

紫外/可见光谱技术无损检测苹果的挤压损伤

尚静, 张艳, 孟庆龙
(贵阳学院, 贵阳 550005)

摘要: **目的** 通过紫外/可见光谱技术结合模式识别算法, 建立挤压损伤苹果的 Fisher 识别模型、K 最近邻 (KNN) 识别模型和偏最小二乘判别分析 (PLS-DA) 识别模型。**方法** 以挤压损伤苹果和无损苹果为研究对象, 采用光谱仪采集 2 种苹果的光谱反射率, 综合比较不同光谱预处理方法 (二阶微分 (SD)、标准正态变换 (SNV) 和多元散射校正 (MSC)) 对各模型识别效果的影响, 并利用主成分分析方法 (PCA) 对预处理后的光谱数据进行降维, 并提取能反映损伤苹果的特征光谱。**结果** 采用主成分分析法选择了累计贡献率超过 99% 的前 7 个主成分 (P_1-P_7) 作为特征光谱数据, 有效地实现了光谱数据的降维; 二阶微分对光谱反射率预处理的效果最好; 3 种判别模型均能满足实际要求, 且 SD+Fisher 和 SD+PLS-DA 识别模型对校正集和预测集样本的总正确识别率均高达 100%。**结论** 研究结果有助于实现挤压损伤苹果的快速识别。

关键词: 光谱技术; 无损检测; 模式识别; 损伤苹果

中图分类号: TB487 文献标识码: A 文章编号: 1001-3563(2019)13-0025-06

DOI: 10.19554/j.cnki.1001-3563.2019.13.004

Detection of Pressed Damage on Apples Based on UV/VIS Spectroscopy

SHANG Jing, ZHANG Yan, MENG Qing-long

(Guiyang University, Guiyang 550005, China)

ABSTRACT: The work aims to establish the discriminant models of Fisher, K nearest neighbor (KNN) and partial least-square discriminant analysis (PLS-DA) for the pressed apples based on ultraviolet radiation/visible spectroscopy technology combined with pattern recognition algorithm. With pressed apple and intact apple as the study object, the optical fiber spectrum system was used to acquire the spectral reflectance of these two kinds of apple. The effect of different spectral pretreatment methods (Second Derivation (SD), Standard Normal Variation (SNV) and Multi-Scatter Calibration (MSC)) on the recognition effect of each model was comprehensively compared, and the principal component analysis (PCA) was applied for the dimensionality reduction of the pretreated spectral data, and the characteristic spectrum that could reflect the pressed apple was extracted. The results showed that the first 7 principal components (P_1-P_7) with cumulative contribution rate of 99% were selected as the characteristic spectral data by the principal component analysis, and the dimensionality reduction of the spectral data was well realized. The preprocessing effect of Second Derivation on spectral reflectivity was the best, and three models could all meet the practical requirements, especially both SD+Fisher and

收稿日期: 2019-03-18

基金项目: 国家自然科学基金 (61505036); 贵州省科技厅联合基金 (黔科合 LH 字[2014]7174 号); 贵州省教育厅青年科技人才成长项目 (黔教合 KY 字[2018]290); 贵州省普通高等学校工程研究中心项目 (黔教合 KY 字[2016]017); 贵阳市财政支持贵阳学院学科与硕士点建设项目【SY-2019】

作者简介: 尚静 (1988—), 女, 硕士, 贵阳学院助教, 主要研究方向为农产品无损检测。

通信作者: 张艳 (1977—), 女, 博士, 贵阳学院教授, 主要研究方向为农产品无损检测。

SD+PLS-DA models had the optimal recognition performance with a total correct recognition rate of 100% for the samples of calibration set and prediction set. The research results are conducive to the fast recognition of pressed apples.

KEY WORDS: spectroscopy technology; nondestructive detection; pattern recognition; pressed apple

我国的苹果产量大,一直高居世界首位,但是出口量却不理想,特别是近10年以来持续走低,远低于世界平均水平^[1]。在如此迫切的情况下,如何保证水果的品质、提升出口量显得尤为重要。由于苹果数量大、种类多、品质参差不齐,在采摘或运输过程中还会因碰撞和挤压造成隐性的表面损伤,这些都很难通过人工分级达到品质筛选的要求。传统的机械分级^[2-3]和化学检测手段都存在一定的缺陷,比如破坏水果本身不利于保存或运输,检测时间较长耗费成本,指标单一检测不全面等。由此可见,开发研制一种快速、无损、高效的苹果表面损伤检测方法在水果分级领域中的应用前景非常可观。

