

不同速度对近红外光谱预测库尔勒香梨品质模型的影响

陈东杰^{1,2}, 姜沛宏^{1,2}, 郭风军^{1,2}, 张玉华^{1,2*}, 张长峰^{1,2}

1. 国家农产品现代物流工程技术研究中心, 山东 济南 250103
2. 山东省农产品贮运保鲜技术重点实验室, 山东 济南 250103

摘要 针对目前库尔勒香梨品质在线分级检测系统存在价格昂贵、结构复杂等问题, 设计了库勒尔香梨内部品质在线无损检测分级系统。基于该系统研究了不同移动速度(0.3 和 0.5 m·s⁻¹)对库勒尔香梨的可溶性固形物含量(solid soluble contents, SSC)和硬度在线预测模型的影响。不同移动速度下, 采集样品相同部位的信息, 所采集光谱存在差异。由于采集的光谱存在差异性, 采用 SG-平滑(Savitzky-Golay smooth)、SG 卷积导数、多元散射校正(MSC)、标准正态能量变换(SNV)、归一化(Normalization)等多种光谱预处理方法进行处理, 基于偏最小二乘法(partial least squares, PLS), 建立移动速度为 0.3 m·s⁻¹(S1)和 0.5 m·s⁻¹(S2)下库勒尔香梨的 SSC 和硬度模型。结果表明: 移动速度为 0.5 m·s⁻¹下, 采用 SG-DER(Savitzky-Golay Derivative)处理光谱图建立 SSC 模型优于 0.3 m·s⁻¹, 其预测集相关系数和预测均方根误差为 0.880 2和 0.391 5°Bri。而在移动速度为 0.3 m·s⁻¹下的结果, 采用 SGS(Savitzky-Golay smooth)处理光谱图建立的 SSC 模型优于 0.5 m·s⁻¹下的结果, 其预测集相关系数和预测均方根误差分别为 0.820 2和 0.470 8 N。后建立两个速度混合模型, 采用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)筛选特征变量, 后采用 PLS, 建立混合速度下硬度和 SSC 预测模型。从建模效果来看 SPA 和 CARS 都可以有效减少建模所用变量数、提高库勒尔香梨在线 SSC 和硬度检测模型的预测能力和运算速度, 增强模型的稳健性等。采用 CARS 方法, 从 501 个光谱中筛选出 24 个变量, 建立了 CARS-PLS 模型, 建立的 SSC 模型较好, 其预测集相关系数和预测均方根误差分别为 0.915 0 和 0.371 9°Bri。采用 SPA 方法, 从 501 个光谱中筛选出 32 个变量, 建立硬度模型较好, 其预测集相关系数和预测均方根误差分别为 0.821 0 和 0.492 0 N。混合速度建立预测品质模型比单一速度建立模型稳健一些。研究表明: 不同移动速度对建立果品品质预测模型产生不同影响, 该研究有助于果品品质在线分选提供技术支持。

关键词 近红外光谱; 库尔勒香梨; 不同移动速度; 在线检测

中图分类号: TS207.3 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2020)06-1839-07

引言

库尔勒香梨(*Pyrus sinkiangensis* Yü)为新疆特色优质果品, 因肉质清脆, 香味独特, 口感合适深受国内外消费者的喜爱。库尔勒香梨外部形状的多样化, 增加了库尔勒香梨商业化分选的难度。传统的人工分级方法远远不能满足果农的要求。其近红外光谱已应用在检测苹果、草莓、柑桔、梨、西瓜^[1-4]等的品质, 是一种很有效经济的分选方法。国内外学

者对近红外光谱在线检测做了很多研究: 李龙等^[5]利用近红外光谱和机器视觉搭建了苹果在线无损检测分级系统, 苹果内部可溶性固形物含量所建立模型的校正集相关系数为 0.950 8, 校正集均方根误差为 0.347%, 预测集相关系数为 0.949, 预测集均方根误差为 0.449%。单个苹果的检测时间为 0.71 s。McGlone^[6]等采用 2 种近红外光谱系统, 采用漫透射, 分别获取运动速度 0.5 m·s⁻¹苹果光谱, 建立干物质含量的预测模型。Sun^[7]等采用近红外光谱在线分析技术, 研究速度对翠冠梨可溶性固形物的影响, 发现运行速度为

收稿日期: 2019-05-23, 修订日期: 2019-09-02

基金项目: 国家重点研发计划项目(2016YFD0400905), 国家重点研发计划项目(2018YFD040130), 山东省农机装备研发创新计划项目(2019YF011), 山东省高等学校科技计划项目(J18KA162)资助

