): Stem upward; (b): Stem downward; (c): Stem towards the transportation direction;

(d): Stem perpendicular to the transportation direction

表 2 基于 4 种摆放姿态透射光谱 (600~900 nm) 提取主成分后的线性判别分析 (LDA)、马氏距离 (MD) 和 K 近邻法 (KNN) 三种判别模型结果的比较

Table 2 Comparison of three discriminant modeling results of LDA, MD and KNN based on PCs extracted from transmission spectra (600~900 nm) of apples with four types of posture

方法	摆放姿态	PCs	训练集			预测集		
			错判个数		准确率 /%	错判个数		准确率 /%
			病果	健康果		病果	健康果	
LDA	竖放柄朝上	37	1	0	98.44	0	3	90.63
	竖放柄朝下	37	0	0	100	2	4	81.25
	横放柄朝输送方向	38	0	2	96.88	1	12	59.38
	横放柄垂直输送方向	37	1	4	92.19	2	10	62.50
MD	竖放柄朝上	13	4	0	93.75	4	0	87.50
	竖放柄朝下	14	2	0	96.88	4	0	87.50
	横放柄朝输送方向	10	2	0	96.88	4	0	87.50
	横放柄垂直输送方向	14	2	0	96.88	4	0	87.50
KNN	竖放柄朝上	12	9	0	85.94	2	2	87.50
	竖放柄朝下	12	12	0	81.25	3	0	90.63
	横放柄朝输送方向	9	11	1	81.25	3	3	81.25
	横放柄垂直输送方向	9	12	0	81.25	2	2	87.50

放和横放柄垂直输送方向摆放的预测集判别准确率也可达到 80% 以上。

通过对 LDA, MD 和 KNN 三种判别模型的比较, 从表 2 中各项数据可以看出 LDA 模型和 KNN 模型竖放柄朝上摆

放和竖放柄朝下摆放的判别效果相比其他两种摆放姿态都更好, 在三种判别模型中效果最佳的是 LDA 模型的竖放柄朝上摆放。

对 600~900 nm 的透射光谱进行中心化预处理后建立的

PLS-DA 模型结果如表 3 所示, 当苹果竖放柄朝上摆放和竖放柄朝下摆放检测时预测集的判别准确率最高为 93.75%,

苹果横放柄垂直输送方向进行检测所建立的判别模型效果相对其他三种摆放姿态较差。

表 3 基于 4 种摆放姿态透射光谱 (600~900 nm) 的偏最小二乘判别分析 (PLS-DA) 结果  
Table 3 PLS-DA results using transmission spectra (600~900 nm) of apples with 4 types of posture

摆放姿态	LVs	训练集			预测集		
		错判个数		准确率/%	错判个数		准确率/%
		病果	健康果		病果	健康果	
竖放柄朝上	7	6	3	85.94	0	2	93.75
竖放柄朝下	7	8	1	85.94	1	1	93.75
横放柄朝输送方向	7	2	1	90.63	4	2	90.63
横放柄垂直输送方向	2	12	0	81.25	4	0	87.50

表 4 基于 4 种摆放姿态透射光谱 (600~900 nm) 的极限学习机 (ELM) 和支持向量机 (SVM) 两种判别模型的比较  
Table 4 Comparison of two modeling results of ELM and SVM using transmission spectra (600~900 nm) of apples with 4 types of posture

方法	摆放姿态	训练集			预测集		
		错判个数		准确率/%	错判个数		准确率/%
		病果	健康果		病果	健康果	
ELM	竖放柄朝上	6	0	90.63	1	2	90.63
	竖放柄朝下	1	7	87.50	2	2	87.50
	横放柄朝输送方向	1	4	92.19	2	1	90.63
	横放柄垂直输送方向	2	5	89.06	1	2	90.63
SVM	竖放柄朝上	0	7	89.06	1	0	96.88
	竖放柄朝下	0	12	81.25	0	4	87.50
	横放柄朝输送方向	0	0	100.00	4	1	84.38
	横放柄垂直输送方向	1	4	92.19	4	1	84.38

600~900 nm 的透射光谱经归一化处理后, 使用 ELM 和 SVM 两种机器学习算法建立的判别模型结果如表 4 所示, 对于 ELM 模型选用 sigmoid 函数为极限学习机的激活函数并设置隐含层神经元个数为 20, 当苹果竖放柄朝上、横放柄朝输送方向和横放柄垂直输送方向的预测集判别准确率都为 90.63%, 有 2 个病果和 1 个健康果被错判, 其余三种摆放姿态的预测集的判别准确率也都高于 85%; 建立 SVM 模型时利用交叉验证的方法获得最优的惩罚参数  $c$  和核函数的参数  $g$ , 通过表 4 可知, 果柄朝上摆放时预测集判别准确率最高为 96.88%, 预测集只有 1 个病果被错判。

