

尺寸差异对脐橙糖度可见近红外光谱检测模型影响

刘燕德, 饶宇, 孙旭东, 姜小刚, 徐海, 李雄, 王观田, 徐佳

华东交通大学机电与车辆工程学院, 江西 南昌 330013

摘要 糖度是评价脐橙内部品质的重要指标之一, 由于水果自身尺寸差异, 造成糖度预测模型稳健性差, 预测精度不高, 因此消除水果尺寸差异带来的影响, 对提高水果分选模型精度具有重要意义。对比分析了脐橙漫透射、多点发射与接收及环形发射与接收漫反射光谱, 其中, 不同检测平台上, 由于光程差的不同, 大果光谱能量均比小果光谱能量要强, 而环形发射与接收漫反射光谱能量要强于其他两种光谱, 漫透射光谱能量最弱, 波峰与波谷位置大致相同。分别建立不同检测方式下脐橙尺寸预测模型, 其中, 漫透射检测方式下尺寸预测模型的预测集相关系数为 0.60, 预测集均方根误差为 3.95 mm, 多点发射与接收漫反射检测方式下尺寸预测模型的预测集相关系数为 0.97, 预测集均方根误差为 1.46 mm, 环形发射与接收漫反射检测方式下小果预测模型的预测集相关系数为 0.96, 预测集均方根误差为 1.73 mm。分别建立三种不同检测方式下大果、小果、混合果以及多元散射校正预处理的混合果糖度预测模型, 小果的糖度预测模型精度均要高于大果和混合果, 其中漫透射检测方式下小果预测模型的预测集相关系数为 0.76, 预测集均方根误差为 0.81°Brix, 多点发射与接收漫反射检测方式下小果预测模型的预测集相关系数为 0.72, 预测集均方根误差为 0.97°Brix, 环形发射与接收漫反射检测方式下小果预测模型的预测集相关系数为 0.72, 预测集均方根误差为 0.93°Brix。经过多元散射校正预处理光谱后, 近红外漫透射光谱的混合果模型精度要优于小果的模型, 模型预测集相关系数为 0.84, 预测集均方根误差为 0.64°Brix, 而在两种漫反射检测方式中, 多混合果模型精度反而降低。实验结果表明: 在漫透射检测方式中, 使用多元散射校正预处理光谱可以消除尺寸差异影响, 在漫反射检测方式中, 先进行尺寸分选, 再进行糖度分选, 也可以避免尺寸差异带来的影响。该研究为大宗水果快速在线分选提供了参考和理论支持。

关键词 脐橙; 近红外; 尺寸差异; 无损检测; 糖度

中图分类号: S24 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2020)10-3241-06

引言

糖度(SSC)是影响水果口感的一个重要因素, SSC 差异在 2°Brix 以上, 水果口感差异明显^[1]。近年来, 人们对口感品质十分重视, 传统外部品质分选难以满足口感品质分选需求, 可见/近红外光谱技术作为一种快速无损检测手段, 目前已经广泛应用于苹果^[2]、梨、桃^[3]等水果的内部品质检测。

崔丰娟^[4]等研究了苹果动态在线检测的传输速度对模型的精度的影响。曹松涛在他的学位论文中报道了对梨样品相关因素对内部品质普适模型的稳健性影响, 样品的尺寸差异、产地差异、品种差异都是主要影响, 通过对光谱进行一

阶、二阶微分预处理, 基本能够消除差异影响, 提高了模型的预测精度。Saranwong^[5]等对比分析了两种不同近红外漫反射仪器在芒果品质检测上的应用。Greensill^[6]等使用自行研发的光学系统, 研究了硬皮甜瓜果实不同果肉位置的 SSC 取样准确度, 结果表明, 果实赤道部位 SSC 取样值与光谱建模模型准确度最高, 能代表水果的真实糖度值。Mizuki^[7]等结合偏最小二乘判别分析, 应用苹果分选机顶部和底部的光谱仪, 建立了苹果内部褐变发生的判别模型。

本工作主要研究了脐橙的尺寸差异对可见近红外漫透射及两种不同漫反射检测方式模型的影响, 对比分析了大果、小果及混合果模型的预测精度, 考察了近红外光谱模型及多元散射校正预处理方法对消除尺寸差异的效果。

