

水果糖度可见-近红外光谱手持式检测装置开发与试验

樊书祥¹, 王庆艳¹, 杨雨森², 李江波¹, 张 驰¹, 田 喜¹, 黄文倩^{1*}

1. 北京农业信息技术研究中心, 北京 100097
2. 西南交通大学利兹学院, 四川 成都 611731

摘 要 基于可见-近红外光谱分析技术开发了手持式水果糖度检测装置, 并用于水果糖度的现场实时分析。硬件系统主要包括微型光谱仪、卤素灯、OLED 显示屏、单片机及驱动电路等。采用 Keil 5 开发工具, 用 C 语言开发单片机程序。配合上位机以 LabView 编写的光谱采集程序, 实现光谱信息的采集。以苹果和大桃作为检测对象, 对装置的检测精度和模型在 2 台装置(主机、从机)间的传递效果进行了探讨。在实验室和果园环境下, 分别获取了苹果、大桃样本在 600~950 nm 范围的可见-近红外光谱。对实验室条件下采集的主机校正集光谱进行分析, 经过平滑、最大值归一化、二阶导数等预处理后, 利用偏最小二乘算法分别建立了苹果、大桃的糖度检测模型。模型导入主机装置后, 对预测集样本进行检验。对苹果、大桃的预测集相关系数和预测均方根误差分别为 0.925, 0.587% 和 0.821, 0.613%。采用分段直接校正和基于典型相关分析算法的模型传递方法将模型由主机传递到从机。对比后发现, 基于典型相关分析算法取得了更好的模型传递结果。从机对苹果、大桃糖度的预测集相关系数和预测均方根误差分别为 0.883, 0.641% 和 0.805, 0.626%。将实验室条件下建立的模型用于树上采集到的水果光谱数据分析, 得到预测集相关系数和预测均方根误差分别为 0.866, 0.741% 和 0.816, 0.627%。整体检测结果表明, 该装置可以满足对苹果、大桃糖度的有效检测, 借助模型传递算法, 实现了模型在不同装置间的共享和有效传递, 且实验室环境下采集的数据建立的模型可以用于树上水果糖度的有效检测, 该装置具有较大的经济价值和应用前景。

关键词 无损检测; 水果; 可见-近红外光谱; 光谱分析; 糖度; 模型传递

中图分类号: S37 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2021)10-3058-06

引 言

可见-近红外光谱分析为代表的无损检测技术在水果糖度的快速、无损检测方面得到了广泛应用^[1]。水果糖度的便携式或手持式装置, 可以实现水果糖度的快速、现场检测, 满足水果样品抽检、定级以至田间管理^[2]。目前, 如久保田株式会社, FELIX 以及 Atago 等都推出了相关产品。但上述国外公司的产品价格较高, 检测样品单一, 不易于后期模型和设备的维护, 使用成本较高。

在近几年的研究中, 国内外研究人员借助小型商品化光谱仪器, 也开发了具有类似功能的样机。样机的开发主要集中在用于水果光谱采集探头的设计以及后续对光谱仪的控制。Yang^[3]、Guo^[4]等开发了用于猕猴桃、油桃糖度检测的

便携式装置, 王凡^[5]、郭志明^[6]等设计了用于番茄糖度、番茄红素等指标分析的检测装置。除了利用商品化光谱仪外, Yu 等^[7]利用线性渐变滤光片自行开发了用于梨果糖度检测的便携式装置。赵娟等^[8]基于所选 10 个特征波长的 LED 和硅光电二极管, 开发了用于苹果糖度检测的便携式装置, 虽然成本有所降低, 但精度有所下降。借助便携式光谱仪可以直接用于树上柑橘^[9]等水果品质的无损分析, 但大都没有形成最终检测装置。分析发现, 上述装置成本较高, 大都针对同种水果进行评价, 且均针对开发的一台装置进行检验, 缺少模型在不同装置间的传递等方面的探讨。

在前期研究中, 本课题组以海洋光学 USB2000+ 光谱仪和嵌入式操作系统为核心, 成功研制了用于多种水果糖度检测的背带式检测装置^[10]。虽然取得了较好的检测结果, 但成本高, 携带不便。本研究继续以微型光谱仪为核心, 开发了

