

# 光谱分析在西甜瓜内部品质无损检测中的研究进展

马本学<sup>1,2\*</sup>, 喻国威<sup>1,2</sup>, 王文霞<sup>1,2</sup>, 罗秀芝<sup>1,2</sup>, 李玉洁<sup>1,2</sup>, 李小占<sup>1,2</sup>, 雷声渊<sup>1,2</sup>

1. 石河子大学机械电气工程学院, 新疆 石河子 832003  
2. 农业部西北农业装备重点实验室, 新疆 石河子 832003

**摘要** 西甜瓜(西瓜和甜瓜)味甘多汁, 营养丰富, 其内部品质的检测对其生产和流通具有重要意义。西甜瓜内部品质的传统检测方法检测时间长, 成本高且为有损检测, 不能满足现代生产的需要。随着光谱分析技术的快速发展, 应用近红外光谱分析和高光谱成像进行西甜瓜内部品质的无损检测已成为研究热点。为跟踪国内外最新研究进展并分析研究现状, 介绍了近红外光谱分析和高光谱成像的技术特点和系统组成, 归纳了光谱信息预处理、变量筛选、模型建立和模型评价等光谱信息解析方法, 综述了近红外光谱分析和高光谱成像在西甜瓜内部品质(可溶性固形物含量、坚实度、总酸含量、成熟度、水分等)无损检测中的应用, 并从技术难点和实际应用两方面讨论了光谱分析技术在西甜瓜内部品质无损检测中的发展趋势, 指出利用深度学习进行光谱信息解析、建立多特征信息融合的综合评价模型、开发基于人工智能与移动端深度融合的快速无损检测系统等将成为新的研究方向。

**关键词** 西甜瓜; 内部品质; 近红外光谱分析技术; 高光谱成像技术; 无损检测; 综述

**中图分类号**: O657.3 **文献标识码**: R **DOI**: 10.3964/j.issn.1000-0593(2020)07-2035-07

## 引言

西甜瓜(西瓜和甜瓜)<sup>[1]</sup>味甘多汁, 营养丰富, 深受广大消费者的青睐。近年来虽然西甜瓜的产量和销量大幅增加, 但是其内部品质参差不齐, 难以辨别, 主要原因在于传统检测方法检测时间长, 成本高且为有损检测, 无法满足现代生产的需要。随着光谱分析技术的快速发展, 近红外光谱分析(near-infrared spectroscopy, NIRS)和高光谱成像(hyperspectral imaging, HSI)<sup>[2-3]</sup>以快速、绿色、无损等优势广泛应用于西甜瓜内部品质无损检测中。西甜瓜内部品质的评价指标主要包括可溶性固形物含量(soluble solids content, SSC)、总可溶性固形物(total soluble solids, TSS)、坚实度(firmness, FM)、水分(moisture)、总酸含量(total acid content, TAC)、成熟度(maturity)等食用指标和类胡萝卜素(carotenoids)等营养指标<sup>[4]</sup>。

为跟踪国内外最新研究进展, 综述了近红外光谱分析和高光谱成像在西甜瓜内部品质无损检测中的应用; 同时, 简要介绍了近红外光谱分析和高光谱成像的技术特点和系统组成, 并归纳了常用的光谱信息解析方法; 最后, 从技术难点和实际应用两方面综合分析了光谱分析技术在西甜瓜内部品

质无损检测中的发展趋势。

## 1 光谱分析技术简介

### 1.1 近红外光谱分析技术

近红外光(near-infrared, NIR)是人们最早发现的非可见光区域, 它的波长范围为780~2 526 nm, 波数范围为12 500~4 000  $\text{cm}^{-1}$ 。近红外光谱属于分子振动光谱, 利用分子含氢基团(C—H, O—H 和 N—H 等)的倍频与合频的吸收特性来获取复杂的光谱信息。典型的近红外光谱分析系统一般由光源、光谱仪、载物台、光纤、计算机等部分组成, 如图1所示。

### 1.2 高光谱成像技术

高光谱成像技术兴起于20世纪80年代末, 结合了光谱学和传统成像或计算机视觉技术, 可以有效地获取样品“图谱合一”的三维数据信息, 包括二维的样品空间图像信息和一维的光谱信息<sup>[5-6]</sup>。典型的高光谱成像系统一般由成像光谱仪、CCD相机、光源、计算机、传输机构等部分组成, 如图2所示<sup>[7]</sup>。