作者简介: 陈东杰, 1988 年生, 国家农产品现代物流工程技术研究中心工程师 e-mail: dongjie613@163.com

* 通讯联系人 e-mail: zllf@163.com

0.3 m · s⁻¹ 的预测效果最好。

针对目前存在这些问题,与北京伟创英图科技有限公司联合研制可见光-近红外光谱可移动式果品在线快速分选系统,该系统自动化程度高,可移动且体积小,对库尔勒香梨内外品质快速在线检测分选集成化高,可满足果农日常分选的需要。基于该系统研究不同运行速度对库尔勒香梨 SSC 和硬度预测模型影响,并建立一种速度补偿模型,以降低移动速度对库尔勒香梨 SSC 和硬度在线检测的影响,建立混合速度的 SSC 和硬度预测模型。

1 实验部分

1.1 样品

库尔勒香梨由济南堤口集团有限责任公司提供,共选取 122 个库尔勒香梨,从冷库取出后,置于室温下 24 h,待样品与室温一致,沿着果品赤道部位进行标记,手动将库尔勒香梨放置于果托上(图 3),对标记区域进行光谱采集,每个样品采集 3 次,其平均光谱用于后续的数据处理分析。

1.2 在线采集系统与光谱采集

如图 1 所示,在线动态检测系统主要包括光谱采集单元、机器视觉系统、数据采集单元、品级分选单元及系统主控单元。本试验只采集库尔勒香梨光谱,并不进行图像采集。近红外光谱系统采用海洋光学(Ocean Optics Inc., USA)USB2000+光纤光谱仪,波长范围 600~1 000 nm,配有 SMA905 标准光纤接口,通过 USB 接口与计算机相连,光源为 12 V/50 W 卤素灯,共有 4 个,分别位于近红外检查箱内部的四个角;光谱仪光纤探头固定于传动链条的下方,采取水平照射底部测量模式,光源水平照射果品;果品置于果托之上,果托中心呈 2 cm 空心圆洞(如图 3),近红外测量探头置于下方,有效避免果品非测量部位造成的漫反射光干扰(如图 2)。

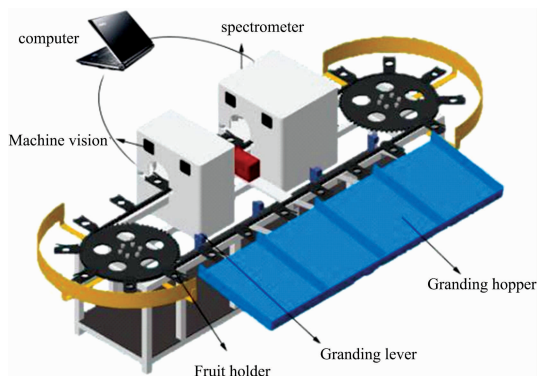


图 1 库尔勒香梨在线检测系统示意图

Fig. 1 Pear quality inspection and classification online system diagram

1.3 理化指标测定

1.3.1 硬度测定

采用 TA5 探头在质构仪 TMS-PRO(美国,FTC)上进 TPA 测试。沿着标记区域进行测试,设置测前速率 1 mm ·

s⁻¹; 测中及测后速率 1 mm · s⁻¹; 压缩程度 25%, 起始力 1 N。

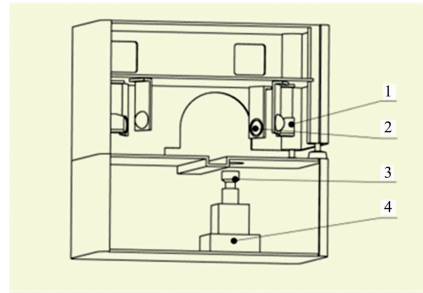


图 2 近红外光谱检测模块结构图

1: 光源; 2: 传感器; 3: 光纤传感器; 4: 光谱仪

Fig. 2 NIR inspection module structure diagram

1: Light; 2: Sensor; 3: Optical fiber Sensor; 4: Spectrometer



图 3 光源套件实物及效果图

Fig. 3 Diagram and rendering of light source modules

1.3.2 可溶性固形物(soluble solids content, SSC)测定

按照 NY/T585—2002 库尔勒香梨进行测定。采集完光谱后,将库尔勒香梨放入研钵中研磨成匀浆,用纱布滤后得到滤液后滴在 Atago(日本,Atago)数显糖度计上,记录读数,测定 3 次,取平均值为库尔勒香梨的 SSC 值。

