

杜明华, 杨 甜, 马 燕, 等. 基于 NIR 高光谱成像技术的番茄叶片叶绿素含量检测[J]. 江苏农业科学, 2022, 50(20): 48–55.
doi:10.15889/j.issn.1002-1302.2022.20.007

基于 NIR 高光谱成像技术的番茄叶片叶绿素含量检测

杜明华, 杨 甜, 马 燕, 张 捷, 吴龙国

(宁夏大学农学院, 宁夏银川 750021)

摘要:利用近红外高光谱成像技术对番茄叶片叶绿素含量的无损检测进行初步探讨。通过高光谱成像系统(900 ~ 1 700 nm)采集了 192 个番茄叶片图像, 基于偏最小二乘回归模型(PLSR)对光谱进行样本集划分, 对原始光谱与 Kubelka – Munk 函数曲线及多种光谱预处理的偏最小二乘回归模型进行对比分析, 优选出多元散射校正(MSC)为预处理方法。采用 5 种方法提取特征波长, 并根据特征波长建立偏最小二乘回归、多元线性回归(MLR)、主成分回归(PCR)3 种模型的叶片叶绿素含量预测模型。结果表明, 建立无信息变量消除法(UVE)挑选特征波长的偏最小二乘回归模型最优, 其预测集的相关系数(R_p)为 0.8495, 均方根误差(RMSEP)为 4.3375。因此, 利用近红外高光谱成像技术提取特征波长进行叶绿素含量检测是可行的, 同时也为今后番茄品质在线检测提供了理论依据。

关键词:高光谱成像技术; 番茄叶片; 叶绿素含量; 快速诊断; 偏最小二乘回归模型

中图分类号:S641.201 **文献标志码:**A **文章编号:**1002-1302(2022)20-0048-08

番茄是我国需求量较高的蔬菜之一, 在全国范围内大面积栽培种植。番茄作为茄科作物之一, 本身所含的谷胱甘肽、维生素、番茄红素、 β -胡萝卜素等营养成分对于人们来说就具有很高的应用价值^[1-2]。因此, 番茄的高质量生产是必不可少的。番茄的生长状况可以通过叶片的叶绿素含量表现出来, 叶绿素含量是番茄重要的生理指标之一^[3-5]。可通过对叶绿素含量的检测从而对微咸水灌溉下番茄生长发育情况、健康状况做出监测。目前, 常见的测定方法往往操作流程繁琐、费时, 只能进行单个叶片检测^[4], 同时传统方法一般具有事后性和破坏性, 不能满足高效技术要求。因此, 无损、高效的检测方法是番茄高质量生产中重要的环节。

高光谱成像技术是近年来无损检测技术发展的新趋势, 将光谱信息和数字成像信息相集成, 主要是利用光谱反射率以及光谱反射波形特征, 从而可以精确地检测到所需样本图像中每一个像素点的光谱数据以及图像信息, 这种技术具有速度快, 效果好等优点, 同时在检测过程中不会破坏样本,

在很大程度上弥补了传统检测方法中存在的不足, 能够为可视化与高效无损检测提供一定的技术依据^[6]。高光谱技术被广泛应用于多个领域。包括农产品质量评估^[7]、品种鉴别^[8]、食品的安全检测^[9]、肉类保鲜等^[10-11]。Zhang 等以 160 个苹果为样品, 提取高光谱数据建立其可溶性固性物含量的预测模型, 结果表明, 预测效果较好的模型是基于 CARS – SPA 组合算法来提取的特征波长所建立的^[12]。Liu 等为了鉴别猕猴桃是否经过膨大剂处理, 采用了高光谱技术来检测并建立了多种模型来实现判别, 结果发现, 此技术可以很明显地判别多品种混合的猕猴桃膨大果^[13]。张静宜等利用高光谱成像技术来判别甜瓜叶片是否感染尾孢叶斑病, 结果显示, 高光谱图像对健康样本的判别率较高(>97%), 可用于健康样本与病变样本的识别^[14]。Feng 等利用近红外高光谱技术对比了不同预处理效果对番茄光谱模型检测精度的影响, 结果发现, 极限学习(ELM)获得了最好的预测结果^[15]。Sohrabi 等采用近红外光谱技术提取了番茄的可溶性固形物以及可滴定酸度的光谱数据, 并基于偏最小二乘回归法建立模型检测。结果表明, 可溶性固性物模型结果最好, 其预测相关系数达到了 0.91^[16]。袁瑞瑞等以灵武长枣为研究对象, 利用高光谱成像技术对长枣的损伤级别进行快速判别, 结果表明, MSC – CARS – PLS – DA 为最优模型组合, 校正集准确率为 77.14%, 预测集准确率为

收稿日期: 2021-09-03

基金项目: 第四批“宁夏青年科技人才托举工程”(编号: TJGC2019065); 宁夏重点研发计划(编号: 2018BBF02012)。

作者简介: 杜明华(1998—), 女, 河南焦作人, 硕士研究生, 研究方向为设施蔬菜栽培。E-mail: minghua980421@163.com。

通信作者: 吴龙国, 博士, 讲师, 硕士生导师, 主要从事设施园艺作物营养精准检测方面的研究。E-mail: wlg@nxu.edu.cn。

89.52%^[17]。邹志勇等对比分析了马铃薯在不同损伤条件下的高光谱不同波段下的图像特征并分别建立模型,结果表明,在建立的分类模型中,使用的轻量梯度提升机+逻辑斯蒂回归达到最高的判别精度 98.86%^[18]。