通过对上述 6 种判别模型预测集判别效果的汇总结果如图 3 所示, LDA 和 SVM 模型的 4 种摆放姿态的预测集判别率最佳的是竖放柄朝上摆放, MD 和 ELM 模型的 4 种摆放姿态的判别效果接近, KNN 模型除了横放柄朝输送方向摆放时判别效果较差其余 3 种摆放姿态的预测集判别率都超过 87%, PLS-DA 模型的最佳摆放姿态为竖放柄朝上摆放和竖放柄朝下摆放, 整体来看竖放柄朝上摆放在不同的判别模型中都有不错的判别效果, 这可能与在线检测装置中光源和检测探头的位置布置有关。

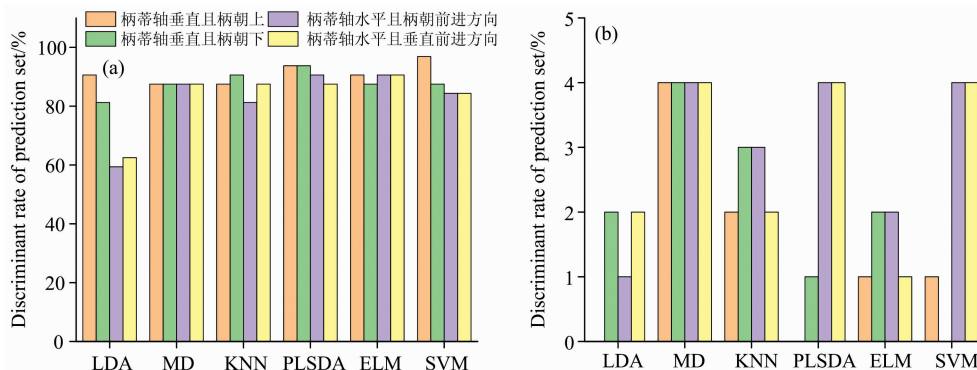


图 3 预测集的 6 种判别模型效果 (a) 及病果错判为健康果个数 (b) 对比图

Fig. 3 Comparison of the performance (a) and the number of moldy core apples incorrectly identified as healthy (b) of the six discriminant models

从建模方法来看 PLS-DA 的判别效果优于其他方法, 其中 PLS-DA 模型的竖放柄朝上和竖放柄朝下的判别准确率都为 93.75%, 通过 6 种判别模型预测集病果错判为健康果个数结果对比, PLS-DA 判别模型的竖放柄朝上的摆放姿态没有误判果, 可以很好的识别出预测集中所有的病果。

苹果霉心病 PLS-DA 判别模型的变量投影重要性指标 (VIP) 得分值分布如图 4 所示, 在 710 nm 附近得分值达到峰值, 因此进一步选择 690~720 nm 作为特征波段建立 PLS-DA 判别模型, 由表 5 所示结果可知, 竖放柄朝上和竖放柄朝下时预测集的判别准确率为 93.75%。4 种摆放姿态在 690~720 nm 波段范围的预测集判别准确率均大于 80%, 这为采用 LED 灯代替卤素灯检测苹果霉心病的可行性提供了依据。

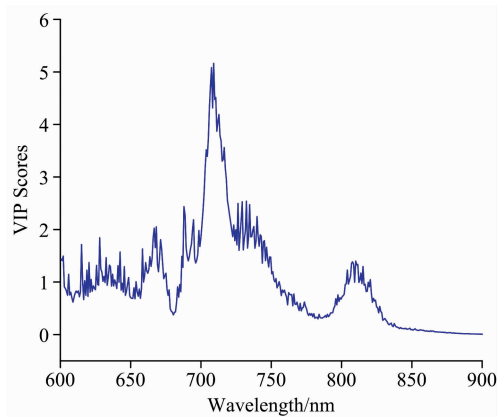


图 4 PLS-DA 判别模型不同波长的 VIP 得分值分布  
Fig. 4 VIP score distribution of PLS-DA model

表 5 基于 4 种摆放姿态透射光谱 (690~720 nm) 的偏最小二乘判别分析 (PLS-DA) 结果  
Table 5 PLS-DA results using transmission spectra (690~720 nm) of apples with 4 types of posture

摆放姿态	LVs	训练集			预测集		
		错判个数		准确率/%	错判个数		准确率/%
		病果	健康果		病果	健康果	
竖放柄朝上	4	4	0	93.75	0	2	93.75
竖放柄朝下	3	13	1	78.13	2	0	93.75
横放柄朝输送方向	4	1	2	95.31	1	4	84.38
横放柄垂直输送方向	3	12	1	79.69	4	0	87.50

### 3 结 论

基于 600~900 nm 波段 4 种摆放姿态的透射光谱提取主成分后建立的 LDA, MD 和 KNN 判别模型都有着不错的判别效果, 其中 LDA 模型对应竖放柄朝上摆放姿态的预测集判别准确率最佳为 90.63%, LDA 和 KNN 模型竖放柄朝上和竖放柄朝下的判别效果都优于横放柄朝输送方向和横放柄垂直输送方向。基于 600~900 nm 建立的 PLS-DA 判别模型 4 种摆放姿态的预测集判别准确率都大于 80%, 并且竖放柄朝上和竖放柄朝下的判别准确率都是大于其余两种摆放姿态。还建立了 ELM 和 SVM 两种机器学习算法判别模型, 从