1.3.3 质量测定

将采集光谱的库尔勒香梨,放置于梅特勒电子天平 ME204/02(美国, METTLER TOLEDO)进行称量,测定三次,取平均值。

1.4 光谱处理与建模方法

采集的近红外原始光谱除含有样品信息外,还包含随机噪声、基线漂移及测定环境背景等干扰因素,为提高模型的稳健性和适用能力,获得高信噪比,降低背景干扰,降低模型复杂度,需对原始光谱进行预处理。采用 SG-平滑(Savitzky-Golay smooth)(窗口宽度为 5 点、拟合 3 次)、SG 卷积导数(窗口宽度为 5 点、拟合 3 次)、多元散射校正(MSC)、标准正态能量变换(SNV)、归一化(normalization)等多种光谱预处理方法对 600~1 000 nm 波长范围的光谱进行预处理,消除或减弱噪声,以提高预测库尔勒香梨品质的精度和使用能力。后采用偏最小二乘回归法(partial least squares regression, PLSR)建立库尔勒香梨品质的 PLS 预测模型。所建立预测模型采用校正集相关系数(correlation coefficient of calibration, r_c)、验证集相关系数(correlation coefficient of vali-

dation, r_p)、校正集均方根误差 (root mean square error of calibration, RMSEC) 和验证集均方根误差 (root mean square error of validation, RMSEP) 进行评定, 相关系数越高, 均方根误差越低, 建立的预测模型越好。

图谱处理采用北京伟创英图科技有限公司 ChemoStudio2018 化学计量学软件和 matlab8.0 对光谱图进行预处理、完成模型导入自行搭建的近红外光谱系统中, 采集光谱前调用已建立好的预测模型, 实现对库尔勒香梨无损快速分级。

2 结果与讨论

2.1 光谱最佳积分时间的确定

在采集库尔勒香梨光谱前, 确定果品品质在线无损检测系统最佳采集参数。将库尔勒香梨置于果托上, 设置样品积分时间 40 ms, 在最高速度下 ($0.5 \text{ m} \cdot \text{s}^{-1}$), 分别设置参比积分时间为 20, 10 和 5 ms, 采集 3 次库尔勒香梨的近红外光谱, 采集的光谱如图 4 所示。设置参比积分时间不同, 其光谱的吸光度也不同, 产生能量也不同。由图 5 可知, 当参比积分时间设置为 20 和 10 ms 时, 在波长 680~840 nm 范围, 参比能量已饱和, 而参比积分为 5 ms, 参比能量值低于饱和值, 故此选设置参比积分时间为 5 ms。

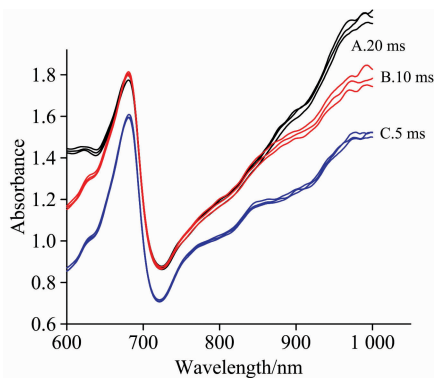


图 4 不同参比积分时间下采集光谱

Fig. 4 Collecting spectra under different reference integral time

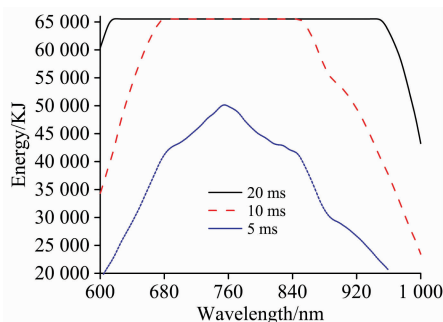


图 5 不同参比积分时间下参比能量图

Fig. 5 The reference energy under different reference integral time

2.2 库尔勒香梨质量参数的测量结果分析

利用自行搭建的在线动态检测系统采集库尔勒香梨光谱, 共采集 122 张光谱图, 采用 Kennard-Stone 分类方法对采集的光谱进行按照 3:1 进行分类, 校正光谱为 85 张, 验证光谱为 37 张; 库尔勒香梨的 SSC 和硬度校正集和验证集的统计分布如表 1 所示。校准集和预测集的 SSC 在 $11.2 \sim 15.0^\circ\text{Bri}$ 和 $12.0 \sim 15.0^\circ\text{Bri}$, 硬度的校正集和预测集为 $4.27 \sim 7.65$ 和 $4.05 \sim 7.50 \text{ N}$ 之间, 为了消除苹果的大小及重量影响, 校准集的重量范围为 $113.26 \sim 138.55 \text{ g}$ 与预测集的 $108.56 \sim 140.80 \text{ g}$ 。