本研究以番茄叶片叶绿素含量为研究对象,结合高光谱成像技术采集样品 900~1 700 nm 光谱图像。使用共生距离法 (SPXY)、随机法 (RS) 和 Kennard-Stone 法划分样本集;之后对比分析基于原始光谱与 Kubelka-Munk 函数曲线建立的模型,选择最优光谱;在此基础上使用 PLSR 权重系数法、竞争性自适应重加权算法、无信息变量信息消除法、连续投影算法提取的特征波长进行偏最小二乘建模对比分析优选最佳特征波长提取方法;最终基于特征波长建立偏最小二乘回归模型、主成分回归模型以及多元线性回归模型进行对比分析。通过化学计量学结合高光谱成像技术可为今后番茄品质在线检测提供理论参考。

1 材料与方法

1.1 材料与仪器

试验于 2020 年 10 月至 2021 年 8 月在宁夏大

学农科实训基地玻璃温室中进行。供试材料采用博美 2 号番茄品种,购置于宁夏贺兰县天缘种业公司。共计 54 株番茄幼苗,于子叶展开 3~4 张时进行移苗,采用栽培基质盆栽的方式进行种植,每盆种植 1 株番茄幼苗。试验设计 6 个灌溉处理:对照组 (CK): 蒸馏水; A: 50 mmol/L NaCl 溶液; B: 100 mmol/L NaCl 溶液; C: 150 mmol/L NaCl 溶液; D: 200 mmol/L NaCl 溶液; E: 250 mmol/L NaCl 溶液,每个处理重复 3 次。

前期确定最适灌水量为 550 mL,定植后每隔 2 d 灌水 1 次,经过 90 d 种植后,分别采摘番茄上、中、下层叶片,每层采摘 2 张,试验共采集 216 张叶片,将采摘下的叶片编号放入塑封袋中,用低温保鲜箱立刻运回实验室,光谱图像采集之后,立即进行化学指标的测定。

仪器:便携式叶绿素测定仪 (SPAD-502Plus, 柯尼卡美能达)。HyperSIS-NIR 高光谱成像系统 (900~1 700 nm,光谱分辨率 5 nm,256 个波段),由高光谱成像光谱仪、CCD 相机、4 个 35W 卤钨灯、计算机数据采集软件 (ENVI4.7) 5 个部分组成^[19] (图 1),北京双利合谱仪器有限公司生产。

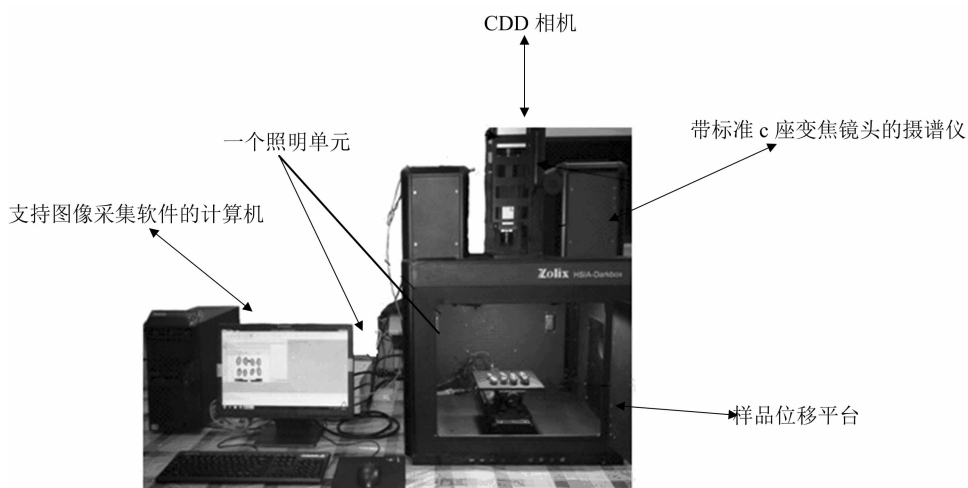


图1 高光谱成像系统

1.2 方法

1.2.1 叶绿素测定方法 叶绿素含量采用便携式叶绿素测定仪 (SPAD-502Plus, 柯尼卡美能达) 进行测量,每株番茄植株分别取上、中、下层各叶片,避开叶脉集中的部位,均匀地选取 3 个不同位置测定 SPAD 值并计算平均值代表该叶片数值。

1.2.2 番茄叶片光谱的测定 为了获取更加真实及清晰的图像,经过试验确定图像采集的参数为相

机曝光时间 25 ms,物距 360 mm,扫描速度 15 mm/s,线扫描实际长度 60 mm。消除试验过程中反射杂散光干扰,使用时,将电控位移平台背景设置为黑色。光谱扫描前需进行黑白校正^[20],其公式如下:

$$R = \frac{R_0 - D}{W - D} \quad (1)$$

式中: R_0 为原始图像; D 为黑板图像; W 为白板图

像; R 为标定后的高光谱图像。

1.3 数据分析方法

使用 The UnscramblerX10.4 对数据进行处理与建模,使用 Matlab2014a 提取特征波长及绘图。高光谱数据自身的不利影响会对处理和建模造成一定的影响。因此在数据处理与分析方面需要和化学计量学方法相结合,对模型优化^[21]。

1.3.1 样本集划分 样本集按 3 : 1 进行校正集和预测集的划分,与此同时结合建立偏最小二乘回归模型对比分析共生距离法、随机法和 Kennard - Stone 法划分样本。选择最佳划分方法,提高后期建模模型的性能。

