

基于高光谱技术的柑橘不同部位糖度预测模型研究

Research on the detection model of sugar content in different position of citrus based on the hyperspectral technology

介邓飞 杨杰 彭雅欣 连裕翔 张登

JIE Deng-fei YANG Jie PENG Ya-xi LIAN Yu-xiang ZHANG Deng

(华中农业大学工学院,湖北武汉 430070)

(College of Engineering, Huazhong Agricultural University, Wuhan, Hubei 430070, China)

摘要:利用高光谱技术研究柑橘不同部位的糖度预测模型,将花萼、果梗和赤道部位的高光谱信息分别建立与其对应部位糖度的预测模型,建立基于偏最小二乘(Least squares regression, PLSR)、主成分回归(Principal component regression, PCR)和多元线性回归(Stepwise multivariate linear regression, SMLR)预测模型,3种预测模型中 PLSR 模型检测效果最好,通过 Norris derivative 预处理方法对花萼光谱数据进行处理后,预测集相关系数 $r_{pre} = 0.950$, 预测集均方根误差 $RMSEP = 0.636$ °Brix。结果表明,采用柑橘不同部位的高光谱信息与对应糖度预测模型是可行的,花萼部位所建立模型的效果优于果梗、赤道部位,因此花萼部位可作为优先选择的光谱检测部位,这对于指导实际检测分级生产中柑橘的摆放位置具有重要意义;采用 PLSR 方法建立柑橘花萼、果梗和赤道部位的高光谱信息与平均糖度的预测模型时,花萼部位模型效果最好,预测集相关系数 $r_{pre} = 0.913$, 预测集均方根误差 $RMSEP = 0.621$ °Brix, 建模效果相较于对应部位光谱与糖度模型差,因此,采用柑橘全部果肉的平均糖度与采集部位光谱建立糖度预测模型具有一定的局限性。

关键词:高光谱技术;柑橘;糖度预测模型;无损检测;

Abstract: Hyperspectral techniques were used to study the sugar content of different parts of citrus, and the sugar content detection models with hyperspectral information of calyx, stem and equator part were established respectively. The results showed that the model established by calyx was better than that of stem and equator. The de-

tection models of partial least squares regression (PLSR), principal component regression (PCR), and stepwise multivariate linear regression (SMLR) were established respectively, and the results of these three models were close. The PLSR model was found to be the best among them, after Norris derivative pretreatment methods were applied, the prediction correlation coefficient (r_{pre}) and the root mean square error of prediction (RMSEP) were 0.950 and 0.636 °Brix. This result inclined that it was feasible to use the hyperspectral technology to detect the sugar content in different parts. The study indicated that the calyx part could be the prior choice for the sugar content detection site in the citrus quality testing, and the conclusion has great significance for the way of citrus place in the actual production. Moreover, the PLSR method was used to establish the model of hyperspectral information and average sugar content in calyx, stem and equator part. The highest prediction r_{pre} and RMSEP of models was in the calyx and only to be 0.913 and 0.621 °Brix, which was not excellent enough. Therefore, it was limited to predict the citrus average sugar content with the hyperspectral information of a certain part.

Keywords: hyperspectral technology; citrus; soluble solids content; nondestructive testing

近年来,随着社会的发展和人们消费观念的不断提高,农产品的品质问题越来越为人们所重视。柑橘作为世界第一大水果,富含 V_C 、类胡萝卜素、类黄酮等功能性物质,为消费者所喜爱。柑橘果实中可溶性糖含量与其风味品质密切相关,是其重要的品质指标。在柑橘果实的发育过程中,柑橘内部糖度的分布受到其生长环境的影响,可能造成不同果实的可溶性糖积累情况差异较大^[1]。在实际生产加工的过程中,对柑橘进行检测与分级是商品化处理中一个重要的步骤,而其糖度是对柑橘进行分级的重要参考指标^[2]。