表 1 库尔勒香梨物理参数的校正与预测集统计

Table 1 Statistics of physical parameters of pear samples in calibration and prediction sets

类型	校正			验证集		
	质量 /g	SSC / $^\circ\text{Bri}$	Firmness /N	质量 /g	SSC / $^\circ\text{Bri}$	Firmness /N
最大值	138.55	15.0	7.65	140.80	15.0	7.50
最小值	113.26	11.2	4.27	108.56	12.0	4.05
均值	123.89	13.49	3.38	119.26	13.50	3.45
标准偏差	20.23	0.83	0.70	24.23	0.77	0.67

采集库尔勒香梨的波长范围为 600~1 000 nm, 所有的光谱为漫透射光谱。图 6 是分别在 S1 ($0.3 \text{ m} \cdot \text{s}^{-1}$) 和 S2 ($0.5 \text{ m} \cdot \text{s}^{-1}$) 采集的平均光谱图, 在移动速度为 S1 和 S2 下, 光谱在 680 nm 处有最大吸收峰, 在 725 nm 处有波谷, 说明不同运行速度并未改变其光谱的吸收峰和波谷的位置。但低速 S1 运行速度下采集光谱的吸光度高于 S2, 说明不同移动速度下, 采集样品相同部位的信息, 所采集光谱存在差异。由于不同速度下的光谱存在的差异性, 采取预处理方法去除或者减小这些差异是十分必要的。

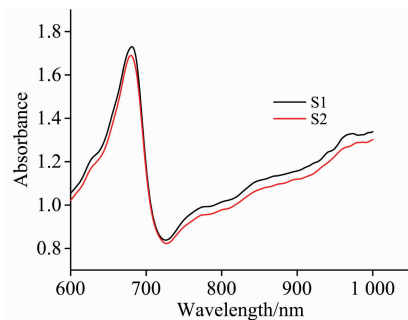


图 6 S1 和 S2 速度下样品的平均光谱

Fig. 6 Average spectra of samples under S1, S2

2.3 库尔勒香梨 SSC 在线检测模型的建立

由表 2 可知, 不同移动速度下, 采用不同预处理方法, 用 PLS 建立的模型预测性能也不一样。与其他预处理方法相比, 移动速度为 $0.3 \text{ m} \cdot \text{s}^{-1}$ 下, 采用 SG-DER 预处理效果建立模型效果最好, 其校正集和验证集相关系数分别为 0.895 1 和 0.875 6, 均方根误差分别为 $0.393 2^\circ\text{Bri}$ 和 $0.413 1^\circ\text{Bri}$ 。移动速度为 $0.5 \text{ m} \cdot \text{s}^{-1}$ 下, 同样采用 SG-DER 预处理建立模型效果最好。其校正集和验证集相关系数分别

为 0.901 4 和 0.880 2 均方根误差分别为 0.380 2°Bri 和 0.391 5°Bri。从建模效果来看, 速度 0.5 m·s⁻¹ 下建立模型略优于 0.3 m·s⁻¹ 下建立的 SSC 预测模型; 预处理方法采用 SG-DER, 在两种速度下建立 SCC 模型都最好, 故此后续采用其他算法建立预测模型, 其预处理方法采用 SG-DER。

表 2 采用不同预处理方法建立 PLS 模型
预测库尔勒香梨的 SSC

Table 2 Prediction results of SSC of pear samples by PLS model using different preprocessing methods

名称	预处理方法	因子数	校正集		预测集	
			RMSEC	r_c	RMSEP	r_p
S1	RAW	8	0.415 4	0.825 6	0.483 4	0.805 1
	SG-DER	8	0.393 2	0.895 1	0.413 1	0.875 6
	SNV	8	0.395 4	0.876 2	0.450 2	0.821 4
	MSC	9	0.406 5	0.886 1	0.435 6	0.863 2
	SGS	9	0.406 2	0.870 2	0.411 2	0.850 3
	NOR	7	0.453 2	0.815 6	0.462 1	0.801 1
S2	RAW	8	0.435 1	0.835 6	0.502 1	0.790 1
	SG-DER	7	0.380 2	0.901 4	0.391 5	0.880 2
	SNV	8	0.405 6	0.851 4	0.452 3	0.825 1
	MSC	9	0.398 9	0.890 1	0.423 5	0.856 2
	SGS	9	0.476 5	0.798 5	0.456 1	0.875 1
	NOR	7	0.486 5	0.815 3	0.491 2	0.801 3