近红外光谱分析技术和高光谱成像技术的区别与联系<sup>[8-10]</sup>, 如表1所示。

收稿日期: 2019-06-10, 修订日期: 2019-10-21

基金项目: 国家自然科学基金项目(61263041)和国家科技支撑项目(2015BAD19B03)资助

作者简介: 马本学, 1970年生, 石河子大学机械电气工程学院教授

e-mail: mbx\_shz@163.com

\* 通讯联系人

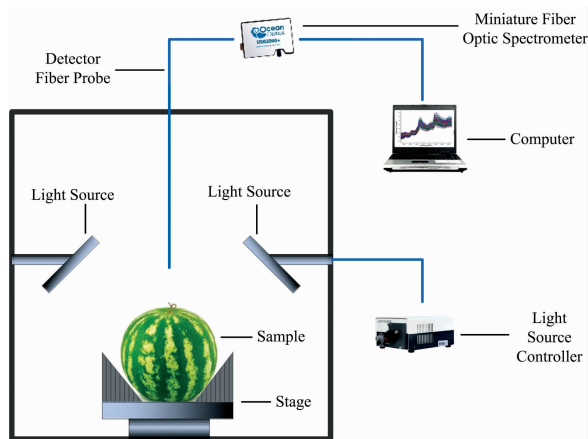


图 1 近红外光谱分析系统示意图  
Fig. 1 Schematic diagram of NIRS system

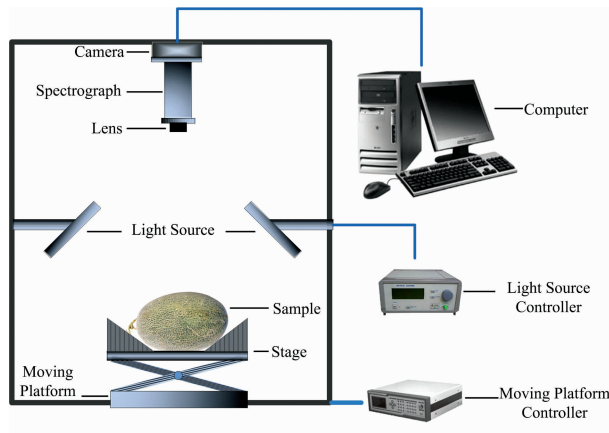


图 2 高光谱成像系统示意图  
Fig. 2 Schematic diagram of HSI system

表 1 近红外光谱分析和高光谱成像的区别和联系  
Table 1 Summary of differences and connections for NIRS and HSI

技术	区别			联系	
	仪器设备	信息获取	应用方向	信息处理	信息检测
NIRS	成本较低 设备较小, 便于携带	只能获取光谱信息 数据采集速度快	只能实现样品化学参数的评估 适用于在线检测	需依赖机器学习、化学计量学等方法建立数学模型进行定量、定性分析 模型鲁棒性和适应性差 数据有效信息挖掘难度大	可实现样品多组分信息的同时检测
HSI	成本较高 设备较大, 适用于室内 需要较高性能的计算机硬件	能够同时获取图像和光谱信息 数据量大且冗余信息较多 数据采集速度慢	可实现样品物理和化学参数的评估并将其可视化 难以实现商业化在线实时检测		

## 2 光谱信息解析方法

### 2.1 光谱信息解析流程

在光谱信息解析过程中, 因光谱系统采集的光谱信息会受到采集环境、光学条件和仪器性能等因素的影响, 故原始

光谱信息除含有被测样品属性的信息外, 还包含大量无关信息, 如噪声、背景干扰、杂散光等。为消除光谱冗余信息, 提高模型的预测精度, 对光谱信息进行预处理并提取有效信息是模型建立前的重要环节。一般的光谱信息解析步骤主要包括光谱信息的采集、光谱信息的预处理、变量的筛选、预测模型的建立、模型的评价等, 其流程图如图 3 所示<sup>[11]</sup>。