最近几年,基于光谱技术的无损检测技术具有快速、无损、可实时在线检测等优点,已经被广泛地应用于农业、食品和医药等诸多领域^[4-6]。国内外研究学者对水果的品质检测分析主要集中在水果内部品质(可溶性物质、硬度等)的定量检测方面^[7-10],而在水果外部品质(表面损伤和缺陷等)方面^[11-13]仍需做更多的努力。针对苹果的大小、外形、颜色、污物、黑斑、外伤等外部品质的研究也有报道。早在2006年,LEFCOUT A M等^[14]采用图像处理技术对黄元帅和红元帅等2种苹果上的污物进行了鉴别。赵杰文等^[15]于2008年采用主成分分析法选择最佳波长预测了苹果的轻微损伤。2012年,程国首等^[16]通过高光谱成像技术对新疆红富士苹果的颜色进行了分级,机器分级结果与人工分级的一致率高达97.7%。Zhu Q等^[17]应用高光谱成像技术结合偏最小二乘法对被小铅球敲击受损后的苹果进行了淤青检测,建立的预测模型的正确预测率在84.9%~91.9%之间,正确识别率相对较低。最近,刘晶晶等^[18]利用高光谱分选仪获取了苹果图像,经过二阶巴特沃斯高通滤波器滤波后对苹果的外部损伤进行了检测,正确率达到95%。总之,虽然对苹果的外部品质进行了一些相关研究,但是在分选效果方面仍有很大的提升空间,需要建立一套完整的方案来高效、快速地实现苹果的等级分选。

文中以“红富士”苹果为研究对象,基于紫外/可见光谱技术结合模式识别算法,分别建立挤压损伤苹果和正常苹果样本的Fisher识别模型、K最近邻(K Nearest Neighbor, KNN)识别模型和偏最小二乘判别分析(Partial Least Square Discrimination Analysis, PLS-DA)识别模型,尝试通过光谱信息结合多种识别模型对轻微损伤苹果进行无损检测,并对比分析不同识别模型检测损伤苹果的正确识别率。

1 实验

1.1 材料

实验所用“红富士”苹果购买于当地的沃尔玛超市,从买回来的苹果中挑选一批完好无损、大小均匀的苹果样本共240个。将240个苹果样本随机分为3组:实验组①、实验组②和对照组③,每组80个。分别对实验组①和实验组②中的苹果样本进行人为损伤,采用电子式压力试验机(CBS-50K)以压头直径为20 mm,试验速度为10 mm/min,最大压力分别为100, 125 N来缓慢挤压2个实验组的苹果,使实验组①中的苹果样本发生1.1~2.3 mm的形变(即在每个苹果样本的表面中心产生相当于采摘或运输过程中造成的轻微损伤),使实验组②中的苹果样本发生1.5~2.7 mm的形变(该实验组的样本比实验组①的样本损伤稍重),并用记号笔将每个苹果样本的损伤区域标记出来。对照组③中的80个苹果样本不做任何处理,为正常苹果样本。分别将实验组①、实验组②和对照组③中的60个样本作为校正集样本(共180个),另外的20个样本作为检验集样本(共60个)。损伤苹果和正常苹果测试区域见图1。

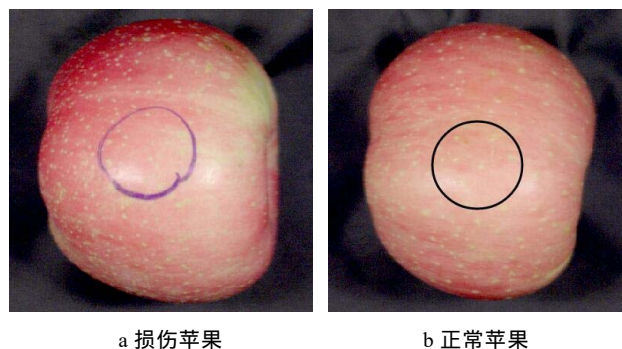


图1 苹果样本测试区域
Fig.1 Test areas of apple sample

1.2 光谱采集系统

该试验使用的光谱采集系统见图2,系统主要由高性能光谱仪QEPro(海洋光学)、光纤R600-7-VIS-125F(由7股光纤紧紧捆扎在一起组成,套环为不锈钢材质,其中6股为照明光纤,中间的1股为读出光纤,光纤的直径为600 μm)、卤钨灯光源HL-2000、反射探头支架RPH-1(用于固定反射探头的表面镀铝的探头架,反射探头与表面呈90°)、RPH-ADP适配器(用于使探头适合于RPH-1型探头架,使探头架可

以固定直径为 3.17 mm (1/8 英寸) 的探头)、标准反射白板 WS-1 和计算机等组成。光谱仪的波长范围为 198.2~1006.4 nm, 共 1044 个波段。光学分辨率 (FWHM) 为 0.8 nm, 信噪比为 1000 : 1, A/D 分辨率为 18 bit, 量子效率为 90%。



图 2 光谱采集系统示意

Fig.2 Schematic diagram of spectral acquisition system

将待测苹果样本紧贴在反射探头支架 RPH-1 表面 (反射探头通过 RPH-ADP 适配器固定在反射探头支架 RPH-1 上, 距离反射探头支架表面约 1 cm), 测量位置见图 1。将 5 次测量结果的平均值作为完好无损和表面有挤压损伤苹果样本的光谱反射率, 并将采集到的光谱数据以文本形式存入计算机。光谱采集系统的积分时间为 110 ms, 扫描平均次数为 8, 滑动平均宽度为 1。在实际应用中, 为了能采集整个苹果样本表面的光谱信息, 可以将待测样本放在旋转的支架上进行光谱的采集。