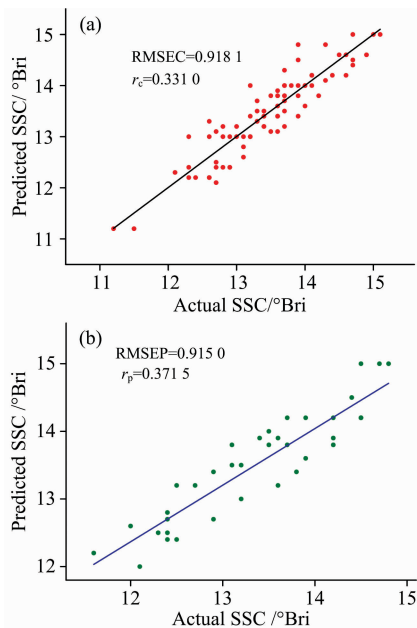


图 7 CARS-PLS 库尔勒香梨 SSC 的预测值和实际值散点图

Fig. 7 Plots of actual versus predicted values of pear SSC obtained with CARS-PLS

2.4 库尔勒香梨硬度在线检测模型的建立

由表 3 可知, 在移动速度为 0.3 和 0.5 m·s⁻¹ 下, 采用 SGS 处理建立模型效果最好。移动速度为 0.3 m·s⁻¹ 下, 其校正集和验证集相关系数分别为 0.820 2 和 0.801 0, 均方根误差分别为 0.450 8 和 0.470 8 N。移动速度为 0.5 m·s⁻¹

的校正集和验证集相关系数分别为 0.792 3 和 0.780 1, 均方根误差分别为 0.486 2 和 0.505 6 N。速度 0.3 m·s⁻¹ 的模型略优于 0.5 m·s⁻¹ 的硬度预测模型, 而在两种速度下采用预处理 SGS, 建立模型都最好, 故此后续库尔勒香梨硬度预测模型采用 SGS 预处理方法。

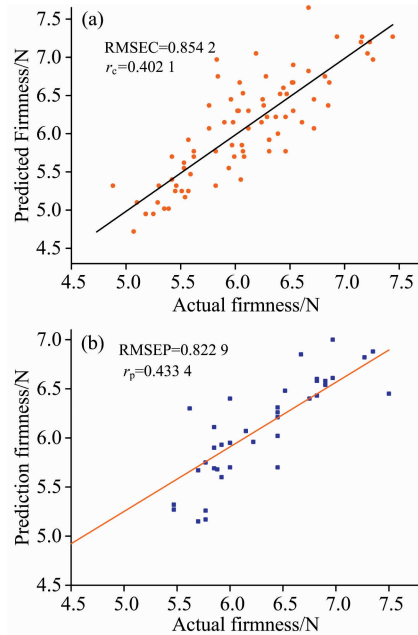


图 8 SPA-PLS 库尔勒香梨硬度的预测值和实际值散点图

Fig. 8 Plots of actual versus predicted values of pear firmness obtained with SPA-PLS

表 3 采用不同预处理方法建立 PLS 模型预测
库尔勒香梨的硬度

Table 3 Prediction firmness values of samples by PLS model using different preprocessing methods

名称	预处理方法	因子数	校正集		预测集	
			RMSEC	r_c	RMSEP	r_p
S1	RAW	10	0.614 3	0.660 2	0.650 6	0.626 0
	SG-DER	11	0.550 5	0.749 0	0.586 9	0.713 2
	SNV	11	0.472 6	0.780 9	0.489 8	0.770 3
	MSC	11	0.580 5	0.710 5	0.540 2	0.736 0
	SGS	10	0.450 8	0.820 2	0.470 8	0.801 0
	NOR	11	0.490 5	0.780 3	0.509 5	0.745 6
S2	RAW	10	0.654 1	0.590 2	0.630 2	0.606 0
	SG-DER	11	0.506 2	0.770 8	0.523 6	0.750 2
	SNV	12	0.534 8	0.743 1	0.556 8	0.725 6
	MSC	12	0.528 9	0.750 2	0.565 2	0.720 3
	SGS	12	0.486 2	0.792 3	0.505 6	0.780 1
	NOR	10	0.603 6	0.689 2	0.632 0	0.645 6

由表 2 可知, 在同一波段范围, 采用相同预处理方法, 用 PLS 建立的 SCC 预测模型, 移动速度为 0.5 m·s⁻¹ 建立模型更为稳定。由表 3 可知, 采用相同预处理方法用 PLS 建立的硬度预测模型, 移动速度为 0.3 m·s⁻¹ 下建立模型更好。说明移动速度对建立库尔勒香梨预测模型产生影响。为