高光谱成像技术兼备图像与光谱两者的优点,在农产品品质 and 安全性评估方面具有巨大应用潜力,已广泛应用于葡萄^[3]、蓝莓^[4]、圣女果^[5]、黄瓜^[6]、枣^[7]、桃子^[8]等农产品品质

基金项目:现代农业(柑橘)产业技术体系建设专项资金项目(编号: CARS-27);中央高校基本科研业务费资助项目(编号: 2662015PY078);国家级大学生创新项目(编号: 201610504057)

作者简介:介邓飞(1982—),男,华中农业大学讲师,博士。
E-mail: dfjie@mail.hzau.edu.cn

收稿日期:2016—12—22

检测^[9-11]。在柑橘品质检测方面, Kim等^[12]利用高光谱成像技术建立了柑橘黑斑病检测模型, 检测精度可达97.9%。Qin等^[13]利用主成分分析(Principal Component Analysis, PCA)方法对柑橘溃疡的高光谱图像进行无损检测分析, 检测准确度为92.7%。而在柑橘糖度的内部品质无损检测方面, 许文丽等^[14]应用USB4000便携式近红外光谱仪获取不同放置方位下的赣南脐橙的可溶性固形物含量, 研究了柑橘的果梗和花萼部位不同夹角的入射光线对光谱图像的影响。

目前的研究中, 大多数是对整个柑橘果肉进行榨汁后测定其糖度, 而柑橘不同部位的光谱特征信息差异较大, 柑橘在光谱检测过程中, 由于光源入射角度的原因某些部位的信息可能不会被光谱仪所接收, 若在不考虑检测部位的情况下, 与光谱信息所建立的糖度检测模型将会对判断柑橘糖度品质产生一定的干扰, 影响其检测精度。基于高光谱技术对柑橘进行糖度无损检测时, 寻找柑橘最具代表性的检测部位, 探索检测部位糖度与光谱信息相关关系至关重要, 适宜的检测部位可以有效获得柑橘内部信息, 从而提高柑橘糖度在线检测精度。本研究拟采用高光谱技术采集柑橘花萼、果梗和赤道部位的透射光谱信息, 分别取对应部位的果肉进行榨汁测定其糖度值, 研究柑橘不同部位糖度与高光谱信息之间的相关关系, 建立糖度预测模型, 探索适宜在线检测与分级的柑橘摆放位置和糖度建模方法, 旨在为柑橘无损检测分级加工设备开发提供技术支持和研究基础。

1 材料与与方法

1.1 试验材料与仪器

新鲜柑橘样本: 采摘于湖北省武汉市江夏区某橘园, 所采摘柑橘大小接近, 外观完好。采摘当天运回实验室后, 进行检查和筛选, 选取试验样本70个, 并对试验样本依次进行编号, 记录其外观尺寸和重量, 在24℃(室温)下存放待测;

手持糖度计: WYT-4型, 福建泉州光学仪器厂;

电子天平: APTP452型, 中国深圳安普特公司;

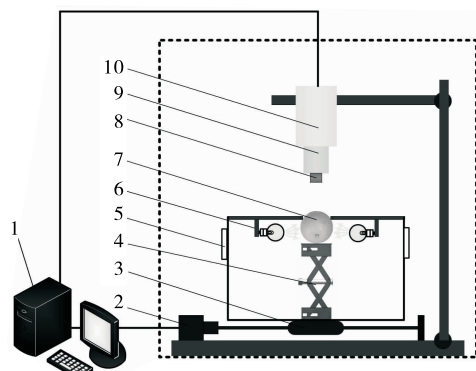
游标卡尺: MNT-150E型, 中国上海美耐特公司;