收稿日期: 2021-03-01, 修订日期: 2021-07-09

基金项目: 北京市科技计划课题(Z191100004019007), 国家自然科学基金项目(31801262), 北京市农林科学院青年基金项目(QNJ201818)资助

作者简介: 樊书祥, 1989 年生, 北京农业信息技术研究中心助理研究员 e-mail: fansx@nrcita.org.cn

* 通讯作者 e-mail: huangwq@nrcita.org.cn

成本更低,更轻便的手持式水果糖度检测装置,并在实验室和果园环境对装置的检测精度进行检验,同时探讨了糖度模型在不同装置间的传递。

1 实验部分

1.1 硬件系统设计

水果糖度手持式可见-近红外光谱检测装置硬件系统主要包括微型光谱仪、卤素灯、单片机及外围电路、锂电池、OLED显示屏等部分。为了减小手持式装置的体积、节约成本,该装置不具备存储光谱数据的功能,因此在采集光谱数据时需要将装置与计算机通过串口相连。待模型构建并导入单片机后,方可单独使用。

光谱仪为微型光谱仪(C11708MA,滨松,日本),波长范围为 640~1 050 nm,该光谱仪体积小,重量轻。采用小型

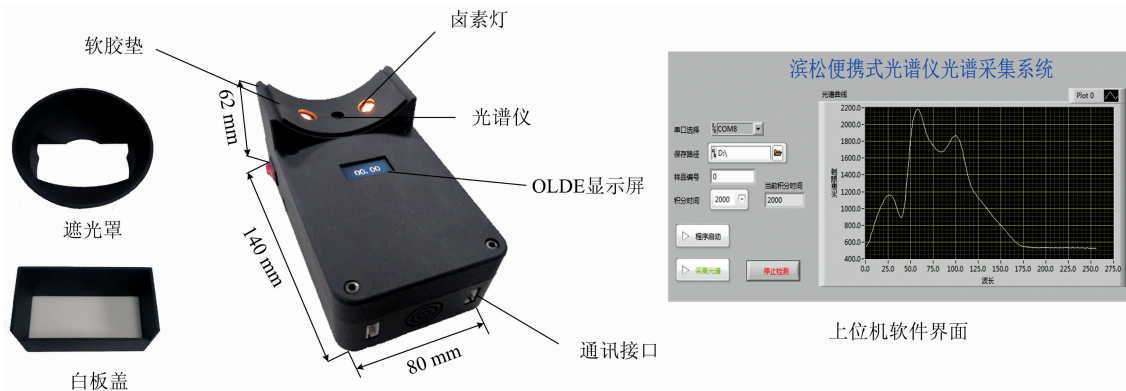


图 1 水果糖度手持式检测装置图

Fig. 1 Image of the developed handheld visible/near infrared device for detecting fruit sugar content

1.2 软件系统设计

检测装置的软件系统基于 Keil 5 平台开发,采用 C 语言进行程序编写。实现水果漫透射可见-近红外光谱信息的采集、数据传输、结果显示等功能。上位机配套 LabView 编写的光谱采集软件,通过串口和手持式检测装置通讯,通过上位机设定手持式装置的积分时间,完成水果光谱信息在指定积分时间下的获取。

1.3 材料及水果光谱采集

以大桃、苹果作为实验对象。样机共开发两台,分别编号为 P-NIR-01 和 P-NIR-02。2020 年 8 月份在北京平谷大桃果园进行实验。借助笔记本电脑,完成树上水果光谱信息采集,采集前获取白参考光谱,随后将遮光罩套在探头上用于树上水果光谱采集。采集水果光谱后对样品进行编号,共采摘大桃 150 个,积分时间设置为 900 ms。田间光谱信息采集只利用了 P-NIR-01 装置。采集完成后,运回实验室。从中随机选取 110 个大桃作为校正集,剩余样本作为预测集。实验室环境下的光谱采集,同时采用上述 2 台装置获取大桃光谱信息,积分时间都设置为 900 ms。

采取同样方法,采用编号为 P-NIR-01 装置于 2020 年 10 月在北京昌平苹果园获取树上苹果光谱信息,积分时间为 600 ms,共采集富士苹果 140 个。从中随机选取 100 个苹果

卤素灯,功率 1.5 W。检测时采用漫透射方式^[5],将水果放置于探头之上,赤道部位紧贴中间接收孔,卤素灯发出的光,经过与水果作用后,被距卤素灯 16 mm 接收孔下方的微型光谱仪接收。