了减少不同移动速度对库尔勒香梨的 SSC 和硬度预测模型的影响, 建立混合速度(S3)预测模型。随机挑选 S1 和 S2 采集的光谱图, 用于预测集各 43 条, 用于校正集各 18 条。分别采用竞争性自适应重加权算法 (competitive adaptive re-weighted sampling, CARS)^[8-9]、连续投影算法 (successive projections algorithm, SPA)^[10-11] 筛选特征波长, 再进行 PLS 建模, 比较 PLS, CARS-PLS 和 SPA-PLS 3 种建模方法对库尔勒香梨的 SSC 和硬度预测效果影响。

2.5 不同建模方法库尔勒香梨 SSC、硬度在线检测模型的影响

由表 4 可知, 混合速度模型下, 采用 CARS 和 SPA 算法建立的 SSC 和硬度预测模型与原始相比, 都可以降低参与建模的主因子数 (如图 10), 而且采用 CARS 和 SPA 算法使得建模的波长数量大大降低, 说明这两种方法都可剔除噪声过大和去除不相关变量的变量, 这与郭志明和刘燕德的研究类似^[12-13]。不同品质指标, 采用相同的算法, 所选的特征变量不相同。对于 SSC 预测模型, 采用 CARS 对库尔勒香梨近红外光谱数据进行特征变量选取, 其中从全波长 401 个光谱中筛选出 24 个变量, 建立了 SSC CARS-PLS 模型, 从全波长 401 个光谱中筛选出 38 个变量, 建立了硬度 CARS-PLS 模型。其 r_p 和 RMSEP 分别为 0.915 0, 0.371 9 和 0.801 3, 0.502 3。采用 SPA 对库尔勒香梨近红外光谱数据进行特征变量选取, 从全波长 401 个光谱中筛选出 26 个变量, 建立了 SSC SPA-PLS 模型, 从全波长 401 个光谱中筛选出 32 个特征变量, 其 r_p 和 RMSEP 和 0.900 2, 0.381 5 和 0.821 0, 0.492 0。

表 4 可溶性固形物建模效果比较

Table 4 Comparison of different soluble solids content models

预处理方法	变量数	主因子数	校正集		预测集	
			RMSEC	r_c	RMSEP	r_p
SSC						
PLS	501	9	0.366 8	0.905 6	0.3915 6	0.902 2
CARS-PLS	24	6	0.331 0	0.918 1	0.371 5	0.915 0
SPA-PLS	26	8	0.361 4	0.910 1	0.381 4	0.900 2
Firmness						
PLS	501	13	0.480 2	0.825 1	0.562 3	0.766 0
CARS-PLS	38	11	0.455 1	0.832 1	0.502 3	0.801 3
SPA-PLS	32	11	0.407 1	0.853 0	0.492 0	0.821 0

采用 CARS 和 SPA 建立的 SSC 和硬度模型优于用全波长建立的预测模型, 其中 SSC 模型中以 CARS-PLS 模型更优, 其 r_p 和 RMSEP 为 0.915 0 和 0.371 9, 硬度模型以 SPA-PLS 模型更优, 其 r_p 和 RMSEP 为 0.821 0 和 0.492 0。由图 10 可知, SSC 模型采用 CARS 选取 24 个特征变量为 606, 669, 690, 697, 706, 711, 719, 726, 735, 745, 749, 773, 784, 792, 854, 867, 879, 886, 893, 906, 920, 931, 941 和 951 nm。由图 11 可知, 硬度模型采用 SPA 选取 32 特征变量 611, 620, 633, 641, 654, 672, 680, 689, 704, 719, 730, 741, 750, 755, 763, 775, 780, 788, 802, 811, 820, 834, 840, 903, 910, 917, 923, 927, 932, 937, 942 和 946 nm。

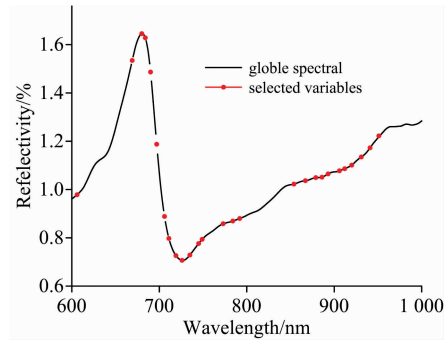


图 9 CARS 选择的特征光谱变量

Fig. 9 Distribution of selected from the global model by CARS for prediction of SSC using the full spectra