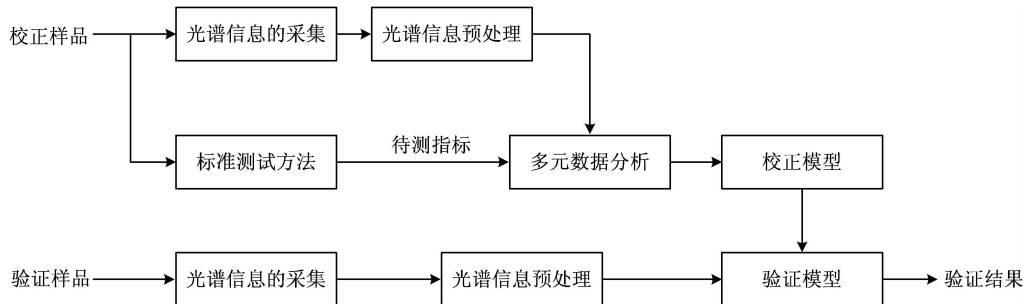


图 3 光谱信息解析流程图  
Fig. 3 Flow chart of spectral information analysis

### 2.2 光谱信息预处理

光谱信息解析首要环节就是对采集的光谱信息进行预处理, 常用光谱信息预处理方法<sup>[12]</sup>主要有平滑(smoothing), 包括移动平均法(moving average, MA)和卷积法(savitzky-golay, SG)等、导数修正(derivative correction), 包括一阶导

数(first-order derivative, 1<sup>st</sup>D)和二阶导数(second-order derivative, 2<sup>nd</sup>D)、归一化(normalization, NOR)、标准正态变换(standard normal variable transformation, SNV)、多元散射校正(multiplicative scattering correction, MSC)、小波变换(wavelet transform, WT)等, 其作用效果如表 2 所示。

表 2 光谱预处理方法及其作用效果

预处理方法	作用效果
Smoothing	消除噪声, 提高信噪比
Derivative Correction	消除基线的平移和漂移、分离重叠峰, 提高光谱灵敏度
NOR	消除光程变化产生的影响
SNV	补偿由颗粒大小和表面散射引起的偏差
MSC	补偿颗粒分布不均匀和颗粒大小引起的散射影响
WT	平滑滤波、基线校正、信息压缩

### 2.3 光谱信息变量筛选

光谱信息中存在的大量冗余信息会降低预测模型的准确性和稳定性。为提高模型运算速度和精度, 进行光谱信息变量筛选是十分必要的。常用变量筛选方法<sup>[13]</sup>主要有变量区间选择算法, 包括移动窗口偏最小二乘法(moving windows partial least squares, MWPLS)和区间偏最小二乘法(interval partial least squares, iPLS)等、无信息变量消除算法(uninformative variable elimination, UVE)、遗传算法(genetic algorithm, GA)、连续投影算法(successive projections algorithm, SPA)和竞争性自适应重加权算法(competitive adaptive reweighted sampling, CARS)等, 其特点如表 3 所示。

表 3 变量筛选方法的特点

变量筛选方法	特点
变量区间选择算法	筛选速度较快, 与单一变量筛选方法相比稳定性较高
UVE	通过定义阈值来筛选有效变量, 筛选变量的速度较快, 稳定性较高
GA	可以较好地保留变量间的组合优势, 适用于全局优化问题
SPA	具有较少的光谱冗余信息和共线性变量, 适用于高共线性光谱信息的变量筛选
CARS	具有平滑降噪、基线校正、压缩信息、计算速度快等特点

### 2.4 光谱信息模型建立

建立样品待测指标的数学模型是光谱信息解析过程中关键的一步, 不同的建模方法会直接影响模型的准确性和稳定性。常用的建模方法<sup>[14]</sup>主要有多元线性回归(multi linear regression, MLR)、主成分回归(principal component regression, PCR)、偏最小二乘回归(partial least squares regression, PLSR)、支持向量机(support vector machine, SVM)、最小二乘支持向量机(least square support vector machine, LS-SVM)等, 其特点如表 4 所示。

### 2.5 光谱信息模型评价

在样品待测指标的预测模型建立后, 需对模型进行评价, 常见的模型预测性能评价标准<sup>[15]</sup>主要有预测集相关系数(correlation coefficient of prediction set,  $R_p$ )、校正集相关系数(correlation coefficient of calibration set,  $R_c$ )、决定系数(coefficient of determination,  $R^2$ )、预测标准偏差(root mean

square error of prediction, RMSEP)、校正标准偏差(root mean square error of calibration, RMSEC)、剩余预测偏差(residual predictive deviation, RPD)等。质量较高的模型具有较高的  $R_p$ ,  $R_c$ ,  $R^2$  和 RPD, 较低且较为接近的 RMSEP 和 RMSEC<sup>[16]</sup>。