收稿日期: 2019-08-02, 修订日期: 2019-12-27

基金项目: 国家自然科学基金项目(31760344), 江西省杰出青年人才资助计划项目(20171BCB23060), 水果光电检测技术能力提升项目(S2016-90)资助

作者简介: 刘燕德, 女, 1967年生, 华东交通大学机电与车辆工程学院教授 e-mail: jxliuyd@163.com

1 实验部分

1.1 材料

实验所用的脐橙样品购于南昌市某水果市场,将脐橙表面擦净,在 25 °C 的室温下静置 24 个小时,依次采集脐橙样品的横纵径、重量、糖度的理化指标如表 1 所示。实验共采用 480 个样品,按照脐橙的横纵径指标划分大小果,其中横径范围在 75~90 mm,纵径在 80~95 mm 为大果共 312 个,横径范围在 60~75 mm,纵径在 65~85 mm 为小果共 168 个,大果建模集为 234 个,预测集为 78 个,小果建模集为 126 个,预测集 42 个,混合果建模集 360 个,预测集 120 个。

表 1 脐橙的理化指标范围统计结果
Table 1 Statistical results of physical and chemical indicators of navel orange

果型	数据集	样品数	横径范围	纵径范围	糖度范围
大果	建模集	234	75~90	80~95	11.4~16.6
	预测集	78	75~90	80~95	11.4~16.4
小果	建模集	126	60~75	65~80	11.5~16.2
	预测集	42	60~75	65~80	11.1~16.5

1.2 装置和光谱采集

可见近红外漫反射光谱采集平台如图 1 所示,实验平台搭建在重量分选装置上的出果口前,考虑实际应用场景,光源发射与接收探头安装在距离果杯 15 cm 处。采用 Ocean Optics 公司的 QE PRO 光谱仪作为实验的光谱采集装置,其波长范围为 350~1 150 nm,光源为 1 个 12 V、100 W 的卤钨灯,采用稳压电源供电。光纤探头分别为环形发射与接收和多点发射与接收漫反射式光纤探头,均设计为将发射光源与接收光源装置集成在同一探头上,通过卤钨灯提供直射光源,从光源接收端口进入光纤,由探头发射端发出光照,光信号经过样品内部后,经过吸收、反射等,由探头接收端口接收到漫反射光信号,经过光纤传输并触发光谱仪进行数

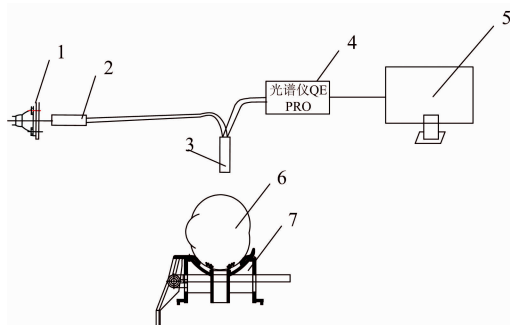


图 1 近红外漫反射在线检测装置

1: 光源; 2: 光源接收端口; 3: 光源发射与接收端口;
4: QE PRO 光谱仪; 5: 电脑; 6: 脐橙样品; 7: 果杯

Fig. 1 Near-infrared diffuse reflection on-line detecting device

1: Light source; 2: Light source receiving port; 3: Light source transmitting and receiving port; 4: QE PRO spectrometer; 5: Computer; 6: Navel orange sample; 7: Fruit cup

据保存,每触发一次,光谱采集软件便储存一条光谱。

开机预热 0.5 h,聚四氟乙烯球作为光谱参比,采集脐橙样品光谱,实验参数为:积分时间 100 ms,速度 5 个·s⁻¹,光强 100 W。光谱采集软件采集光谱,分别标记面光谱,取平均光谱作为该样品可见近红外漫反射光谱。

漫透射光谱采集装置如图 2 所示,采用海洋光学公司的 QE65000 光谱仪,波长范围为 350~1 150 nm,光源为 10 个 12 V、100 W 的卤钨灯,均匀分布在传送带两侧,通过稳压电源供电。光纤探头位于果杯正下方,光信号从样品表面进入,经过透射到达样品下方的光纤探头,探头每接收到光信号一次,光谱仪触发一次,则光谱采集软件保存一条光谱。果杯设计有弹性避光圈与透光孔,避免杂散光与外界干扰光被光纤接收到。