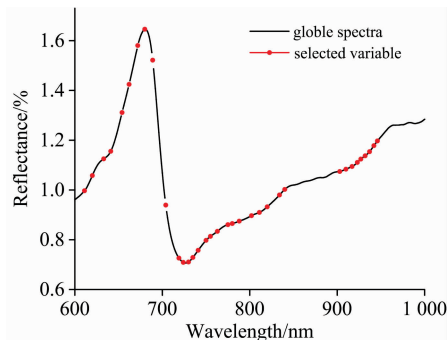


图 10 SPA 选择的特征光谱变量

Fig. 10 Spectral variables of selected by SPA for prediction of firmness using the full spectra

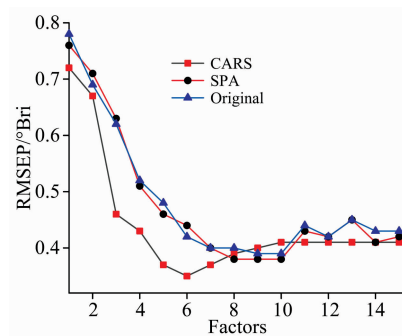


图 11 SSC 模型主因子数

Fig. 11 Principal componet factor of SSC model

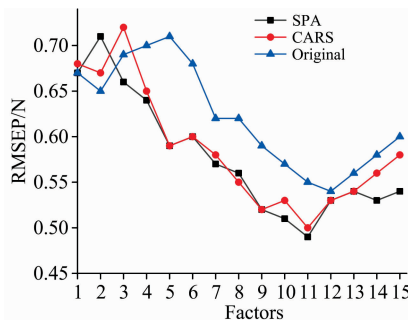


图 12 硬度模型主因子数

Fig. 12 Principal componet factor of firmness model

3 结 论

设计了可用于检测库尔勒香梨内部品质的在线无损检测系统,该系统主要由光谱采集单元、数据采集单元、品级分选单元及系统主控单元组成。基于该系统,利用可见光近红外漫透射技术研究不同移动速度对在线检测库尔勒香梨品质影响。采用不同预处理方法对两种不同速度下采集光谱进行处理,采用 PLS 建立库尔勒香梨 SSC 和硬度模型,结果表明:用 SG-DER 方法、SGS 预处理建立 SSC 预测模型、硬度模型更为稳健。在两种不同移动速度下,移动速度为 $0.5 \text{ m} \cdot \text{s}^{-1}$ 建立 SSC 预测模型优于 $0.3 \text{ m} \cdot \text{s}^{-1}$,而在 $0.3 \text{ m} \cdot \text{s}^{-1}$

建立的硬度模型优于 $0.5 \text{ m} \cdot \text{s}^{-1}$ 。说明不同移动速度对库尔勒香梨品质预测产生影响。建立混合速度(S3)模型,采用 CARS 方法,从 501 个光谱中筛选出 24 个变量,建立了 SCAR-PLS 模型,其中建立的 SSC 模型较好,其 r_p 和 RMSEP 分别为 0.915 0 和 0.371 9。采用 SPA 方法,从 501 个光谱中筛选出 26 个变量,建立硬度模型较好,其 r_p 和 RMSEP 分别为 0.821 0 和 0.492 0。

本文研究了在线移动速度对库尔勒香梨品质检测的影响,为研制升级在线分选系统提供基础,未来的研究将采用品种果品,研究多种不同移动速度对建立果品品质预测模型的影响。进一步分析,以实现更准确的和稳健模型。