1.3.2 预处理方法选择 在图像采集过程中,由于仪器自身的局限性,本身会带有一些噪音和暗电流等不利因素。同时也在一定程度上参杂着自非品质信息的影响。例如光谱曲线的基线漂移现象,多重共线性问题等。因此,为了避免这些不利因素对所获取样本光谱曲线的影响,以及保持图像的不失真需要对原始的光谱信息进行预处理,从而提取一些较为有用的信息^[22-25]。本研究使用平均平滑法 (Moving Average Smoothing)、归一化处理 (Normalize)、基线校准 (Baseline)、标准正态变化 (Standard Normal Variate)、去趋势化 (Detrend)、多元散射校正 (Multiplicative Scatter Correction) 6 种方法进行预处理,并建立偏最小二乘回归模型进行比较分析。

1.3.3 特征波长提取 大量冗余的光谱信息和光谱维数也为后期的建模提高了难度,增加了工作量。提取特征波长成为一种筛选数据极为重要且有效的方法。本研究在保留原始有效信息的前提下采用 PLSR 权重系数法 (β 权重系数)、竞争性自适应重加权算法 (CARS)、无信息变量信息消除法 (UVE)、连续投影算法 (SPA) 对预处理最优的样本进行光谱特征波长的提取。

1.3.4 预测模型的建立 对选取特征波长后的光谱信息进行建模。分别建立基于番茄叶片叶绿素含量定量分析的偏最小二乘回归模型 (PLSR)、多元线性回归 (MLR)、主成分回归 (PCR) 预测模型。模型性能的评价依据相关系数 (R) 及均方根误差 ($RMSEP$) 2 个指标。 R 值越大与 $RMSEP$ 值越小则越能表明模型性能好。 R 值表现了预测样本值与样本真实值之间的拟合状态,其值与拟合程度成正比,与模型效果成正比,而 $RMSEP$ 值则反映了样本

离散程度, $RMSEP$ 值与模型性能成反比,越小反映模型预测能力越好^[26]。为保证模型预测效果试验采用交叉验证方法对建模效果进行校正。

2 结果与分析

2.1 样本划分方法的选择及模型的建立

为了建立稳健的模型,需对样本集进行划分。本研究采用 RS、KS、SPXY 方法进行校正集与预测集样本划分,结果如表 1、表 2 所示。

表 1 不同样本划分方法对叶绿素的 PLSR 模型结果统计

方法	主成分数	校正集				预测集	
		R_c	$RMSEC$	R_{cv}	$RMSECV$	R_p	$RMSEP$
KS	9	0.906 9	3.700 2	0.832 9	4.882 4	0.756 2	5.402 3
RS	2	0.785 5	5.230 9	0.767 5	5.419 3	0.698 3	6.411 5
SPXY	9	0.894 0	3.980 8	0.820 6	5.087 6	0.808 1	4.668 4

从表 1 可以得出, SPXY 法的预测集相关系数 R_p 大于 0.8, 高于 KS 法和 RS 法。数值之和也可用来评价模型效果, 其值越大说明精确度越高^[27]。KS 法之和为 1.6631, RS 法为 1.483 8, SPXY 法为 1.702 1, SPXY 法的 R_c 与 R_p 之和数值也是 3 种方法数值最高的。并且采用 SPXY 法建立的模型中, 预测集的 $RMSEP$ 值均小。综上考虑, 选择 SPXY 法划分番茄叶片叶绿素样本集最为合适。

表 2 样本番茄叶片叶绿素统计

样本类别	样本数 (个)	最大值	最小值	均值	方差
校正集	128	46	8.30	29.15	8.89
预测集	64	46	8.40	31.33	7.83

从表 2 可以看出, 使用 SPXY 法对样本中叶绿素值进行划分是可行的。对于 192 个番茄叶片叶绿素值样本, 取 2/3 样本作为校正集, 1/3 样本作为预测集, 并且预测集的各指标值包含在校正集的指标值范围之内。校正集和预测集的方差值也较为理想, 由此可以说明所选样本数据集的划分具有代表性。

2.2 原始光谱特征波长下不同光谱参数的建模分析

为了研究全波段 (900 ~ 1 700 nm) 下 Kubelka - Munk 处理对建模的影响, 利用 Unscrambler X10.4 软件对原始光谱的反射率进行转换以及 Kubelka - Munk 函数转换, 研究不同光谱参数对番茄叶片叶绿素含量的预测。基于 PLSR 分别对 R、K - M 光谱曲线建立模型对比分析, 结果见表 3。

表 3 原始光谱特征波长下不同光谱参数的 PLSR 模型

光谱参数	主成分数	校正集				预测集	
		R_c	$RMSEC$	R_{CV}	$RMSECV$	R_p	$RMSEP$
R	9	0.894 0	3.980 8	0.820 6	5.087 6	0.808 1	4.668 4
K-M	9	0.880 7	4.209 8	0.785 6	5.547 7	-0.321 7	47.909 3

从表 3 可以看出, Kubelka - Munk 处理后建模的 R_p 值小于原始光谱建模的 R_p , 说明处理后的光谱数据建模的预测效果并不是太理想。原始光谱的 R_c 值大于 Kubelka - Munk 处理的 R_c 值。但二者之间的差异较小, 说明 2 个模型的性能差异不大, 而 Kubelka - Munk 处理只是模型性能做了细微的调改。因此, 可以采用原始光谱进行数据处理, 不仅