开机预热 0.5 h,聚四氟乙烯球作为光谱参比,采集脐橙样品光谱,实验参数为:积分时间 100 ms,速度 5 个·s⁻¹,光强 1 000 W。采用光谱采集软件采集光谱,分别标记面光谱,取平均光谱作为该样品可见近红外漫反射光谱。

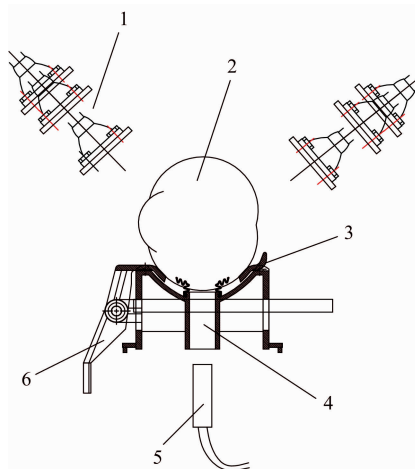


图 2 近红外漫透射在线检测装置

1: 光源; 2: 样品; 3: 遮光圈;
4: 光路; 5: 光纤探头; 6: 弹跳爪

Fig. 2 Near-infrared diffuse transmission on-line detecting device

1: Light source; 2: Sample; 3: Shading ring;
4: Optical path; 5: Fiber optic probe; 6: Bouncing claw

1.3 糖度含量判定

脐橙糖度含量由折射式数字糖度计(ATAGO PAL- α , 日本 ATAGO 公司)测定,测量前,用纯净水洗净并擦干糖度计,并进行标定,糖度值为 0°Birx,测量时,使用一次性纱布将标记部分约 5 mm 深果肉滤出少量果汁,滴于糖度计的测试窗口,重复 3 次,取 3 次的平均值作为糖度真实值,所得值以质量分数表示。

2.4 数据处理

采用 matlab2012a 将光谱数据转换为数据矩阵,再通过 Unscrambler 进行 PLS 建模。建模前先进行光谱数据预处理,再使用偏最小二乘法(PLS)分别建立漫透射及两种不同漫反射式的糖度模型。偏最小二乘法中的相关系数 r 和预测

集均方根误差 RMSEP、建模集均方根误差 RMSEC 计算公式如式(1)一式(3)所示

$$r = \frac{\sqrt{\sum_{i=1}^n (\hat{y}_i - y_i)^2}}{\sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2}} \quad (1)$$

$$RMSEC = \sqrt{\frac{1}{n_c - 1} \sum_{i=1}^{n_c} (\hat{y}_i - y_i)^2} \quad (2)$$

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n_p} (y_i - \hat{y}_i)^2}{n_p}} \quad (3)$$

式中: n 为样品总数; n_c 为校正集样品个数; n_p 为预测集样品个数; y_i 为校正集第 i 个样品的真实值, \hat{y}_i 为预测集的第 i 个样品的预测值。

2 结果与讨论

2.1 不同检测方式的尺寸差异影响分析

近红外光谱分析是一种应用于水果内部品质无损检测的重要手段,但由于样品的相关因素如样品温度、大小、产地、品种等的影响,模型的稳健性受到较大影响,其中,样品大小是一个重要的影响因素。近红外漫透射和漫反射检测方式不同大小样品光路图如图 3 所示,其中图 3(a)为漫透射检测大、小尺寸的脐橙光路图,在相同光照路径下,光在大果中的光程要比小果中的光程更长,光的能量衰减严重,光纤探头接收到的能量强度更低,而小果由于它的小尺寸,光能快速透过,光纤探头接收到的能量衰减少,强度强,携带的有效信息多。图 3(b)为漫反射检测大、小尺寸的脐橙光路图,在相同光照路径下,光在大果中的光程要比小果中的光程要短,可能穿过的果肉组织不多,携带的有效信息少,而在小果中,光程较长,穿过的果肉组织较多,携带的有效信息也多。