References

- [1] Xu X, Xu H, Xie L, et al. *Journal of Food Measurement and Characterization*, 2018, 13(5): 506.
- [2] Chen H, Liu Z, Cai K, et al. *Vibrational Spectroscopy*, 2017, 94: 7.
- [3] Lee H, Kim M S, Lim H S, et al. *Biosystems Engineering*, 2016, 148: 138.
- [4] SUN Tong, MO Xin-xin, LIU Mu-hua(孙 通,莫欣欣,刘木华). *Spectroscopy and Spectral Analysis(光谱学与光谱分析)*, 2018, 38(5): 1406.
- [5] LI Long, PENG Yan-kun, LI Yong-yu, et al(李 龙,彭彦昆,李永玉,等). *Transactions of the Chinese Society of Agricultural Engineering(农业工程学报)*, 2018, 34(9): 275.
- [6] McGlone V A, Martinsen P J. *Journal of Near Infrared Spectroscopy*, 2004, 12(1): 37.
- [7] Sun T, Lin H, Xu H, et al. *Postharvest Biology and Technology*, 2009, 51(1): 86.
- [8] Xie C, Xu N, Shao Y, et al. *Spectrochimica Acta Part A: Molecular & Biomolecular Spectroscopy*, 2015, 149: 971.
- [9] Fan S, Guo Z, Zhang B, et al. *Food Analytical Methods*, 2016, 9(5): 1333.
- [10] Yu X, Lu H, Di W. *Postharvest Biology & Technology*, 2018, 141: 39.
- [11] Huang Y P, Lu R F, Chen K J. *Journal of Food Engineering*, 2018, (222): 185.
- [12] GUO Zhi-ming, HUANG Wen-qian, CHEN Quan-sheng, et al(郭志明,黄文倩,陈全胜,等). *Modern Food Science & Technology(现代食品科技)*, 2016(9): 147.
- [13] LIU Yan-de, SHI Yu, CAI Li-jun, et al(刘燕德,施 宇,蔡丽君,等). *Transactions of the Chinese Society of Agricultural Engineering(农业机械学报)*, 2013, 44(9): 138.

Effects of Prediction Model of Kolar Pear Based on NIR Diffuse Transmission under Different Moving Speed on Online

CHEN Dong-jie^{1,2}, JIANG Pei-hong^{1,2}, GUO Feng-jun^{1,2}, ZHANG Yu-hua^{1,2*}, ZHANG Chang-feng^{1,2}

1. National Engineering Research Center for Agricultural Products Logistics, Ji'nan 250103, China

2. Shandong Key Laboratory of Storage and Transportation Technology of Agricultural Products, Ji'nan 250103, China

Abstract With the aim of solving problems related to cost, and the complicated structure of the online grading and inspection system for detecting the quality of pears, the online non-destructive system was designed for inspecting and classification of the internal quality of pears. Based on the system, the effects of prediction models on the Soluble Solids Content (SSC) and firmness of pears were researched under the different moving speeds (0.3 and $0.5 \text{ m} \cdot \text{s}^{-1}$). Collected spectra from the same position of the pear were discrepancy at different moving speeds. Due to the discrepancy in the collected spectra, adapting spectral pre-processing methods, as SG-smoothing, SG-convolution derivative, multiple scattering correction (MSC), standard normal energy transformation (SNV), Normalization, was to eliminate differences. Adopt Partial Least Squares (PLS), prediction models of SSC and hardness for Korla Pears were established at moving speeds of $0.3 \text{ m} \cdot \text{s}^{-1}$ (S1) and $0.5 \text{ m} \cdot \text{s}^{-1}$ (S2). The results showed that the established SSC prediction model at $0.5 \text{ m} \cdot \text{s}^{-1}$ was more effective than $0.3 \text{ m} \cdot \text{s}^{-1}$ by using SG-DER (Savitzky-Golay Derivative) processing spectrogram. The correlation coefficient of the prediction set, and the root mean square errors of prediction were to be 0.8802 and 0.3915°Brix respectively. However, when the moving speed was $0.3 \text{ m} \cdot \text{s}^{-1}$, established the SSC model, by adapting SGS (Savitzky-Golay Smooth) processing spectrogram, was more robust than at $0.5 \text{ m} \cdot \text{s}^{-1}$. Its correlation coefficient of the prediction set, and the root mean square errors of prediction were to be 0.8202 and 0.4708 N respectively. Afterwards two speed hybrid prediction models were established. Competitive adaptive re-weighted sampling (CARS) and Successive projections algorithm (SPA) were used to select the characteristic variables, and PLS was used to establish hardness and SSC prediction models at mixed speeds. In view of the perspective of the model effect, SPA and CARS effectively reduced the number of variables, improving the online prediction ability and processing data speed, and enhancing the robustness of the model. Using CARS to select 24 variables from a total of 501, then which established the CARS-PLS model. Establishing the SSC prediction model was more efficient, and its correlation coefficient of the prediction set and root mean square errors of prediction were calculated as 0.9150 and 0.3719°Brix respectively. Using SPA to select, 32 variables were selected from a set of 501, and a firmness model was established. The correlation coefficient of the prediction set and the root mean square errors of prediction were ascertained as 0.8210 and 0.4920 N respectively. Establishing predictive quality model at the mixing speed is more robust than at the single speed. The research showed that the different moving speeds have different effects on the fruit quality prediction models. The research provides technical support for on-line classification of fruit quality.

Keywords Near-infrared spectroscopy; Korla pear; Different movement speeds; On-line inspection

(Received May 23, 2019; accepted Sep. 2, 2019)

* Corresponding author