预测集判别准确率和病果错判个数观察效果最佳都是竖放柄朝上的姿态。

综合以上建立的 6 种判别模型来看, 竖放柄朝上摆放在不同的判别模型中都有不错的判别效果, 这可能与在线检测装置中光源和检测探头的位置有关。从建模方法来看, 4 种摆放姿态整体判别效果最好的 PLS-DA 模型, 再通过其 VIP 得分值提取特征波段 690~720 nm 后重新建立 PLS-DA 判别模型, 竖放柄朝上和竖放柄朝下的预测集判别准确率都达到 93.75%, 其中竖放柄朝上摆放检测到的光谱建立的模型筛选病果的能力达到最佳, 全部病果都能被判别出, 因此竖放柄朝上摆放检测是一种合理有效的苹果霉心病在线检测姿态。

### References

- [1] Gao LL, Zhang Q, Sun X Y, et al. Plant Disease, 2013, 97(4): 510.
- [2] Khatiwada B P, Subedi P P, Hayes C, et al. Postharvest Biology and Technology, 2016, 120: 103.
- [3] Shenderoy C, Shmulevich I, Alchanatis V, et al. Food and Bioprocess Technology, 2010, 3(1): 79.
- [4] Sun X D, Liu Y D, Li Y F, et al. Postharvest Biology and Technology, 2016, 116: 80.
- [5] Mogollon M R, Jara A F, Contreras C, et al. Postharvest Biology and Technology, 2019, 116: 111060.
- [6] Zhou Z Y, Lei Y, Su D, et al. International Journal of Agricultural and Biological Engineering, 2016, 9(6): 148.
- [7] Tian S J, Zhang J H, Zhang Z X, et al. Infrared Physics and Technology, 2019, 100: 117.
- [8] Hu Q X, Tian J, Fang Y. International Journal of Pattern Recognition and Artificial Intelligence, 2019, 33(12): 1950020.
- [9] Cortes V, Cubero S, Blasco J, et al. Food and Bioprocess Technology, 2019, 12(6): 1021.
- [10] Dong X G, Dong J, Li Y L, et al. Computers and Electronics in Agriculture, 2018, 156: 669.
- [11] LIU Yan-de, XIAO Huai-chun, SUN Xu-dong, et al(刘燕德, 肖怀春, 孙旭东, 等). Spectroscopy and Spectral Analysis(光谱学与光谱

- 分析), 2018, 38(2): 52.
- [12] Ziegler J U, Leitenberger M, Longin C F H, et al. *Journal of Food Composition and Analysis*, 2016, 51: 30.
- [13] Teye E, Amuah C L Y, McGrath T, et al. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2019, 217: 147.
- [14] Yang M H, Xu D Y, Chen S C, et al. *Sensors*, 2019, 19(2): 263.
- [15] Parastar H, van Kollenburg G, Weesepeol Y, et al. *Food Control*, 2020, 112: 107149.

## Optimization of Fruit Pose and Modeling Method for Online Spectral Detection of Apple Moldy Core

QIN Kai<sup>1</sup>, CHEN Gang<sup>2</sup>, ZHANG Jian-yi<sup>1,2</sup>, FU Xia-ping<sup>1\*</sup>

1. Faculty of Mechanical Engineering and Automation, Zhejiang Sci-Tech University, Hangzhou 310018, China
2. Zhejiang DEKFELLER Intelligent Machinery Manufacturing Co., Ltd., Hangzhou 310014, China

**Abstract** Moldy core of apples is a fungal disease that affects many commercially popular cultivars of apples. It is difficult to distinguish moldy core of the fruit from its appearance until the fruit is cut open. The objective of this study was to detect moldy core of apples by visible near-infrared spectroscopy (NIRS). The discrimination effects of four kinds of apple on-line transportation postures were compared: the apple stem upward, the apple stem downward, the apple stem towards the transportation direction, and the apple stem perpendicular to the transportation direction. Principal component analysis (PCA) was used to extract the principal components from the transmission spectra of 600~900 nm, and then linear discriminant analysis (LDA), Mahalanobis distance (MD) and  $k$ -nearest neighbor (KNN) models were established for comparison. The partial least squares discriminant analysis (PLS-DA) model was established after the central pretreatment of 600~900 nm. Two machine learning algorithms, extreme learning machine (ELM) and support vector machine (SVM) were also used to predict moldy core of apples. The best modeling method is PLS-DA. The accuracy rate of stem upward and stem downward was 93.75%, and the accuracy of the other two postures were more than 85%. Then according to VIP (variable importance in projection) scores, the characteristic band 690~720 nm was extracted, and the model was rebuilt. The best result of the four postures was apple stem upward. The accuracy rate of the prediction set was 93.75%. The results showed that PLS-DA could be used as an effective method to distinguish moldy core of apples, and the stem upward can be used as an effective posture for on-line detection of moldy core of apples.

**Keywords** Apple; Moldy core; Spectrum; Posture; Modeling method

(Received Oct. 20, 2020; accepted Feb. 22, 2021)

\* Corresponding author