单片机选用 STM32 单片机,自行设计光谱仪驱动电路。实现光谱按设定积分时间进行获取。为增加光谱稳定性,获取 3 次水果光谱的平均值作为最终光谱。串口用于和上位机软件通讯。OLED 显示屏用于显示水果糖度预测值。电源容量 6 000 mah 锂电池,工作时间约为 2.5~3 h。在采集可见-近红外光谱时常常进行白参考校正。本研究针对开发的装置结构,设计了如图 1 所示的白参考校正模块。另外,该装置配有遮光罩用于室外水果的光谱采集,减少阳光等杂散光对光谱采集影响。图 1 为开发的手持式水果糖度可见-近红外光谱检测装置实物图以及配套的上位机光谱采集界面。

作为校正集。运回实验室后采用 2 台装置进行光谱信息采集,采集参数同样设置为 600 ms。

在实验室采集条件下,研究了模型在上述 2 台装置间的传递。选择 P-NIR-01 作为主机(也称作源机),P-NIR-02 作为从机(也称作目标机)。从校正集光谱中选取一定数量样本作为标准样品,用于确定主机光谱与从机光谱的对应关系。主机对应的校正集用于模型构建,主机的预测集用于对模型的检测精度进行评价,从机的预测集用于模型传递的评价。

1.4 水果糖度实测值获取

光谱采集完成后,从光谱采集的赤道部位切取厚度为 10 mm 左右带果皮果肉,经纱布过滤挤汁后,滴到折光仪(PAL-1, ATAGO, 日本)镜面,读数并记录。结果如表 1 所示。

表 1 水果糖度实测值

Table 1 Statistics of fruit sugar content measurement

水果	集合	样本数	最小值 /%	最大值 /%	平均值 /%	标准差 /%
苹果	校正集	100	8.40	15.25	12.56	1.35
	预测集	40	7.90	14.35	12.61	1.36
大桃	校正集	110	10.50	16.80	12.73	1.01
	预测集	40	9.60	14.50	12.70	1.05

1.5 光谱数据处理

由于采集的光谱信息在波段首尾两端均存在较大噪声,因此选择 600~950 nm 范围进行后续分析。首先参照式(1)对原始光谱数据进行白校正得到透射率光谱(Transmittance)。

$$\text{Transmittance} = \frac{\text{Sample}}{\text{White}} \quad (1)$$

式(1)中, Sample 为采集到的水果原始光谱, White 为白板盖放置在探头上方时获取的白参考光谱。由于原始光谱数据中存在诸多噪声、毛刺等干扰信号,采用包括 5 点平均平滑、归一化、标准变量变换、一阶导数(窗口宽度 21)、二阶导数(窗口宽度 37)对光谱进行处理。

分段直接校正算法(piecewise direct standardization, PDS)是在直接校正算法基础上开发的用于模型传递的算法^[11]。采用 Kennard-Stone 算法从主机校正集光谱中,分别选择 5, 10, ..., 75, 80 个样本作为标准样进行比较。同时考察窗口宽度(w)为 3, 9, 15, 21, 31 对传递结果的影响。基于典型相关分析(canonical correlation analysis, CCA)的模型传递算法是依据 CCA 分析而提出的。借助 CCA 可以提取两组光谱之间的线性关系,实现光谱在主机与从机间的传递,