高光谱成像光谱仪: Specim V10E型, 芬兰SPECIM公司。

1.2 试验方法

1.2.1 光谱采集 柑橘试验样本的高光谱图像信息通过漫透射的方式获取。高光谱漫透射成像系统由光谱仪、镜头、光源、电动位移平台、风扇和电脑组成, 装置示意图见图1。高光谱成像仪是由CCD摄像头(Andor, Clara, Britain)、图像光谱仪(SPECIM, V10E-CL, Finland)和成像镜头组成, 样品室光源采用4盏50W卤素灯, 风扇用于样品室的散热。光谱仪的光谱范围是390~1055nm, 分辨率为2.8nm, 像素为1392×1040。

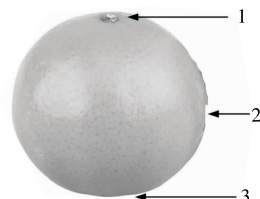
在对样本进行试验之前, 对试验装置的各项参数进行调试以获得较高信噪比高光谱图像, 调整完毕后, 进行信息采集试验。试验过程中采集柑橘花萼、果梗和赤道部位的光谱信息, 光谱采集部位示意图见图2, 曝光时间设置为100ms,



1. 电脑 2. 电动机 3. 电动位移台 4. 升降台 5. 风扇 6. 光源
7. 柑橘 8. 镜头 9. 光谱仪 10. 相机

图1 高光谱透射成像系统

Figure 1 The acquisition of the transmission spectra of citrus by hyperspectral imaging system



1. 果梗部位 2. 赤道部位 3. 花萼部位

图2 柑橘不同采集位置示意图

Figure 2 Different positions of citrus used to collect spectra

电动位移平台移动速度设置为2mm/s, 通过Spectral SECN-V17E软件记录柑橘高光谱漫透射图像。

1.2.2 理化指标的测定 高光谱图像采集完毕之后, 进行糖度理化指标测定。采用糖度计测定柑橘可溶性固形物, 将蒸馏水滴定在棱镜玻璃面上, 盖上盖板, 检查示数并进行调零操作, 将柑橘的花萼、果梗和赤道部位的果肉分别榨汁测定其可溶性固体物(Soluble Solids Content, SSC)含量, 并对柑橘样本均进行重复试验。

1.2.3 分析方法和软件 利用ENVI 4.6(美国Exelis Visual Information Solutions公司)对采集到的柑橘高光谱信息进行处理, 将样本与背景分离作为感兴趣区域(Region of Interest, ROI)以减少冗余, 对所有样品进行类似处理, 计算ROI的平均光谱, 作为定量分析的输入数据。采用TQ Analyst 8.0(美国Thermo公司)和MATLAB 2010a(美国MathWorks公司)软件进行数据处理。通过相关系数(r)、模型的校正均方根误差(Root mean square error of calibration, RMSEC)和模型的预测均方根误差(Root mean square error of prediction, RMSEP)对模型的稳定性和准确性进行评价, 模型的相关系数 r 越大越好(包括校正集相关系数 r_{cal} 和预测集相关系数 r_{pre}), RMSEC和RMSEP越小越好, 所建的模型效果越好^[15-16]。

2 结果与分析

2.1 光谱分析

柑橘高光谱图两端噪音较大, 去除首尾噪音部分, 选取

550~950 nm 范围内的光谱,见图 3,共计 313 个波段进行后续分析。由图 3 可知,柑橘高光谱图在 650,720,810 nm 附近有 3 个波峰,在 680 nm 和 760 nm 附近有两个波谷。

2.2 柑橘糖度品质指标测量结果

由表 1 可知,柑橘花萼部位的变异系数值大于果梗部位和赤道部位,说明花萼部位的可溶性固形物的分布离散性较大。随着生长时间的延长,柑橘样品中可溶性固形物的量不断增加,变化较大的品质指标更能够显示出柑橘糖度品质的变化,可以预见柑橘花萼部位的高光谱检测效果应优于果梗部位和赤道部位。