表 4 建模方法的特点

建模方法	特点
MLR	计算简单, 适用于变量较少的回归分析
PCR	降低光谱信息维度以排除原始光谱中相互重叠的信息, 在一定程度上滤除了噪声, 较好地解决了多元线性回归中的多重共线性问题
PLSR	在建模过程中结合了多元线性回归分析、主成分分析和典型相关分析等方法的特点, 不仅可以简化信息结构, 提取关键信息, 建立鲁棒性较好的定量分析模型, 同时可以更好地解决变量的多重相关性问题
SVM	具有优秀的泛化能力, 建立的模型鲁棒性较好, 能较好地解决少量样本的分类问题
LS-SVM	具有较强的非线性处理能力和泛化能力, 能进行快速地多元建模

通过对常用光谱信息解析方法的总结可以看出, 机器学习算法已逐渐应用于光谱信息的变量筛选和模型建立, 虽然提高了模型精度, 但是模型的普适性和实用性仍不能满足实际需要。现阶段, 深度学习作为机器学习的分支迅速崛起, 深度网络的规模和精度也在不断提高, 并且持续成功地应用于各类实际问题。为进一步提高模型的精度和泛化能力, 使其能够满足现代生产的需要, 应用深度学习算法进行光谱信息解析将是今后发展的必然趋势之一。

## 3 光谱分析在西甜瓜内部品质无损检测中的应用

### 3.1 近红外光谱分析的应用

表 5 列举了近红外光谱分析在西甜瓜内部品质无损检测中的应用。

SSC 含量作为西甜瓜内部品质的重要评价指标, 其定量分析一直是研究热点。介邓飞等<sup>[18]</sup>以“麒麟”西瓜为样品, 研究了不同变量筛选方法(等间隔平均光谱法、等间隔抽取光谱法、SPA)和不同建模方法(PLSR, MLR 和 PCR)对 SSC 含量预测模型精度的影响。结果表明, 采用等间隔抽取光谱法结合 SPA 建立的 PLSR 模型精度较高( $R_p$  为 0.828, RMSEC 为 0.589, RMSEP 为 0.611)。西甜瓜的 SSC 含量在不同部位存在明显差异, 检测部位差异是影响西甜瓜 SSC 含量预测模型精度的重要因素。为探究西甜瓜 SSC 含量的最佳检测部位, 介邓飞等<sup>[19]</sup>分别利用瓜梗、赤道和瓜脐部位的近红外漫透射光谱信息结合 PLSR 和 LS-SVM 建立了西瓜单一检测部位的 SSC 含量预测模型。结果表明, 瓜脐为最佳单一检测部位(LS-SVM 模型:  $R_p$  为 0.768, RMSEP 为 0.731、

PLSR 模型:  $R_p$  为 0.823, RMSEP 为 0.652)。Zhang 等<sup>[26]</sup>以哈密瓜为研究对象, 选用不同变量筛选方法(CARS, UVE, CARS-SPA 和 UVE-SPA), 对比分析了线性 PLS 和非线性 LS-SVM 建模方法对基于单一检测部位的哈密瓜 SSC 含量预测模型的影响。结果表明, 无论是线性模型还是非线性模型, 基于赤道部位光谱信息建立的哈密瓜 SSC 含量预测模型性能最佳(UVE-SPA-PLS 模型:  $R_p$  为 0.914 3, RMSEP 为 0.835 9、CARS-SPA-LSSVM 模型:  $R_p$  为 0.913 4, RMSEP 为 0.895 8)。由此可见, 不同品种西甜瓜的最佳检测部位不

同。为减小检测部位差异对西甜瓜 SSC 含量预测模型性能的影响, 钱曼等<sup>[22]</sup>利用“京秀”西瓜三个检测部位的近红外漫反射光谱信息结合 CARS 和 PLS, 建立了西瓜 SSC 含量的混合(赤道-瓜脐-瓜梗)预测模型( $R_p$  为 0.905, RMSEP 为 0.629), 其性能优于单一(瓜脐)预测模型。上述研究表明, 对于不同品种的西甜瓜选择适宜的近红外光谱检测部位并结合变量筛选方法, 可以有效提高西甜瓜 SSC 含量预测模型的性能。

表 5 近红外光谱分析在西甜瓜内部品质无损检测中的应用

Table 5 Applications of NIRS for non-destructive detection of internal quality in watermelon and muskmelon