表 4 不同预处理方法对叶绿素含量的 PLSR 模型结果统计

类型	主成分数	校正集				预测集	
		R_c	$RMSEC$	R_{CV}	$RMSECV$	R_p	$RMSEP$
原始数据	9	0.894 0	3.980 8	0.820 6	5.087 6	0.808 1	4.668 4
MA	10	0.896 4	3.938 2	0.838 1	4.862 0	0.832 4	4.355 9
Normalize	7	0.859 3	4.544 6	0.788 8	5.494 9	0.811 3	4.667 4
Baseline	8	0.900 7	3.861 4	0.848 8	4.715 7	0.802 0	4.705 8
SNV	9	0.902 2	3.833 6	0.838 1	4.867 4	0.840 7	4.251 3
Detrend	9	0.915 9	3.568 0	0.834 2	4.925 8	0.764 1	5.178 3
MSC	9	0.902 9	3.820 3	0.823 1	5.070 7	0.849 3	4.143 0

从表 4 可以看出, 对于番茄叶片的中叶绿素指标来说, 经过 MSC 法对样本值进行预处理后建立的 PLSR 模型与原始数据相比, 具有较好的模型参数, R_c 和 R_p 值较之前有所提高, $RMSEC$ 和 $RESEP$ 都有所减小。综上所述, MSC 为番茄叶片叶绿素样本的最佳预处理方法。

2.4 特征波长的选取

2.4.1 PLSR 权重系数法提取特征波长 β 系数是反映评价指标重要程度的量化系数, 系数越大, 说明该指标对总目标影响越大。由图 2 可知, 基于对叶绿素进行 MSC 预处理后建立的 PLSR 模型进行特征提取, 第 4 成分具有最显著的特征, 选为最优主成分数用于提取特征波长; 在 900 ~ 1 700 nm 波段下, 叶绿素的吸收很弱, 光谱吸收特征不明显。所以在 β 系数图中按照局部绝对值最大原则选取 7 个叶绿素特征波长为 1 008、1 202、1 226、1 369、1 515、1 628、1 657 nm。有效地降低了光谱的维数, 方便后期建立模型。

2.4.2 竞争自适应重加权算法提取特征波长 竞争性自适应重加权算法即采用自适应重加权采样

减少了数据的前期处理, 而且能够获得较优的校正模型与预测模型。

2.3 预处理方法的选择

前期为了获得真实清晰的图像进行黑白校正, 这只是简单的对光源强度分布不均以及在图像采集过程中自身暗电流等不利因素进行了消除。而样品的背景色和杂散光等都还会对试验结果产生一定的误差, 为了提高模型的精确性, 需要对光谱进行一定的处理。而在对比分析中, 要充分考虑各个变量之间的影响, 而偏小二乘法能够综合考虑变量之间的关系, 能够在多重共线性条件下进行回归建模, 因此可以利用 PLSR 模型进行光谱预处理的结果比较, 结果见表 4。

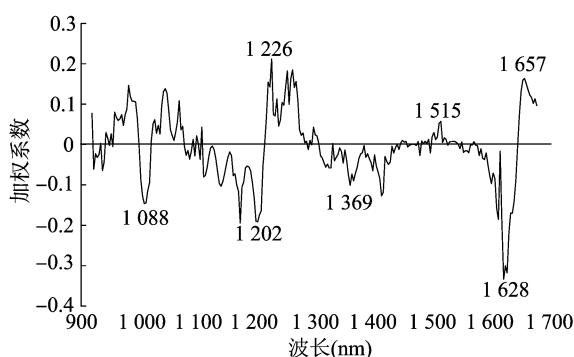


图 2 叶绿素 PLSR 权重系数法提取特征波长

(ARS)技术在模型中选择并保留下回归系数绝对值大的波段数, 同时, 去掉权重小的波段数, 此算法在计算取值的过程中是采用交互验证法选出均方根误差值最低的子集, 此算法所取值覆盖范围广, 因此可有效地寻出最优与最具代表性的变量组合。利用 CARS 方法提取 900 ~ 1 700 nm 范围内的特征波长, 得到结果见图 3。由图 3 可知, 在采样运算过程中, 变量数的下降过程呈现由快变慢的趋势, 图 3 - B 是基于自适应加权采样选择特征波长建立偏最小二乘模型预测效果的显示。图 3 - C 回归系数