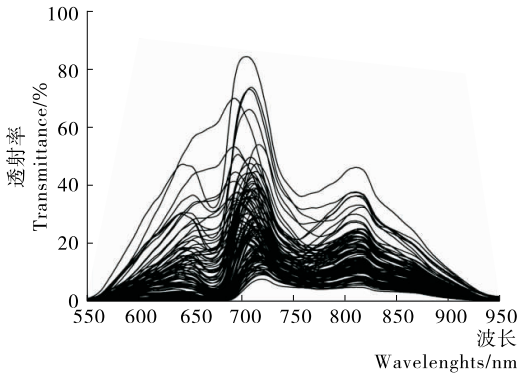


图 3 柑橘样本的高光谱图

Figure 3 The spectra of citrus obtained by hyperspectral imaging

表 1 柑橘可溶性固形物品质参数测量结果

Table 1 The measurement of the SSC quality parameter

检测部位	范围/°Brix	平均值/°Brix	标准方差/°Brix	变异系数/%
花萼	10.1~15.0	12.607	1.480	11.74
果梗	9.8~13.0	11.304	0.871	7.71
赤道	10.1~14.0	12.153	1.215	10.00

2.3 异常光谱剔除

TQ Analyst 软件中 Spectrum Outlier 功能可实现对异常光谱的剔除。通过计算每个光谱与样本光谱平均值之间的差值,将差值大于平均光谱标准差值的 0.95 置信区间的异常值剔除。使用 Leverage 功能实现对浓度异常样本的剔除,通过杠杆值和学生残差查验方式,将靠近上、下边缘残差或杠杆值较大的样本暂定为异常样本,之后对这些样本逐次回收计算校验。将所有样本均设为建模集,使用 PLSR 方法建模分析,根据模型的性能好坏确定是否为异常浓度样本。柑橘样本的花萼、果梗、赤道部位光谱分析时,需使用上述方法将异常样本剔除。

2.4 光谱预处理

在光谱数据中存在许多不同类型的随机噪声,诸如相机中的读出噪声和暗电流,其显著影响图像数据。在建立预测模型之前,对平均光谱数据应用预处理方法,以减少高频随机噪声、样本中的不均匀性和表面散射的影响。预处理方法包括多元散射校正(Multiplicative signal correction, MSC)、标准正态变量校正(Standard normal variable, SNV)、一阶导

数(First derivative)、二阶导数(Second derivative)、Savitzky-Golay 平滑和 Norris derivative 平滑等。本研究采用花萼部位建模效果的差异选取预处理方法,预处理结果见表 2。

表 2 经不同光谱预处理方法后对花萼部位建立 PLSR 模型结果

Table 2 The modeling effects by different spectral preprocessing methods application for SSC detection

预处理方法	校正集		预测集	
	r_{cal}	RMSEC/°Brix	r_{pre}	RMSEP/°Brix
Raw	0.932	0.705	0.834	0.985
MSC	0.923	0.744	0.897	0.802
SNV	0.926	0.730	0.887	0.842
First derivative	0.909	0.754	0.892	0.828
Second derivative	0.926	0.732	0.912	0.759
Savitzky-Golay	0.920	0.739	0.918	0.753
Norris derivative	0.971	0.461	0.950	0.636

对样本数据采用不同光谱预处理方法后,对花萼部位建立 PLSR 模型,通过比较不同预处理方法下模型效果来确定预处理方法。经对比,原始光谱模型效果最差,预测集相关系数 $r_{pre}=0.834$,RMSEP=0.985 °Brix,而 Norris derivative 平滑处理后效果最优,预测集相关系数 $r_{pre}=0.950$,RMSEP=0.636 °Brix。综上所述,对于该试验的预处理方法,Norris derivative 平滑时建模效果最好,因此,后续的模型分析都是基于 Norris derivative 预处理方法进行。

2.5 预测模型的建立

2.5.1 不同预测部位糖度的光谱预测模型比较 分别应用 PLSR、PCR、SMLR 方法建立了柑橘花萼部位,赤道部位和果梗部位糖度的预测模型,通过模型预测值与真实值的结果比较,得到建模效果见表 3。