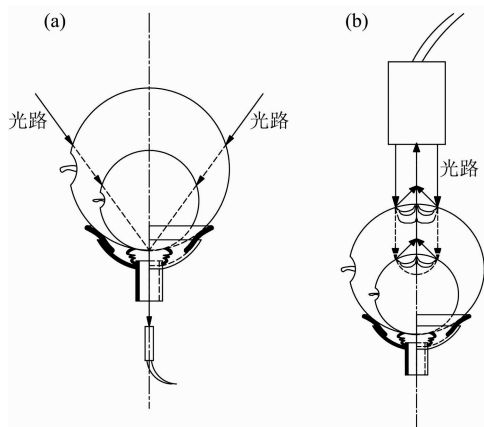


图 3 漫透射(a)和漫反射(b)上大、小尺寸的脐橙光路示意图

Fig. 3 Schematic diagrams of light paths of diffuse transmission (a) and diffuse reflection (b) for large and small navel oranges

2.2 不同检测方式的光谱对比分析

对样品的尺寸进行测量,每个样品的最大横径测量 5 次,取平均值作为样品的横径指标,考察尺寸差异对脐橙糖度在线检测精度的影响。由图 4 可知,无论是漫透射检测还是漫反射检测,小果的光谱能量强度都要比大果强。对同一样品的三种不同检测方式的光谱进行对比分析。两种漫反射光谱的能量要比漫透射光谱强,而环形发射与接收漫反射光谱又要比多点发射与接收漫反射光谱的能量强。漫反射光谱在 690 nm 左右存在吸收峰,漫透射光谱的峰在 700 和 825 nm 左右,从两种检测方式的光谱波形上来看,漫透射光谱要比漫反射光谱多一个吸收峰,且漫透射光谱集中在 600~850 nm 范围内,漫反射光谱范围则在 500~900 nm 内,比漫透射光谱范围大一点。对比分析三种不同检测方式,漫透射检测方式中,光源位于样品两侧,接收装置位于样品下方,果杯上的橡胶遮光圈能有效消除样品表面反射光及其他干扰光进入接收器;在漫反射检测方式中,光源与接收装置集成在同一光纤上,环形发射与接收的漫反射检测方式中,由于它的环形光源的结构特点,光在样品表面形成一个圆形光斑,可以有效的将外界干扰光及样品表面反射光隔离在光斑外面,达到克服杂散光的目的;多点发射与接收漫反射检测方式中,由于它是点光源,点接收方式,所以它克服样品表面反射光和其他干扰光的能力要比前两种检测方式稍差。

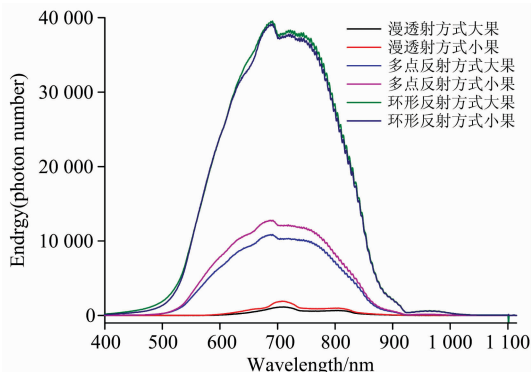


图 4 三种检测方式的大小果光谱对比

Fig. 4 Spectral comparison of fruit with different sizes using three detection methods

2.3 不同检测方式下尺寸预测模型建立

不同样品尺寸在某一波长范围光下具有一定的衰减规律,实验选用波长范围为 900~936 nm 建立近红外光谱尺寸预测模型。可见/近红外光谱波长范围 320~500 nm 是可见光波段,受颜色变化影响较大;500~900 nm 是光谱的有效信息集中波段,用来建立糖度预测模型;936~1150 nm 波段噪声和无效信息较多,故选用 900~936 nm 建立尺寸预测模型。不同检测方式下,尺寸模型预测散点图如图 5 所示。由图 5 可知,漫透射检测方式下尺寸模型的预测集相关系数为 0.60,预测集均方根误差 3.95°Brix,截距 $B=81.27$;多点发射与接收漫反射检测方式下尺寸模型的预测集相关系数为 0.97,预测集均方根误差 1.46°Brix,截距 $B=94.3$;环形发射与接收漫反射检测方式下尺寸模型的预测集相关系数为

0.96, 预测集均方根误差 1.73°Brix, 截距 $B=70.73$ 。由不同的检测方式下尺寸预测散点图可知, 漫透射检测方式的尺寸模型预测精度较低, 两种漫反射检测方式的尺寸模型预测精度较高, 可以有效的预测样品的尺寸大小。