品种	检测指标	最优预测模型	检测部位	模型性能评价	文献
西瓜	SSC/(°Brix)	MC-UVE-GA-PLS	瓜脐	$R_p=0.845$ , RMSEP=0.574	[17]
		SPA-PLSR		$R_p=0.828$ , RMSEP=0.611	[18]
		PLSR		$R_p=0.823$ , RMSEP=0.652	[19]
		MC-UVE-SMLR		$R_p=0.70$ , RMSEP=0.33	[20]
	Lycopene/(mg · 100 g <sup>-1</sup> FW)	PLS	混合	$R_p=0.751$ , RMSEP=0.439, RPD=1.13	[21]
	SSC/(°Brix)			$R_p=0.862$ , RMSEP=0.717, RPD=1.83	
	Moisture/%	CARS-PLS	混合	$R_p=0.939$ , RMSEP=0.276, RPD=2.79	[22]
	SSC/(°Brix)			$R_p=0.905$ , RMSEP=0.629	
	Lycopene/(mg · kg <sup>-1</sup> )			$R^2=0.877$ , RMSECV=15.68	
	$\beta$ -Carotene/(mg · kg <sup>-1</sup> )			$R^2=0.822$ , RMSECV=0.81	
TSS/%	PLS	整瓜	$R^2=0.836$ , RMSECV=0.8	[23]	
Maturity/%			Calibration set: CCR=90.2%		
甜瓜	SSC/(°Brix)	C-RPP	瓜脐	Prediction set: CCR=88.1%	[24]
	FM/(kgf · cm <sup>-2</sup> )			$R^2=0.938$ , RMSEP=0.83, RPD $\approx$ 3.0	
哈密瓜	SSC/(°Brix)	GA-PLS	赤道	$R^2=0.741$ , RMSEP=0.35	[25]
		UVE-SPA-PLS		$R_p=0.9143$ , RMSEP=0.835 9	
		CARS-SPA-LSSVM		$R_p=0.9134$ , RMSEP=0.895 8	[26]

Note: “MC” is Monte Carlo; “FW” is fresh weight; “RMSECV” is root mean square error of cross validation; “SMLR” is stepwise multiple linear regression; “C-RPP” is corrected-ratio of peaks method; “CCR” is correct classification rate

研究开发西甜瓜内部品质快速无损检测设备对西甜瓜的质量管理和市场竞争起着至关重要的作用, 国内外应用近红外光谱分析已构建多种西甜瓜内部品质快速无损在线检测系统。Jie 等<sup>[20]</sup>采集 697~920 nm 范围内西瓜瓜脐部位的近红外漫透射光谱信息, 利用基线偏移校正(baseline offset correction, BOC)进行光谱预处理, 建立了 MC-UVE-SMLR 预测模型( $R_p$  为 0.70, RMSEP 为 0.33), 构建了西瓜 SSC 含量在线检测系统。Tamburini 等<sup>[23]</sup>采集 900~1 700 nm 范围内的完整西瓜的近红外漫反射光谱信息, 利用导数修正和 SNV 进行光谱预处理, 建立了 PLS 模型, 构建了西瓜内部品质在线检测系统, 可实现对番茄红素( $R^2$  为 0.877, RMSECV 为 15.68),  $\beta$  胡萝卜素( $R^2$  为 0.822, RMSECV 为 0.81)和 TSS( $R^2$  为 0.836, RMSECV 为 0.8)的检测。少数公司研发的西甜瓜内部品质在线检测设备已达到工业化标准, 如日本三井金属矿业株式会社的西瓜糖度无损检测生产线(效率为 120 pcs · min<sup>-1</sup>, 误差为  $\pm 0.5^\circ$  Brix), 意大利萨克米公司研发的西甜瓜内部品质(糖度、酸度、成熟度和空心度)在线检测系统(效率为 300 pcs · min<sup>-1</sup>)等<sup>[27]</sup>。上述系