路径则反映了特征波长变量回归系数的变化趋势，* 号线反映了 *RMSECV* 最小位置，* 号线之后部分有效信息被去除，*RMSECV* 值持续增大。叶绿素依

据 *RMSECV* 最小值原则选择 11 个特征波长，模型拟合效果最佳，数据压缩率为 95.7%，有效去除了环境干扰及不相干信息，但可能有效信息也被去除。

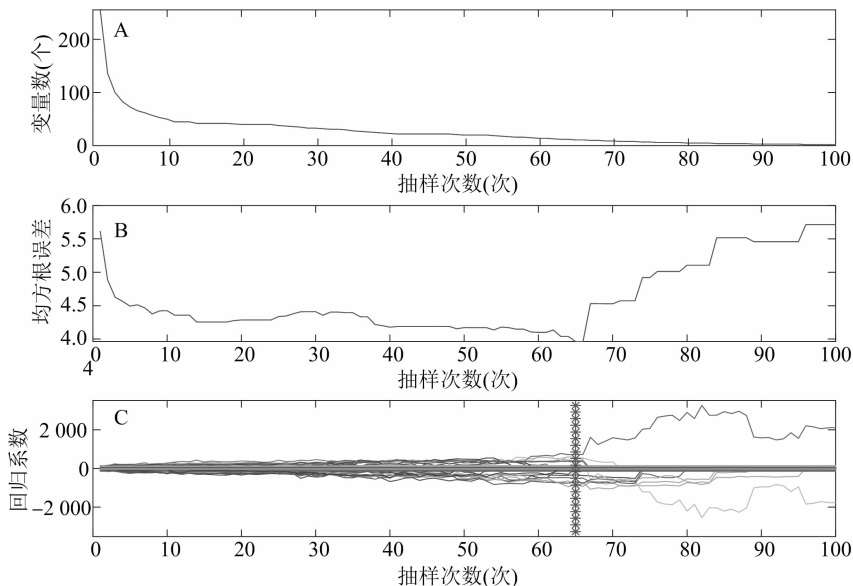


图3 叶绿素 CARS 提取特征波长

2.4.3 无信息变量消除变换法 (UVE) 提取特征波长 如图 4 所示,在 900 ~ 1 700 nm 波段范围内,在主成分数是 10 时 *t* - value 值模型拟合效果最佳。

图中左侧表示 256 个输入变量的分布曲线,而右侧图表示随机变量的分布曲线。最终,对于叶绿素选取了 96 个特征波长。

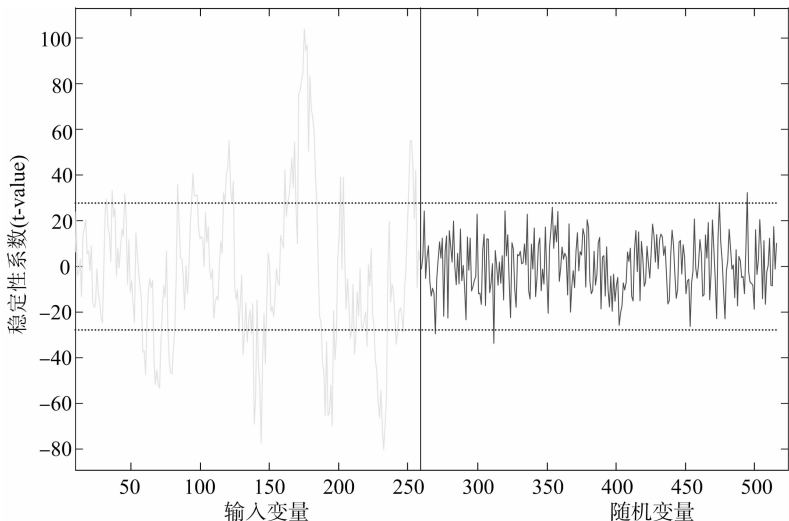


图4 叶绿素 UVE 提取特征波长

2.4.4 连续投影算法 (SPA) 提取特征波长 选择 $m - min = 10$, $m - max = 30$,图 5 - A 表示特征变量数与 *RMSE* 的变化曲线,可以得出 *RMSE* 最小值与特征变量数的关系,综合分析可以确定叶绿素特征波长数为 14 个,*RMSE* 值为 4.134 3。图 5 - B 为 SPA 算法所选取叶绿素的 14 个波段数。

2.4.5 UVE + SPA 组合算法提取特征波长 通过 UVE 法提取 900 ~ 1 700 nm 波段范围内的特征波长为 96 个,但与其他方法相比发现数据量仍然较大,为了降低维数,减少冗余,可以对数据进行进一步的提取与优化。在本试验中将 UVE 法和 SPA 法相结合来提取特征波长 (图 6)。