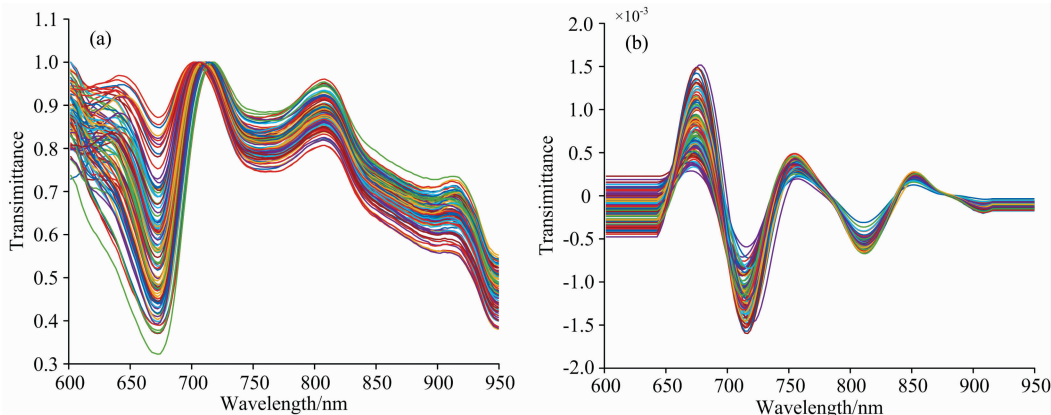


图 2 经过预处理后的苹果及大桃的透射率光谱

(a): 苹果; (b): 大桃

Fig. 2 The transmittance spectra after preprocessing methods

(a): Apple; (b): Peach

2.1.2 装置预测精度验证

直接利用主机装置对预测集的样本进行预测,采集前,首先获取白参考光谱。然后将预测集样本依次放置在探头上方,点击触发按键,即可在显示屏获取水果糖度预测值,读数后记录。同时,原始光谱数据通过串口传输到电脑端保存。待所有水果预测完成后,采用 1.4 节方法,获取预测集样本检测部位实际糖度值并记录。检测结果如图 3(a, b) 所示。

对于苹果糖度检测, R_p 和 RMSEP 分别为 0.92 和 0.587%。对比前期文献检测结果^[12],证明该装置可满足苹果糖度的快速、有效分析。而对大桃糖度预测,与前期近红外光谱大桃糖度检测结果相比略差,可能是由于在该研究中,大桃糖度主要集中在 11%~15% 这一较窄范围造成预测

并消除噪声以及干扰信息的影响^[12]。

本研究模型分析采用偏最小二乘算法,并以校正集样本 10 折交互验证均方根误差(RMSECV)最小值来确定最佳预处理方法和最佳建模主因子数^[13]。模型的评价采用校正集相关系数(R_c)和校正均方根误差(RMSEC),以及预测相关系数(R_p)和预测均方根误差(RMSEP)。

2 结果与讨论

2.1 实验室条件下水果糖度光谱分析

2.1.1 主机校正集光谱分析

对主机采集到的苹果、大桃校正集光谱数据,经过预处理算法后的 PLS 建模比较发现,对于苹果采用 5 点平滑结合最大值归一化的预处理方式结果较好[图 2(a)],交互验证相关系数和 RMSECV 分别为 0.885 和 0.628%,最佳主因子数为 9。对于大桃采用 37 点的二阶导数处理方式结果较好[图 2(b)],交互验证相关系数和 RMSECV 分别为 0.813 和 0.590%,最佳主因子数为 11。将模型系数及相应的预处理算法写入检测程序以对预测集样本进行直接预测。

相关系数偏低。另外,大桃品种的差异、所用光谱仪的波段范围也会对检测精度产生影响。

2.1.3 主机模型到从机传递结果

图 4 为 PDS 算法传递后,不同标准样本数量和不同窗口宽度对应的从机 R_p 值的变化。从图中可以看出,不同参数组合对 R_p 值的影响较大。对于苹果糖度模型传递结果[图 4(a)],当窗口宽度为 21,标准样本数量为 50 时, R_p 值最大为 0.891,此时对应的 RMSEP 为 0.657%。而对于大桃糖度模型的 PDS 传递结果[图 4(b)],当窗口宽度为 31,标准样本数量为 65 时,此时 R_p 最高为 0.778, RMSEP 为 0.725%。

图 5 为经过 CCA 算法传递后,从机对预测集样本的预测结果。对于苹果糖度模型的传递结果,当标准样本数量为 50 时,此时对应最优结果, R_p 和 RMSEP 分别为 0.883 和