统及设备均为大型装置, 不便于小型商户和消费者使用, 开发基于智能移动终端的快速检测系统是未来发展方向之一。

### 3.2 高光谱成像技术的应用

表 6 列举了高光谱成像技术在西甜瓜内部品质无损检测中的应用。

近些年来利用高光谱成像技术对西甜瓜 SSC、FM 的定量分析较多, 李锋霞等<sup>[28]</sup>以哈密瓜为对象, 选取 500~820 nm 波段范围内的高光谱图像信息, 分别研究了不同光谱预处理方法(1<sup>st</sup>D, 2<sup>nd</sup>D, MSC 和 SNV)和不同建模方法(PLS, SMLR 和 PCR)对哈密瓜 SSC 和 FM 预测模型的影响。结果显示, 1<sup>st</sup>D 和 SNV 结合 PLS 建立的哈密瓜 FM 预测模型的预测效果最佳( $R_c$  为 0.873,  $R_p$  为 0.646, RMSEC 为 4.18, RMSEP 为 6.4)。高光谱成像所获取的光谱信息冗余度较大, 与被测组分无关的信息会削弱模型的预测精度。因此, 选择包含相对较多有效信息的变量来建立预测模型是十分必要的。Sun 等<sup>[7]</sup>分别比较了基于不同变量筛选方法(SPA, CARS 和 GA)结合 PLS 建立的哈密瓜 SSC, FM 和 TAC 预测模型的精度。结果表明, CARS-PLS 模型实现了对哈密瓜

SSC( $R_p$  为 0.960 6, RMSEP 为 0.381 6, RDP 为 3.598)、FM( $R_p$  为 0.867 1, RMSEP 为 20.05, RPD 为 1.996)和 TAC( $R_p$  为 0.912 5, RMSEP 为 0.026 3, RPD 为 2.445)的准确预测。Sun 等<sup>[30]</sup>利用 PLSR, PCA, SVM 和人工神经网络(artificial neural network, ANN)分别建立了甜瓜 SSC 和 FM 预测模型,其中 PLSR 模型的预测性能最佳。为进一步

提高模型精度,采用加权回归系数法筛选变量后结合 PLSR 重新建立了甜瓜 SSC 和 FM 预测模型。结果表明,经变量筛选后建立的甜瓜 SSC 模型( $R^2$  为 0.775 5, RMSEP 为 1.187 1)和 FM 模型( $R^2$  为 0.355 5, RMSEP 为 525.293 2)的预测精度均得到了提高。

表 6 高光谱成像在西甜瓜内部品质无损检测中的应用

Table 6 Applications of HSI for non-destructive detection of internal quality in watermelon and muskmelon

品种	检测指标	最优预测模型	检测部位	模型性能评价	文献
哈密瓜	FM/N	PLS		$R_c=0.873$ , $R_p=0.646$ RMSEC=4.18, RMSEP=6.40	[28]
	SSC/(°Brix)	MSC-CARS-SPA-SVM		$R_p=0.9404$ , RMSEP=0.4027, RPD=2.941	[29]
	FM/(g·cm <sup>-2</sup> )			$R_p=0.8253$ , RMSEP=35.22, RPD=1.771	
	Maturity/%	CARS-PCA-SVM	赤道	Calibration set: Accuracy=95% Prediction set: Accuracy=94%	
	FM/(g·cm <sup>-2</sup> )	CARS-PLS		$R_p=0.8671$ , RMSEP=20.05, RPD=1.996	
	SSC/(°Brix)	SNV-CARS-PLS		$R_p=0.9606$ , RMSEP=0.3816, RPD=3.598	[7]
TAC/(g·kg <sup>-1</sup> )			$R_p=0.9125$ , RMSEP=0.0263, RPD=2.445		
甜瓜	SSC/(°Brix)	PLSR	混合	$R^2=0.7755$ , RMSEP=1.1871	[30]
	FM/(g·cm <sup>-2</sup> )			$R^2=0.3555$ , RMSEP=525.2932	

上述研究均是对内部品质的单一指标进行检测,忽略了各指标间的相关性,因此可能会对西甜瓜内部品质的综合评判造成影响。SSC 和 FM 作为西甜瓜成熟度的关键表征因子,孙静涛等<sup>[29]</sup>分别采用 MSC 和 SNV 对光谱信息进行预处理并选择 SPA, CARS 和 SPA-CARS 对变量进行筛选后,对比分析了基于全光谱、SSC 或 FM 单一特征和基于 PCA 特征融合的哈密瓜成熟度 SVM 判别模型的准确率。结果显示,基于 CARS-PCA-SVM 特征融合和全光谱 SNV-SVM 的哈密瓜成熟度判别模型的准确率较高,其校正集和预测集的判别准确率分别为 95%和 94%。

研究表明,利用变量筛选方法可有效降低高光谱信息冗余度,简化模型,提高模型预测精度。建立特征信息融合的成熟度判别模型为今后利用高光谱“图谱合一”的特性综合评判西甜瓜的内部品质提供了理论依据。