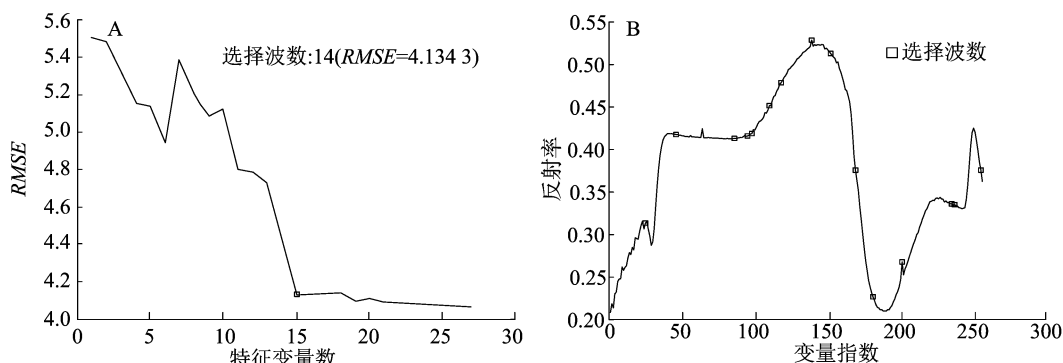


图5 叶绿素 SPA 提取特征波长

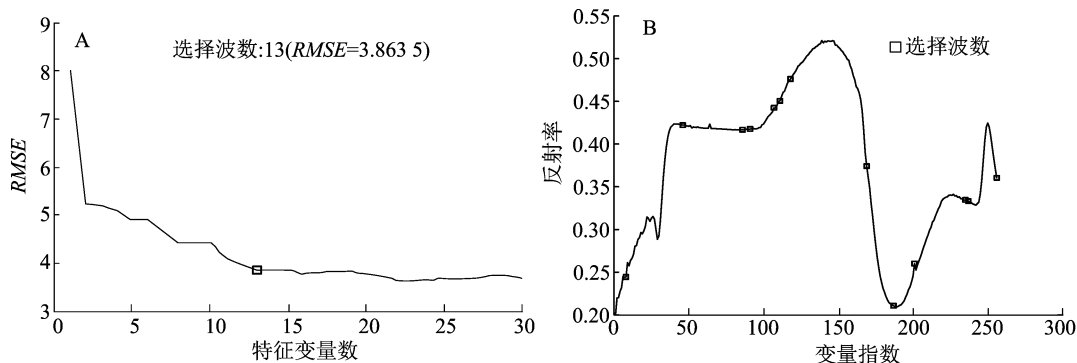


图6 叶绿素 UVE+SPA 提取特征波长

由图 6 可知,在 900 ~ 1 700 nm 波段对提取出的特征波长进一步采用 SPA 法,组合法 UVE + SPA 与 UVE 法相比,数据由 96 个减少到 13 个并且叶绿素的 RMSEC 值为 3.863 5,与 UVE 法相比,大大压

缩了数据,方便后续的建模与分析。

2.4.6 不同方法提取特征波长及对比分析 为了优选最佳的特征波段进行模型建立,对 5 种特征波长提取的波段进行分析,结果见表 5。

表 5 特征波长选取统计

特征提取方法	数量(个)	所选波长(nm)
β coefficient of PLSR	7	1 008, 1 202, 1 226, 1 369, 1 515, 1 628, 1 657
CARS	11	993, 1 002, 1 109, 1 261, 1 318, 1 425, 1 428, 1 467, 1 550, 1 589, 1 655
UVE	96	1 002, 1 014, 1 041, 1 044, 1 065, 1 080, 1 086, 1 103, 1 106, 1 109, 1 112, 1 115, 1 118, 1 133, 1 136, 1 139, 1 142, 1 145, 1 154, 1 184, 1 187, 1 190, 1 193, 1 196, 1 252, 1 255, 1 258, 1 261, 1 264, 1 267, 1 273, 1 285, 1 294, 1 297, 1 303, 1 312, 1 318, 1 321, 1 327, 1 330, 1 333, 1 336, 1 342, 1 384, 1 392, 1 395, 1 398, 1 401, 1 404, 1 413, 1 416, 1 419, 1 422, 1 425, 1 428, 1 431, 1 434, 1 437, 1 440, 1 443, 1 446, 1 449, 1 452, 1 464, 1 467, 1 470, 1 473, 1 476, 1 479, 1 482, 1 485, 1 491, 1 494, 1 503, 1 509, 1 524, 1 527, 1 533, 1 536, 1 559, 1 577, 1 580, 1 583, 1 586, 1 589, 1 592, 1 595, 1 598, 1 601, 1 622, 1 625, 1 649, 1 652, 1 655, 1 658, 1 664
SPA	14	1 047, 1 166, 1 192, 1 202, 1 238, 1 261, 1 324, 1 363, 1 413, 1 449, 1 509, 1 610, 1 616, 1 670
UVE + SPA	13	934, 1 046, 1 166, 1 181, 1 229, 1 240, 1 261, 1 413, 1 467, 1 509, 1 610, 1 616, 1 673

由表 5 可知,5 种特征波长提取方法中, β coefficient of PLSR 提取了 7 个特征波长,占总波长的 2.7%;CARS 提取 11 个,占总波长的 4.3%;UVE 提取 96 个,占总波长 37.5%;SPA 提取 14 个,占总波长的 5.5%;UVE + SPA 提取 13 个,占总波长的 5.1%。由于 UVE 法选取的波长数较多,所以和 SPA 组合使用,可以减低冗余,减少维数。采用这 5 种方法选取特征波长,各个波段均有挑选,具有一定的合理性。

2.5 建模方法的比较分析

为了对比不同建模方法对提取特征波长的建模效果,采用多元线性回归、主成分回归、偏最小二乘回归方法进行模型对比分析,结果见表 6。

由表 6 可知,(1)与偏最小二乘回归模型对比,UVE 挑选出来的特征波长建立的模型 R_c 和 R_{cv} 最大, RMSEC 和 RMSECV 最小,校正模型优于其他模型。在预测能力上,采用 UVE 法提取的特征波长建立的模型有最大的 R_p 值 0.849 5 和最低的 RMSEP

表 6 不同特征提取方法建模效果对比

方法	特征提取方法	主成分 分数	校正集				预测集	
			R_C	$RMSEC$	R_{CV}	$RMSECV$	R_P	$RMSEP$
PLSR	β coefficient of PLSR	7	0.739 9	5.798 4	0.698 4	6.178 5	0.820 4	5.181 8
	CARS	8	0.815 7	5.007 4	0.780 6	5.419 1	0.818 0	5.003 7
	UVE	6	0.862 4	4.448 9	0.814 0	5.110 1	0.849 5	4.337 5
	SPA	5	0.798 6	5.280 9	0.741 7	5.907 5	0.825 0	5.146 8
	UVE + SPA	5	0.772 9	5.396 6	0.718 3	5.932 5	0.818 0	5.386 3
MLR	β coefficient of PLSR	0	0.739 9	5.998 6	0.698 3	6.202 8	0.820 4	5.118 2
	CARS	0	0.817 1	5.242 2	0.775 2	5.023 0	0.818 0	5.010 2
	UVE	0	0.975 4	3.930 7	0.551 1	9.891 7	0.694 5	7.540 5
	SPA	0	0.818 5	5.364 0	0.757 9	5.778 7	0.820 4	5.367 5
	UVE + SPA	0	0.798 5	5.425 0	0.727 4	5.898 4	0.846 3	5.160 1
PCR	β coefficient of PLSR	7	0.739 9	5.798 4	0.698 4	6.178 5	0.820 4	5.118 2
	CARS	8	0.849 7	4.489 9	0.787 8	5.276 0	0.849 7	4.489 9
	UVE	10	0.828 7	4.913 7	0.787 1	5.421 9	0.805 5	4.876 6
	SPA	10	0.778 8	5.503 5	0.721 5	6.065 2	0.795 0	5.423 5
	UVE + SPA	8	0.449 6	5.629 3	0.699 0	6.096 2	0.829 1	5.247 6