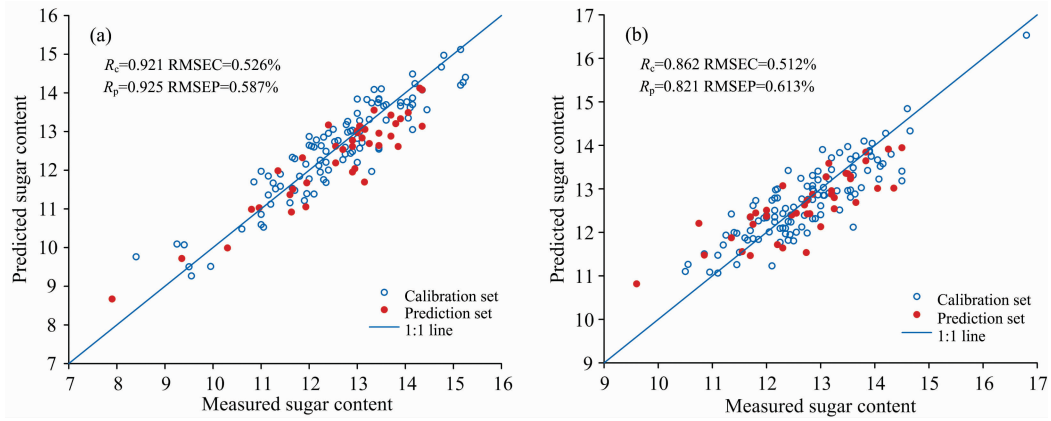


图 3 苹果和大桃糖度真实值和预测值散点图

(a): 苹果; (b) 大桃

Fig. 3 Scatter plots of predicted vs. measured sugar content

(a): Apple; (b): Peach

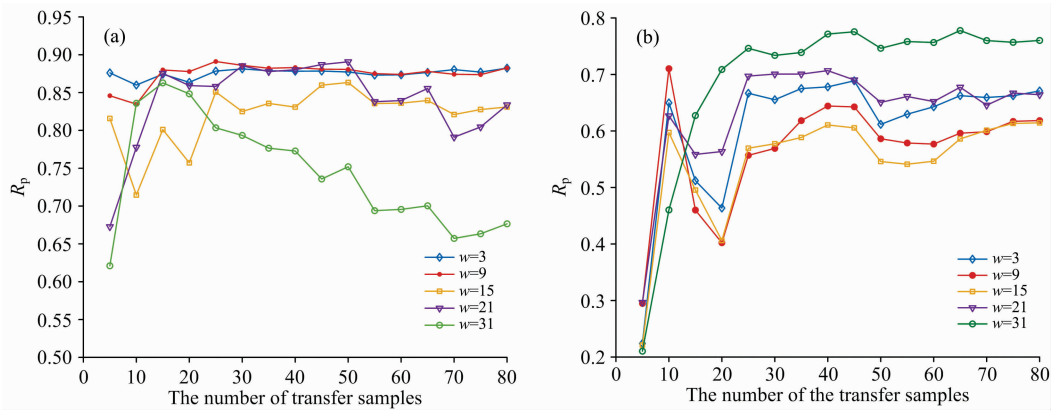


图 4 PDS 算法后从机 R_p 在不同窗口宽度和标准样本数量下的变化

(a): 苹果; (b) 大桃

Fig. 4 Variation of R_p obtained by piecewise direct standardization (PDS) algorithm with different numbers of transfer samples and different window sizes

(a): Apple; (b): Peach

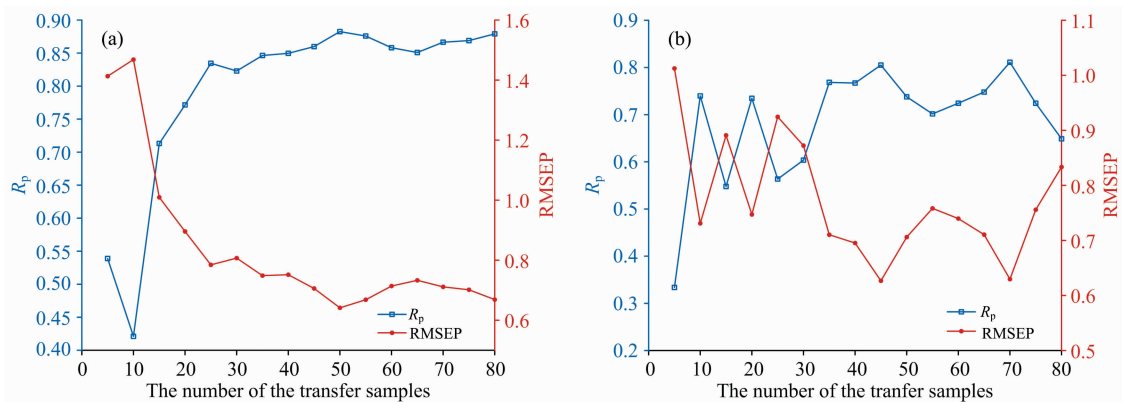


图 5 采用 CCA 算法后从机 R_p 和 RMSEP 随不同标准样本数量的变化

(a): 苹果; (b): 大桃

Fig. 5 Variation of R_p and RMSEP obtained by CCA algorithm with different numbers of transfer samples