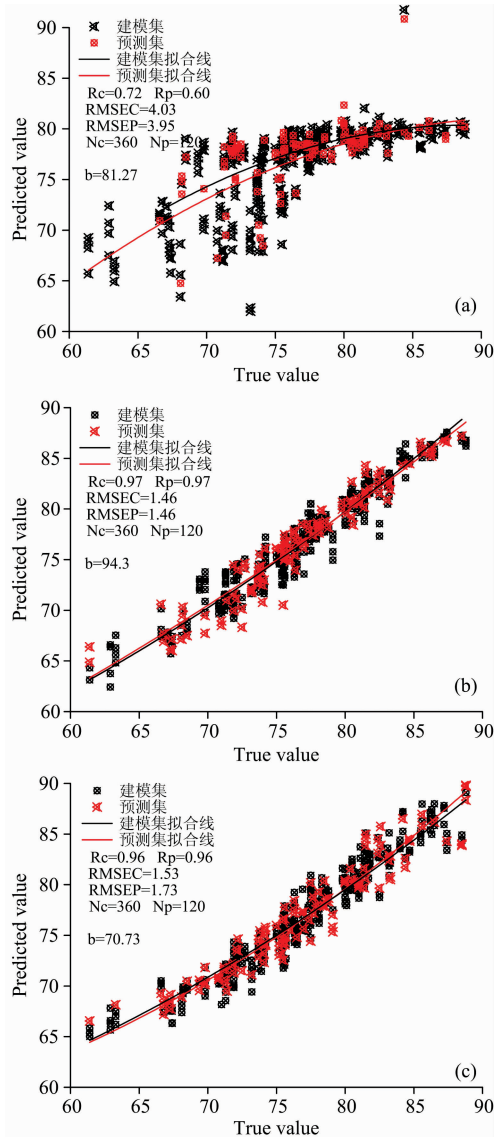


图 5 不同检测方式下尺寸模型预测结果散点图

(a): 漫透射检测方式尺寸预测散点图;

(b): 多点发射与接收漫反射检测方式尺寸预测散点图;

(c): 环形发射与接收漫反射检测方式尺寸预测散点图

Fig. 5 Scatter plot of dimensional model prediction results under different detection modes

(a): Diffuse transmission detection method size prediction scatter plot; (b): Multi-point transmit and receive diffuse reflection detection method size prediction scattered points; (c): Circular emission and reception diffuse reflection detection method size prediction scatter plot

2.4 糖度预测模型建立

王动民(本刊 34 卷 9 期)利用 MSC 方法对光谱进行预处理, 基于不同波段的光谱数据会得到不同的校正结果, 进而

影响所得定标模型的可靠性。分别建立三种不同检测方式的不同果型的 PLS 模型, 模型预测结果如表 2 所示, 通过预测集相关系数 R_p 来评价模型的预测能力。由表 2 可知, 三种不同的近红外光谱检测方式下, 小果的 PLS 模型预测能力要强于大果和混合果, 其中漫透射方式的 R_p 为 0.76, 多点发射与接收方式的 R_p 为 0.72, 环形发射与接收检测方式的 R_p 为 0.72。在相同光照路径下, 光在大果中的光程要比小果的长, 衰减严重, 接收到的光谱能量强度低, 因此携带的有效信息比小果光谱的要少很多, 所以小果的模型预测精度相对较高。

表 2 不同尺寸样品集的不同检测方式 PLS 建模集预测结果

Table 2 PLS modeling prediction results for different size sample sets by using different detection methods

检测方式	果型	R_c	RMSEC	R_p	RMSEP	PCs
漫透射	大果	0.82	0.55	0.69	0.73	8
	小果	0.86	0.66	0.76	0.81	7
	混合果	0.83	0.59	0.73	0.82	10
多点发射与接收漫反射	大果	0.88	0.51	0.84	0.64	7
	小果	0.79	0.64	0.65	0.86	8
	混合果	0.84	0.63	0.72	0.97	7
环形发射与接收漫反射	大果	0.71	0.78	0.66	0.85	8
	小果	0.65	0.83	0.62	0.89	7
	混合果	0.83	0.61	0.64	0.78	10
MSC-PLS	大果	0.88	0.56	0.72	0.93	9
	小果	0.71	0.77	0.62	0.88	9
	混合果	0.69	0.80	0.59	0.91	8