由表 3 可知,对 PLSR、PCR 和 SMLR 3 种建模方法而言,PLSR 方法所建模型的效果最好。在此建模方法下,柑橘

表 3 不同建模方法可溶性固形物指标的建模效果

Table 3 The results of different models for SSC determination

处理 方法	检测 部位	校正集		预测集	
		r_{cal}	RMSEC/°Brix	r_{pre}	RMSEP/°Brix
PLSR	花萼	0.971	0.461	0.950	0.636
	果梗	0.830	0.611	0.753	0.654
	赤道	0.935	0.574	0.883	0.660
PCR	花萼	0.935	0.673	0.869	0.964
	果梗	0.846	0.590	0.731	0.737
	赤道	0.947	0.520	0.867	0.761
SMLR	花萼	0.937	0.678	0.929	0.692
	果梗	0.906	0.482	0.811	0.585
	赤道	0.910	0.663	0.905	0.592

的花萼部位所建立高光谱信息与糖度模型的预测集相关系数最优($r_{pre} = 0.950$, $RMSEP = 0.636$ °Brix), 赤道部位所建模型的预测集相关性其次($r_{pre} = 0.883$, $RMSEP = 0.660$ °Brix), 果梗部位所建模型预测集相关性相对最差($r_{pre} = 0.753$, $RMSEP = 0.654$ °Brix)。花萼部位的高光谱信息与糖度建模效果最优, 可将其作为预测柑橘糖度的检测部位, 分析结果对指导柑橘无损检测分级加工过程中的摆放位置具有现实意义。

2.5.2 平均糖度与光谱特性信息检测模型比较 为了探索花萼、果梗和赤道部位光谱与柑橘全部果肉平均糖度的建模关系, 将花萼、果梗和赤道部位所测糖度取平均值, 作为柑橘平均糖度值, 并将各检测位置光谱信息与平均糖度建立预测模型。

根据 2.5.1 的结论, PLSR 方法下所建立模型效果较好, 利用经 Norris derivative 预处理后柑橘各位置高光谱对柑橘平均糖度建立 PLSR 预测模型。

由表 4 可知, 花萼、果梗和赤道部位高光谱信息和平均糖度关系较对应部位糖度模型建模效果差一些。其中花萼部位的相关性($r_{pre} = 0.913$, $RMSEP = 0.621$ °Brix)略高于果梗和赤道部位, 但不足以作为预测柑橘平均糖度的直接条件。研究表明, 采用柑橘整体果肉榨汁取平均糖度对应高光谱信息进行建模预测的方法缺乏合理性。

表 4 PLSR 建模方法平均糖度指标的建模效果

Table 4 The results of PLSR models for SSC determination

检测部位	校正集		预测集	
	r_{cal}	RMSEC/°Brix	r_{pre}	RMSEP/°Brix
花萼	0.930	0.529	0.913	0.621
果梗	0.947	0.485	0.868	0.647
赤道	0.917	0.568	0.892	0.604

3 结论

本试验采用高光谱技术采集柑橘花萼、果梗和赤道部位的高光谱图像, 将花萼、果梗和赤道部位的果肉分别榨汁测定其 SSC 值, 分别建立对应部位光谱的糖度预测模型。柑橘样本漫透射光谱经 Norris derivative 预处理后建模效果较好, 基于 PLSR、PCR、SMLR 3 种建模方法中, PLSR 预测结果优于 PCR 和 SMLR 预测结果。柑橘不同部位所建立的 PLSR 糖度预测模型性能差异较明显, 花萼部位的预测性能优于果梗和赤道部位, 其预测集相关系数 $r_{pre} = 0.950$, 预测集均方根误差 $RMSEP = 0.636$ °Brix。试验结果表明, 在基于高光谱技术的柑橘内部品质检测中, 花萼部位对应光谱糖度模型的预测效果最好, 花萼部位可作为优先选择的光谱检测部位, 这对指导实际检测分级加工生产中柑橘的摆放方式具有重要意义。研究同时表明, 利用 PLSR 建立花萼、果梗和赤道部位高光谱信息与平均糖度的预测模型, 模型相关性相对较差。因此, 对柑橘果肉整体榨汁取平均糖度值与某一检测部位的高光谱信息建立预测模型的方法具有一定的局限性。