## 4 研究与展望

光谱分析作为一种现代无损检测技术,与传统检测技术相比具有快速、绿色、无损等优势,在西甜瓜内部品质无损检测中具有广阔的应用前景。

(1)光谱分析所获取的信息量较大,其中包含的许多与

被测组分无关的冗余信息会增加光谱信息处理时间,降低模型精度。获取光谱有效信息、优化预测模型一直是光谱分析技术未来发展的重要方向,而机器学习则是实现模型优化的重要手段。深度学习作为机器学习的分支在挖掘大数据集的有效信息方面发展迅猛,为解决光谱分析模型鲁棒性和适应性差等问题提供了新的思路和方法,应用深度模型预测西甜瓜的内部品质将是未来技术发展的必然趋势。

(2)目前对于西甜瓜内部品质的无损检测多是针对单一指标(大多数为影响口感风味的食用指标的定量检测),针对综合指标(如成熟度)的研究较少。研究西甜瓜内部品质评价指标间的内在联系,提取各指标相对应的变量建立多特征信息融合的西甜瓜内部品质综合评价模型将是今后的研究重点之一。

(3)尽管国内外部分学者都构建了基于光谱分析的西甜瓜内部品质无损在线检测系统,但目前大多停留在实验室阶段,应用于商业化生产线较少且多为大型装置,不利于加工企业和消费者使用。随着人工智能、大数据、移动互联网和云计算的不断发展,研究开发基于人工智能与移动终端深度融合的西甜瓜内部品质快速无损检测系统将成为新的研究方向之一。

## References

- [1] Ministry of Agriculture of the People's Republic of China(中华人民共和国农业部). NY/T 427—2016 Green Food—Watermelon and Muskmelon(NY/T 427—2016 绿色食品西甜瓜), 2016.
- [2] Tian H Q, Wang C G, Zhang H J, et al. Sensor Letters, 2012, 10(1): 570.
- [3] Wang A C, Fu X P, Xie L J. Food Analytical Methods, 2014, 8(6): 1403.
- [4] Wang H L, Peng J Y, Xie C Q, et al. Sensors, 2015, 15(5): 11889.

- [ 5 ] ElMasry G M, Nakauchi S. *Biosystems Engineering*, 2016, 142: 53.
- [ 6 ] Hussain A, Pu H B, Sun D W. *Trends in Food Science & Technology*, 2018, 72: 144.
- [ 7 ] Sun J T, Ma B X, Dong J, et al. *Journal of Food Process Engineering*, 2017, 40(3): e12496.
- [ 8 ] Arendse E, Fawole O A, Magwaza L S, et al. *Journal of Food Engineering*, 2018, 5(10): 22481.
- [ 9 ] Porep J U, Kammerer D R, Carle R. *Trends in Food Science & Technology*, 2015, 46(2): 211.
- [10] Li J L, Sun D W, Cheng J H. *Comprehensive Reviews in Food Science and Food Safety*, 2016, 15(5): 897.
- [11] PENG Yan-kun(彭彦昆). *Nondestructive Optical Technology for Agro-Food Quality and Safety Assessment(农畜产品品质安全光学无损快速检测技术)*. Beijing: Science Press (北京: 科学出版社), 2016. 2.
- [12] Wang H L, Peng J Y, Xie C Q, et al. *Sensors*, 2015, 15(5): 11889.
- [13] Yun Y H, Li H D, Deng B C, et al. *Trends in Analytical Chemistry*, 2019, 113: 102.
- [14] LIU Jun, WU Meng-ting, TAN Zheng-lin, et al(刘 军, 吴梦婷, 谭正林, 等). *Journal of Wuhan Institute of Technology(武汉工程大学学报)*, 2017, 39(5): 496.
- [15] Sirisomboon P. *Materials Today: Proceedings*, 2018, 5(10): 22481.
- [16] CHU Xiao-li(褚小立). *Practical Manual for Near Infrared Spectroscopy(近红外光谱分析技术实用手册)*. Beijing: China Machine Press (北京: 机械工业出版社), 2016.
- [17] Jie D F, Xie L J, Fu X P, et al. *Journal of Food Engineering*, 2013, 118(4): 387.
- [18] JIE Deng-fei, XIE Li-juan, RAO Xiu-qin, et al(介邓飞, 谢丽娟, 饶秀勤, 等). *Transactions of the Chinese Society of Agricultural Engineering(农业工程学报)*, 2013, 29(12): 264.
- [19] JIE Deng-fei, CHEN Meng, XIE Li-juan, et al(介邓飞, 陈 猛, 谢丽娟, 等). *Transactions of the Chinese Society of Agricultural Engineering(农业工程学报)*, 2014, 30(9): 229.
- [20] Jie D F, Xie L J, Rao X Q, et al. *Postharvest Biology and Technology*, 2014, 90: 1.
- [21] Qi S Y, Song S H, Jiang S N, et al. *Journal of Innovative Optical Health Sciences*, 2014, 7(4): 1350034.
- [22] QIAN Man, HUANG Wen-qian, WANG Qing-yan, et al(钱 曼, 黄文倩, 王庆艳, 等). *Spectroscopy and Spectral Analysis(光谱学与光谱分析)*, 2016, 36(6): 1700.
- [23] Tamburini E, Costa S, Rugiero I, et al. *Sensors*, 2017, 17(4): 746.
- [24] Jie D F, Zhou W H, Wei X. *Scientia Horticulturae*, 2019, 257: 108718.
- [25] Lu J, Qi S Y, Liu R, et al. *Journal of Innovative Optical Health Sciences*, 2015, 8(6): 1550032.
- [26] Zhang D Y, Xu L, Wang Q Y, et al. *Food Analytical Methods*, 2019, 12(1): 136.
- [27] Jie D F, Wei X. *Computers and Electronics in Agriculture*, 2018, 151: 156.
- [28] LI Feng-xia, MA Ben-xue, HE Qing-hai, et al(李锋霞, 马本学, 何青海, 等). *Acta Photonica Sinica(光子学报)*, 2013, 42(5): 592.
- [29] SUN Jing-tao, MA Ben-xue, DONG Juan, et al(孙静涛, 马本学, 董 娟, 等). *Spectroscopy and Spectral Analysis(光谱学与光谱分析)*, 2017, 37(7): 2184.
- [30] Sun M J, Zhang D, Liu L, et al. *Food Chemistry*, 2017, 218: 413.