糖度预测模型公式如式(4)所示。

$$y = \sum_{i=1}^n \beta_i \gamma_i + b \quad (4)$$

其中, n 为参与建模的光谱变量数, β 为糖度相关能量谱强度, b 为模型的截距。

2.5 脐橙糖度最佳模型建立和预测

通过以上多种模型的对比分析, 漫透射检测方式模型预测效果最好, 多元散射校正预处理后, 模型预测效果有所优化, 可以消除尺寸差异带来的影响。采用多元散射校正预处理漫透射光谱, 再利用偏最小二乘(PLS)建立漫透射及漫反射的预测模型如图 6 所示。由图 6 可知, 经过 MSC 校正的漫透射光谱预测模型散点距离拟合线较为集中, 预测精度较高, 建模相关系数为 0.88, 预测相关系数为 0.84, 预测集均方根误差 RMSEP 为 0.64, 截距 $b=14.05$ 。回归系数与近红外漫透射大小果光谱曲线对比如图 7 所示, 光谱变量在 PLS 定量模型中贡献率越大, 回归系数越大。所用的光谱变量与回归系数加权求和再加上截距, 及 PLS 模型预测值, 实现了对样品的糖度值的预测。

在漫反射检测方式中, 先进行尺寸分选, 区分出大小果后, 再进行大小果的糖度分选方式, 也可以消除尺寸差异带来的影响。

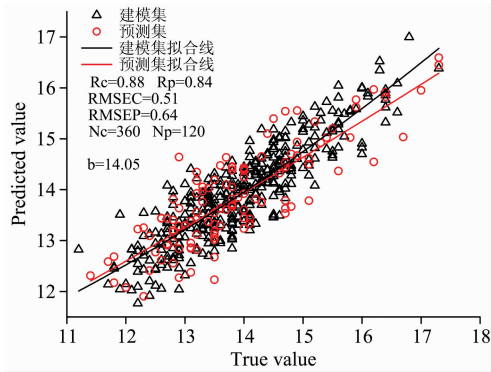


图 6 漫透射 MSC-PLS 建模散点图

Fig. 6 Diffuse transmission MSC-PLS modeling scatter plot

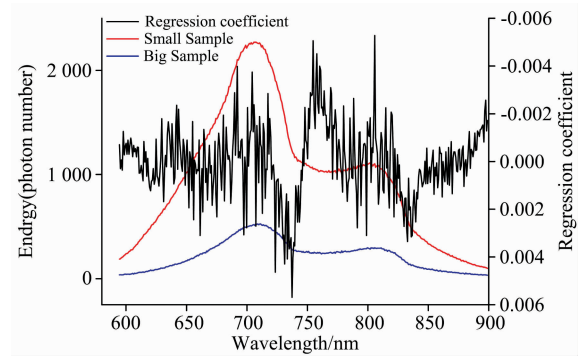


图 7 回归系数与漫透射大小果光谱对比

Fig. 7 Comparison of regression coefficient and diffuse transmission size

3 结 论

研究了样品尺寸差异对脐橙漫透射及漫反射在线无损检测精度的影响。分析了脐橙的尺寸差异在漫透射和漫反射装置上的不同影响以及脐橙三种不同检测方式的近红外光谱特性。分别建立了三种检测方式下的大果、小果、混合果、经 MSC 预处理后的混合果糖度 PLS 模型和尺寸 PLS 预测模型,其中小果的糖度 PLS 模型预测效果均最佳,在漫透射糖度预测模型中,经 MSC 预处理后,混合果预测模型精度有

所提高,预测效果要优于小果,模型的预测相关系数为 0.84,预测均方根误差为 0.64°Brix;在不同检测方式下的尺寸预测模型中,漫透射检测方式要比两种漫反射检斥方式的尺寸模型预测精度低,在漫反射检测方式下,先采用尺寸模型进行尺寸进行分选,然后再使用糖度模型进行糖度分选,最终实现不同检测方式下不同尺寸脐橙的同时在线分选。论文研究为不同果型的水果糖度同时在线无损检测和分选提供了参考和理论支持。