参考文献

- [1] SONG S Y, LEE Y K, KIM I J. Sugar and acid content of Citrus prediction modeling using FT-IR fingerprinting in combination with multivariate statistical analysis[J]. Food Chemistry, 2015, 190: 1 027-1 032.
- [2] ALFATNIA M S M, SHARIFF A R M, ABDULLAH M Z, et al. The application of internal grading system technologies for agricultural products-Review[J]. Journal of Food Engineering, 2013, 116(3): 703-725.
- [3] CHEN Shan-shan, ZHANG Fang-fang, NING Ji-feng, et al. Predicting the anthocyanin content of wine grapes by NIR hyperspectral imaging[J]. Food Chemistry, 2015, 172: 788-793.
- [4] YANG Ce, LEE W S, GADER P. Hyperspectral band selection for detecting different blueberry fruit maturity stages[J]. Computers and Electronics in Agriculture, 2014, 109(109): 23-31.
- [5] CHO B K, KIM M S, BAEK I S, et al. Detection of cuticle defects on cherry tomatoes using hyperspectral fluorescence imagery[J]. Postharvest Biology and Technology, 2013, 76: 40-49.
- [6] CEN Hai-yan, LU Ren-fu, ZHU Qi-bing, et al. Nondestructive detection of chilling injury in cucumber fruit using hyperspectral imaging with feature selection and supervised classification[J]. Postharvest Biology and Technology, 2016, 111: 352-361.
- [7] 徐爽, 何建国, 易东, 等. 基于高光谱图像技术的长枣糖度无损检测[J]. 食品与机械, 2012, 28(6): 168-170.
- [8] SUN Ye, GU Xin-zhe, SUN Ke, et al. Hyperspectral reflectance imaging combined with chemometrics and successive projections algorithm for chilling injury classification in peaches[J]. LWT-Food Science and Technology, 2016, 75: 557-564.
- [9] 刘燕德, 张光伟. 高光谱成像技术在农产品检测中的应用[J]. 食品与机械, 2012, 28(5): 223-226, 242.
- [10] LIU Dan, ZENG Xin-an, SUN Da-wen. Recent developments and applications of hyperspectral imaging for quality evaluation of agricultural products: a review[J]. Critical Reviews in Food Science and Nutrition, 2015, 55(12): 1 744-1 757.
- [11] WU Di, SUN Da-wen. Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review-Part II: Applications[J]. Innovative Food Science and Emerging Technologies, 2013, 19: 15-28.
- [12] KIM D G, BURKS T F, RITENOUR M A, et al. Citrus black spot detection using hyperspectral image analysis[J]. Agricultural Engineering International: CIGR Journal, 2013, 15(3): 171-180.
- [13] QIN Jian-wei, BURKS T F, KIM M S, et al. Citrus canker detection using hyperspectral reflectance imaging and PCA-based image classification method[J]. Sensing and Instrumentation for Food Quality and Safety, 2008, 2(3): 168-177.
- [14] 许文丽, 孙通, 吴文强, 等. 脐橙放置方位对近红外光谱检测结果的影响[J]. 光谱学与光谱分析, 2012, 32(11): 3 002-3 005.
- [15] 朱向荣, 李高阳, 苏东林, 等. 基于近红外光谱与组合间隔偏最小二乘法的稻米镉含量快速检测[J]. 食品与机械, 2015, 31(4): 43-46, 50.
- [16] 丁姣, 蔡建荣, 张海东, 等. 近红外结合 Si-ELM 检测食醋品质指标[J]. 食品与机械, 2012, 28(1): 93-96.