## Recent Advances in Spectral Analysis Techniques for Non-Destructive Detection of Internal Quality in Watermelon and Muskmelon: A Review

MA Ben-xue<sup>1,2\*</sup>, YU Guo-wei<sup>1,2</sup>, WANG Wen-xia<sup>1,2</sup>, LUO Xiu-zhi<sup>1,2</sup>, LI Yu-jie<sup>1,2</sup>, LI Xiao-zhan<sup>1,2</sup>, LEI Sheng-yuan<sup>1,2</sup>

1. College of Mechanical and Electrical Engineering, Shihezi University, Shihezi 832003, China

2. Key Laboratory of Northwest Agricultural Equipment, Ministry of Agriculture, Shihezi 832003, China

**Abstract** Watermelon and muskmelon are sweet, juicy and rich in nutrients. There is great significance in manufacture and circulation for its internal quality detection. The traditional detection methods for internal quality of watermelon and muskmelon are inefficient, long time, high cost and destructive, which can not meet the needs of modern production. With the rapid development of spectral analysis techniques, near-infrared spectroscopy (NIRS) and hyperspectral imaging (HSI) for the internal quality non-destructive detection of watermelon and muskmelon has become a research hotspot. In order to track national and international progress of research, this paper presents the technical characteristics and system composition of NIRS and HIS. The spectral information analysis methods are concluded, including spectral information preprocessing, variable selection, model establishment and evaluation. Afterwards, the recent progress of NIRS and HSI in the non-destructive detection for the internal quality (soluble solids content, firmness, total acid content, maturity and moisture, etc.) of watermelon and muskmelon is

summarized. Finally, the future trends of spectral analysis techniques in the internal quality non-destructive detection of watermelon and muskmelon are discussed from the technical difficulties and practical applications. This review indicates that the following aspects are identified as the direction of future research, using deep learning methods to analyze spectral information, establishing comprehensive evaluation model of multi-feature information fusion, and developing the rapid non-destructive detection system based on the deep integration of artificial intelligence and mobile terminal.

**Keywords** Watermelon and muskmelon; Internal quality; Near-infrared spectroscopy; Hyperspectral imaging; Non-destructive detection; Review

(Received Jun. 10, 2019; accepted Oct. 21, 2019)

\* Corresponding author