(a): Apple; (b): Peach

0.641%，如图 5(a)所示，与 PDS 传递结果相差不大。而对于大桃糖度模型传递结果，当选取 45 个样本作为标准样本用于传递时取得结果最好，此时 R_p 和 RMSEP 分别为 0.805 和 0.626%，见图 5(b)。与 PDS 传递结果相比有了较大提升。经过模型传递算法，将主机的模型用于从机糖度预测，结果是可行的，虽然检测精度略有下降，但从后续设备的批量开发，减少建模工作量、提升开发效率还是具有积极意义的。

2.2 树上水果糖度光谱分析

将树上采集到的水果光谱数据代入上述实验室条件下建

立的模型。图 6(a,b)分别展示了使用实验室环境下构建的苹果、大桃糖度预测模型对树上 140 个水果糖度的检测结果。对大桃的检测精度与实验室条件检测结果基本一致。对于苹果糖度的预测，除去图中一个异常样本外， R_p 提高到 0.894，RMSEP 下降到 0.633%，接近实验室条件下预测精度。上述结果证明，利用开发的检测装置及配套的遮光罩，以及相应的光谱预处理方法，可以实现实验室环境下构建的模型对树上水果糖度的有效检测。

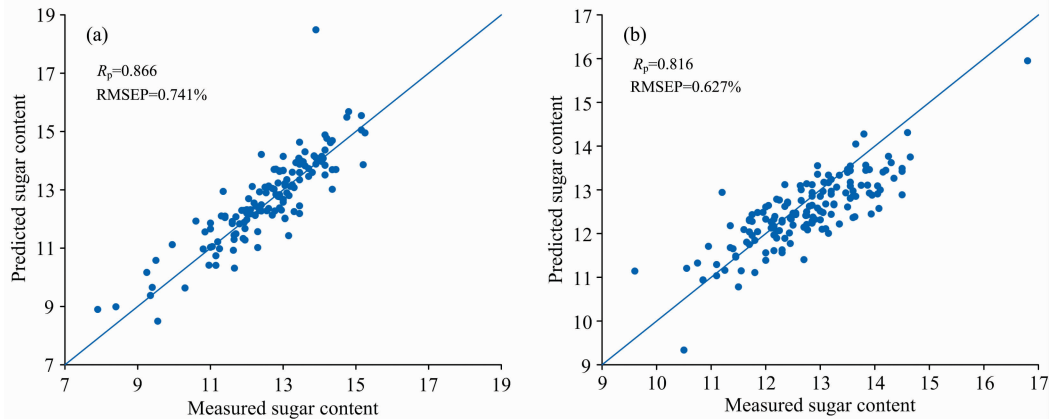


图 6 苹果和大桃糖度检测真实值和预测值散点图

(a): 苹果; (b): 大桃

Fig. 6 Scatter plots of predicted vs. measured sugar content

(a): Apple; (b): Peach

3 结论

基于微型光谱仪开发的手持式水果糖度便携式检测装置，可用于苹果、大桃糖度的无损、快速分析，在实验室环境下对苹果和大桃的预测集 R_p 和 RMSEP 分别为 0.925，0.587%和 0.821，0.613%。相对于分段直接校正算法，基于 CCA 模型传递方法，在该样机上取得了更好的模型传递结果。针对主机建立的用于苹果和大桃糖度预测的模型，传递

到从机后， R_p 和 RMSEP 分别为 0.883，0.641%和 0.805，0.626%，见图 5(b)。可以满足从机对苹果、大桃糖度的有效检测，实现了可见-近红外光谱糖度模型在该装置间进行共享和传递。配合设计的遮光罩，直接采用实验室条件下的光谱数据构建的模型预测树上水果糖度，对苹果、大桃糖度预测的 R_p 和 RMSEP 分别为 0.866，0.741%和 0.816，0.627%，可以实现实验室条件下建立的模型对田间树上水果糖度的检测。