References

- [1] SUN Tong, MO Xin-xin, LIU Mu-hua, et al(孙 通, 莫欣欣, 刘木华). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2018, 38(5): 1406.
- [2] YANG Hao, XIONG Zhi-xin, CHEN Tong(杨 浩, 熊智新, 陈 通). Chinese Journal of Analysis Laboratory(分析实验室), 2018, 37(2): 163.
- [3] TANG Chang-bo, FANG Li-gang(唐长波, 方立刚). Jiangsu Agricultural Science(江苏农业科学), 2013, 41(11): 331.
- [4] CUI Feng-juan, ZHA Jian-wen(崔丰娟, 阚建文). Journal of Agricultural Mechanization Research, 2010, 32(11): 170.
- [5] Saranwong S, Sornsrivichai J, Kawano S. Journal of Near Infrared Spectroscopy, 2003, 11(1): 175.
- [6] Greensill C V, Walsh K B. Measurement Science and Technology, 2000, 11(12): 1674.
- [7] Mizuki Tsuta, Masatoshi Yoshimura, Kazuya Matsubara, et al. Japan Journal of Food Engineering, 2019, 2(17): 1.

Size Effect on the Near-Infrared Spectroscopy Detection Model of Navel Orange

LIU Yan-de, RAO Yu, SUN Xun-dong, JIANG Xiao-gang, XU Hai, LI Xiong, WANG Guan-tian, XU Jia
School of Mechanical and Vehicle Engineering, East China Jiaotong University, Nanchang 330013, China

Abstract Brix is one of the important indicators for evaluating the internal quality of navel orange. Due to the difference in the size of the fruit itself, the sugar content prediction model is poor in robustness and the prediction accuracy is not high. Therefore, eliminating the influence of fruit size effect is of great significance for improving the accuracy of fruit sorting model. The diffuse transmission, multi-point emission and reception, and circular emission and reception diffuse reflectance spectra of navel orange were compared and analyzed. Among different detection platforms, the spectral energy of the big fruit was stronger than that of the small fruit due to the difference of optical path difference, and the circular emission was obtained. The energy of the diffuse reflection spectrum is stronger than that of the other two spectra. The diffuse transmission spectrum energy is the weakest, and the peaks and troughs are roughly the same. The prediction model of orange navel size under different detection methods is established respectively. Among them, the prediction coefficient of the size prediction model under the diffuse transmission detection mode is 0.60, the root mean square error of the prediction set is 3.95 mm, and the size of the multi-point transmission and reception diffuse reflection detection mode. The prediction set correlation coefficient of the prediction model is 0.97. The prediction set RMS error is 1.46 mm, and the prediction set correlation coefficient of the small fruit prediction model under the ring emission and reception diffuse reflection detection mode is 0.96, and the prediction set RMS error is 1.73 mm. The mixed fructose prediction models of large fruit, small fruit, mixed fruit and multi-scattering correction pretreatment were established under three different detection methods. The precision of the sugar content prediction model of small fruit was higher than that of large fruit and mixed fruit, and diffuse transmission detection. The correlation coefficient of the prediction set of the small fruit prediction model is 0.76, the root means square error of the prediction set is 0.81°Brix, and the correlation coefficient of the prediction set of the small fruit prediction model under the multi-point transmission and reception diffuse reflection detection mode is 0.72. The square root error is 0.97°Brix, and the prediction set correlation coefficient of the small fruit prediction model under the ring emission and reception diffuse reflection detection mode is 0.72, and the prediction set RMS error is 0.93°Brix. After multi-scattering correction pre-processing spectra, the hybrid fruit model of near-infrared diffuse transmission spectrum is better than the small fruit model. The correlation coefficient of the model prediction set is 0.84, and the root means square error of the prediction set is 0.64°Brix. In the diffuse reflection detection mode, the accuracy of the multi-mixed fruit model is reduced. The experimental results show that in the diffuse transmission detection method, the multi-scatter correction pre-processing spectrum can eliminate the effect of the size effect. In the diffuse reflection detection method, the size sorting is performed first, followed by the sugar separation, which can also avoid the size effect. This study provides reference and theoretical support for the rapid online sorting of bulk fruits.

Keywords Navel orange; Near infrared; Size effect; Nondestructive testing; Sugar content

(Received Aug. 2, 2019; accepted Dec. 27, 2019)