值 4.337 5,说明该模型预测能力的稳定性优于其他模型。综合各个参数,采用 5 种方法提取的特征波长建立的 PLSR 模型中 UVE 挑选出来的波长所建立的是最优模型。(2)多元线性回归模型对比,在 5 种提取特征波长的方法中 UVE 法的 R_C 值最大 $RMSEC$ 值最小,但 R_{CV} 值较小,说明预测能力的稳定性不强。而 SPA 法的 R_C 值次于 UVE,但 R_{CV} 值较高。从综合参数角度来看,SPA 法建模效果是 5 中

方法中最佳的,所以,选择 SPA 建立 MLR 模型。(3)主成分回归模型对比,CARS 提取特征波长的方法,均具有最高的 R_C 、 R_{CV} 和 R_P 值,最低的 $RMSE$ 、 $RMSECV$ 、 $RMSEP$ 值。因此,无论是校正能力还是预测能力,CARS 效果都是 5 种方法中最佳的。

2.6 最优特征波长模型比较

通过光谱的预处理,提取特征波长等方法,可以得出最优的特征波长的建模方法(表 7)。

表 7 最优特征波长模型对比

方法	主成分 分数	校正集				预测集	
		R_C	$RMSEC$	R_{CV}	$RMSECV$	R_P	$RMSEP$
PLSR	6	0.885 2	4.134 2	0.811 9	5.322 2	0.809 2	4.636 6
UVE - PLSR	6	0.862 4	4.448 9	0.814 0	5.110 1	0.849 5	4.337 5
SPA - MLR	0	0.818 5	5.364 0	0.757 9	5.778 7	0.820 4	5.367 5
CARS - PCR	8	0.813 1	5.038 5	0.745 0	5.785 6	0.849 7	4.489 9

由表 7 可知,使用 UVE 法建立的 PLSR 模型与其他 2 个相比,模型性能参数较好,拥有较高的 R_C 值 0.862 4 和较低的 $RMSEC$ 值 4.448 9,预测模型的 R_P 在 3 种模型中也有较好效果,具有一定的校正能力和预测能力,并且与全波段建立的 PLSR 模型的 R_C 值较为接近,差异小。综合参数来看,UVE - PLSR 模型代表全波段建模是具有可行性的,UVE - PLSR 模型结果见图 7。

3 结论

本试验提出了基于近红外高光谱成像技术对

叶绿素含量的快速检测。将成像技术与化学计量学方法相结合,建立了番茄叶片叶绿素含量的 PLSR 模型。对叶绿素采用 SPXY 法进行样本划分。并基于原始光谱与 Kubelka - Munk 处理光谱参数建立模型比较,对比分析可知经 Kubelka - Munk 处理后的模型参数低于原始光谱参数。因此,试验采用原始光谱进行后期的数据分析。最终建立并分析了基于特征波长的 PLSR、MLR、PCR 模型,并对模型的性能进行了评价。其中 UVE - PLSR 最优,其预测模型的 R_P 值为 0.849 5, $RMSEP$ 为 4.337 5。本试验利用 NIR 高光谱成像技术对番茄叶片叶绿素进行

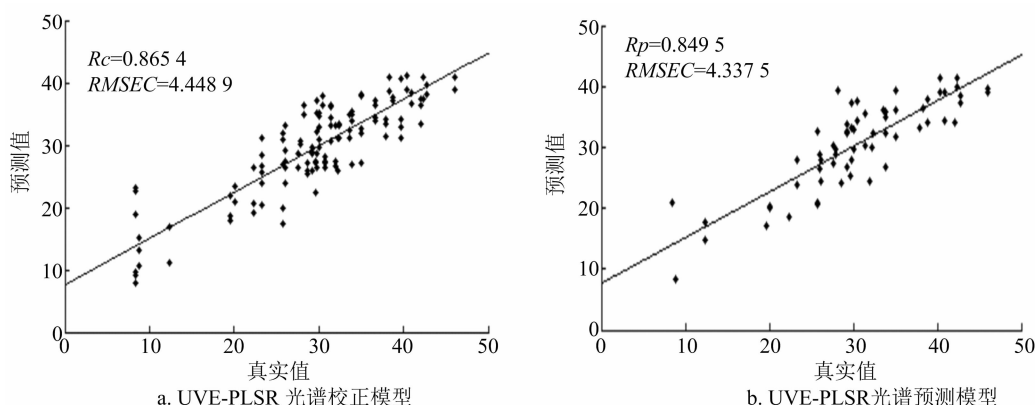


图7 UVE-PLSR 建模结果

无损检测是可行的。能为今后高光谱成像技术应用于番茄品质的在线检测提供参考,也可为推动宁夏回族自治区地方番茄产业快速发展提供技术支持。

参考文献:

- [1] 霍建勇. 中国番茄产业现状及安全防范[J]. 蔬菜, 2016(6): 1-4.
- [2] Feng L, Zhang M, Bhandari B, et al. A novel method using MOS electronic nose and ELM for predicting postharvest quality of cherry tomato fruit treated with high pressure argon [J]. Computers and Electronics in Agriculture, 2018, 154: 411-419.
- [3] 袁小康, 周广胜, 王秋玲, 等. 不同灌溉量夏玉米叶绿素含量的高光谱特征及其反演[J]. 生态学报, 2021, 41(2): 543-552.
- [4] 李艺迪, 顾宸瑞, 江慧欣, 等. 转基因金叶银中杨叶色及生长变异分析[J]. 植物研究, 2020, 40(6): 897-905.
- [5] 卢纯, 张亚红, 李青. LED不同光质补光对日光温室冬春茬番茄生长及光合特性的影响[J]. 江苏农业科学, 2020, 48(8): 127-134.
- [6] 符凯娟, 冯全, 陈佰鸿. 基于图像特征的草莓叶片含水率检测模型[J]. 林业机械与木工设备, 2020, 48(8): 27-34.
- [7] Siedliska A, Baranowski P, Zubik M, et al. Detection of fungal infections in strawberry fruit by VNIR/SWIR hyperspectral imaging [J]. Postharvest Biology and Technology, 2018, 139: 115-126.
- [8] 孙红, 刘宁, 吴莉, 等. 高光谱成像的马铃薯叶片含水率分布可视化[J]. 光谱学与光谱分析, 2019, 39(3): 910-916.
- [9] Thomas S, Kuska M T, Bohnenkamp D, et al. Benefits of hyperspectral imaging for plant disease detection and plant protection: a technical perspective[J]. Journal of Plant Diseases and Protection, 2018, 125(1): 5-20.
- [10] 何鸿举, 王魏, 李波, 等. 近红外高光谱快速无接触评估冷鲜猪肉脂质氧化[J]. 食品与机械, 2020, 36(8): 117-122.
- [11] 闫晓静. 基于光谱和图像信息的进口冰鲜与解冻牛肉品质检测研究[D]. 镇江: 江苏大学, 2019.
- [12] Zhang D Y, Xu Y F, Huang W Q, et al. Nondestructive measurement of soluble solids content in apple using near infrared hyperspectral imaging coupled with wavelength selection algorithm [J]. Infrared Physics & Technology, 2019, 98: 297-304.
- [13] Liu D Y, Guo W C. Identifying CPPU-treated kiwifruits using near-infrared hyperspectral imaging technology [J]. Food Analytical Methods, 2017, 10(5): 1273-1283.
- [14] 张静宜, 陈锦超, 傅霞萍, 等. 甜瓜尾孢叶斑病的高光谱成像检测[J]. 光谱学与光谱分析, 2019, 39(10): 3184-3188.
- [15] Feng L, Zhang M, Adhikari B, et al. Nondestructive detection of postharvest quality of cherry tomatoes using a portable NIR spectrometer and chemometric algorithms [J]. Food Analytical Methods, 2019, 12(4): 914-925.
- [16] Sohrabi M M, Ahmadi E, Monavar H M. Nondestructive analysis of packaged grape tomatoes quality using PCA and PLS regression by means of fiber optic spectroscopy during storage [J]. Journal of Food Measurement and Characterization, 2018, 12(2): 949-966.
- [17] 袁瑞瑞, 刘贵珊, 何建国, 等. 可见近红外高光谱成像对灵武长枣定量损伤等级判别[J]. 光谱学与光谱分析, 2021, 41(4): 1182-1187.
- [18] 邹志勇, 吴向伟, 陈永明, 等. 低温冷冻和机械损伤条件下马铃薯高光谱图像特征响应特性研究[J]. 光谱学与光谱分析, 2019, 39(11): 3571-3578.
- [19] 吴龙国, 何建国, 刘贵珊, 等. 基于近红外高光谱成像技术的长枣含水量无损检测[J]. 光电子·激光, 2014, 25(1): 135-140.
- [20] Polder G, van der Heijden G W A M, Keizer L C P, et al. Calibration and characterisation of imaging spectrographs [J]. Journal of Near Infrared Spectroscopy, 2003, 11(3): 193-210.
- [21] 孙红, 陈香, 孙梓淳, 等. 基于透射光谱的玉米叶片含水率快速检测仪研究[J]. 农业机械学报, 2018, 49(3): 173-178.
- [22] 冯愈钦, 吴龙国, 何建国, 等. 基于高光谱成像技术的长枣不同保藏温度的可溶性固形物含量检测方法[J]. 发光学报, 2016, 37(8): 1014-1022.
- [23] 王松磊, 吴龙国, 康宁波, 等. 基于高光谱图谱融合技术的宁夏滩羊肉嫩度检测方法研究[J]. 光电子·激光, 2016, 27(9): 987-995.
- [24] 王彩霞, 王松磊, 贺晓光, 等. 高光谱技术融合图像信息的牛肉品种识别方法研究[J]. 光谱学与光谱分析, 2020, 40(3): 911-916.
- [25] 王松磊, 吴龙国, 王彩霞, 等. 可见近红外高光谱快速诊断番茄叶片含水量及其分布[J]. 光电子·激光, 2019, 30(9): 941-950.
- [26] 段萌, 杨伟才, 毛晓敏. 覆膜条件下水分亏缺对春小麦光合特性影响及光响应模型比较研究[J]. 农业机械学报, 2019, 49(1): 219-227.
- [27] 董金磊, 郭文川. 采后猕猴桃可溶性固形物含量的高光谱无损检测[J]. 食品科学, 2015, 36(16): 101-106.