References

- [1] LIU Yan-de, XU Hai, SUN Xu-dong, et al(刘燕德, 徐海, 孙旭东, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2020, 40(3): 922.
- [2] Yan H, Siesler H W. NIR News, 2018, 29 (7): 8.
- [3] Yang B, Guo W, Li W, et al. Journal of Food Process Engineering, 2019, 42(2): e12982.
- [4] Guo W, Li W, Yang B, et al. Journal of Food Engineering, 2019, 257: 1.
- [5] WANG Fan, LI Yong-yu, PENG Yan-kun, et al(王凡, 李永玉, 彭彦昆, 等). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2017, 33(19): 295.
- [6] GUO Zhi-ming, CHEN Quan-sheng, ZHANG Bin, et al(郭志明, 陈全胜, 张彬, 等). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2017, 33(8): 253.
- [7] Yu X, Lu Q, Gao H, et al. Journal of Near Infrared Spectroscopy, 2016, 24 (1): 69.
- [8] ZHAO Juan, QUAN Peng-kun, ZHANG Meng-sheng, et al(赵娟, 全朋坤, 张猛胜, 等). Transactions of The Chinese Society of Agricultural Machinery(农业机械学报), 2019, 50(4): 326.
- [9] Cavaco A M, Pires R, Antunes M D, et al. Postharvest Biology and Technology, 2018, 141: 86.

- [10] Fan S, Wang Q, Tian X, et al. *Biosystems Engineering*, 2020, 193: 138.
- [11] Workman J J. *Applied Spectroscopy*, 2018, 72(3): 340.
- [12] LI Xin, BIN Jun, FAN Wei, et al(李 鑫, 宾 俊, 范 伟, 等). *Chinese Journal of Analytical Chemistry(分析化学)*, 2017, 45(7): 958.
- [13] Donis-González I R, Valero C, Momin M A, et al. *Agronomy*, 2020, 10(1): 148.

Development and Experiment of a Handheld Visible/Near Infrared Device for Nondestructive Determination of Fruit Sugar Content

FAN Shu-xiang¹, WANG Qing-yan¹, YANG Yu-sen², LI Jiang-bo¹, ZHANG Chi¹, TIAN Xi¹, HUANG Wen-qian^{1*}

1. Beijing Research Center for Information Technology in Agriculture, Beijing 100097, China

2. Southwest Jiaotong University-University of Leeds Joint School, Chengdu 611731, China

Abstract A handheld portable device for fruit sugar content was developed based on visible/near-infrared spectral analysis. The device consists of a micro-spectrometer, halogen lamps, OLED screen and microcontroller. The real-time analysis and control software of the microcontroller was written in C language with the help of the Keil 5 development tool. Combined with the spectrum acquisition program written by LabView, the spectra of fruit samples were collected by the developed device. Apples and big peaches were used to explore the detection accuracy of the device and the transfer of the model between two devices (master and slave). The visible-near infrared spectra of the apple and peach were collected in the spectral range of 600~950 nm under laboratory conditions and in the field. The spectral data of calibration set collected by the master device under laboratory conditions were preprocessed by smoothing, maximum normalization, second derivative and other preprocessing methods, followed by the sugar content models developed using partial least squares algorithm for apples and peaches respectively. The models were then imported to the custom software, making it possible for the master device to predict the sugar content of apples or peaches directly. The correlation coefficient and the root mean square error of the prediction set were 0.925, 0.587% and 0.821, 0.613% for apples and peaches, respectively. The models were transferred from the master device to the slave device by using the piecewise direct standardization (PDS) and canonical Correlation Analysis (CCA) algorithm. After comparison, it was found that better model transfer results were achieved based on the CCA algorithm. The correlation coefficient and root mean square error of the prediction set were 0.883, 0.641% and 0.805, 0.626% for apples and peaches, respectively. The model established under laboratory conditions was used to analyze the fruit spectral data collected on the tree, the correlation coefficient and root mean square error of the prediction set were 0.866, 0.741% and 0.816, 0.627% for apples and peaches, respectively. The results showed that the developed device had considerable potential to detect fruit sugar content under lab conditions, and in the field. With the help of the model transfer algorithm, the model can be shared and effectively transferred between different devices. The developed device could meet the demand for rapid, non-destructive, and on-site detection of internal fruit quality.

Keywords Nondestructive detection; Fruit; Visible-near infrared spectrum; Spectral analysis; Sugar content; Model transfer

(Received Mar. 1, 2021; accepted Jul. 9, 2021)

* Corresponding author