1.3 光谱预处理方法

通常采用的光谱预处理方法包括二阶微分 (Second Derivation, SD)、标准正态变换 (Standard Normal Variation, SNV) 和多元散射校正 (Multi-scatter Calibration, MSC) 等^[19]。其中, 二阶微分用来消除线性背景平移; SNV 可以用来校正光谱中因散射带来的误差; 在 MSC 预处理方法中, 认为光谱数据集中的每条光谱都应该与“理想”光谱呈线性关系, 但是真正的“理想”光谱是无法得到的, 通常可以采用样本光谱数据集的平均光谱来近似表示。试验采用 Matlab R2016b 软件来实现光谱的预处理。

1.4 建模方法

常用有监督模式的识别方法^[19]包括偏最小二乘判别分析法 (PLS-DA)、Fisher 判别法以及 K 最近邻法 (KNN)。其中, PLS-DA 将自变量 X (样本的特征数据) 与 Y (样本的类别信息, 以 0, 1 为元素的矩

阵) 的每一列数据分别进行偏最小二乘回归分析, 进而获得每个样本的类别信息矩阵, 然后再根据该类别信息矩阵中每一列数据与 1 接近的程度来判别样本属于哪类, 当样本属于其中某类时, 则该样本在矩阵 Y 中对应列的元素值为 1, 否则等于 0。

Fisher 判别法是一种把高维空间的点向低维空间投影的方法, 通过在 2 类样本集点群之间寻找一个最优分割面, 使得样本集点在该分割面上的投影达到最大限度的分割, 同时使每类样本内部的离散度达到最小, 将满足这个使类间距离最大而类内距离最小的准则称为 Fisher 准则。即将样本点的特征记为 $x_{ik,g}$, 此处 $i=1, 2 \dots p$ (p 为特征个数), $g=1, 2$ (g 表示类别), $k=1, 2 \dots m_g$ (k 为模式类别序号, m_g 为 g 类样本数), 样本总数 $n=n_1+n_2$, m_1 为类 1 的样本总数, m_2 为类 2 的样本总数, Fisher 判别函数为:

$$D_{gk}(x) = \sum_{i=1}^p \omega_i x_{ik,g} \quad g=1, 2; k=1, 2 \dots n_g \quad (1)$$

式中, $\omega = (\omega_1, \omega_2 \dots \omega_p)^T$ 。

在 KNN 判别分析中, 将训练集中每类样本点的重心求出, 然后判别未知样本点与各类重心的距离。未知样本点距哪一类重心距离最近, 未知样本就属于哪类。文中采用一种简化的算法——类重心法, 建立 KNN 识别模型, 即将训练集中每类样本点的重心求出, 然后判别未知样本点与各类重心的距离。未知样本点距哪类重心距离最近, 就属于哪类。若用几个特征对模式进行分类, 则对于 1 个 g 类问题, 第 j 类模式的重心为 c_j ($c_{j1}, c_{j2} \dots c_{jp}$)。

$$c_{ji} = \frac{1}{n_j} \sum_{k=1}^{n_j} x_{kj} \quad i=1, 2 \dots p; j=1, 2 \dots g \quad (2)$$

式中: i 为特征序号; j 为类序号; n_j 为第 j 类模式的样本数目。

1.5 数据分析

样本集的光谱反射率采集由 OceanView (Ocean Optics, copyright 2013) 软件来完成, 采集完光谱数据后在 Matlab R2016b 软件中对样本的原始光谱进行预处理, 并建立识别模型。

2 结果与分析

2.1 苹果样本的光谱反射率及预处理

为了消除原始光谱起始端和末端噪声的影响, 选择波长范围在 204.7~997.6 nm 内的 1024 个波段作为有效的光谱区域。该试验进一步利用不同的光谱预处理方法 (SD, SNV, MSC) 对原始光谱数据进行预处理, 以消除光谱数据中存在的基线漂移和噪声等对光谱的干扰, 获得更高信噪比的光谱数据, 有助于提高建立识别模型的稳健性。苹果样本的原始光谱反射率

(图 3a), 以及经过 SD (图 3b), SNV (图 3c) 和 MSC (图 3d) 预处理后的光谱反射率曲线见图 3。

2.2 主成分分析

苹果样本的原始光谱数据以及经过 SD, SNV 和 MSC 预处理后的光谱数据再经过 PCA 变换后的第 1

主成分和第 2 主成分得分分布见图 4。从图 4a 中可以看出, 原始光谱直接经过 PCA 变换后, 第 1 和第 2 主成分得分分布是分散开的, 不适合直接用来建立识别模型; 通过二阶微分对原始光谱进行预处理, 再经 PCA 变换后得到的分类结果具有非常好的分类效果, 见图 4b。

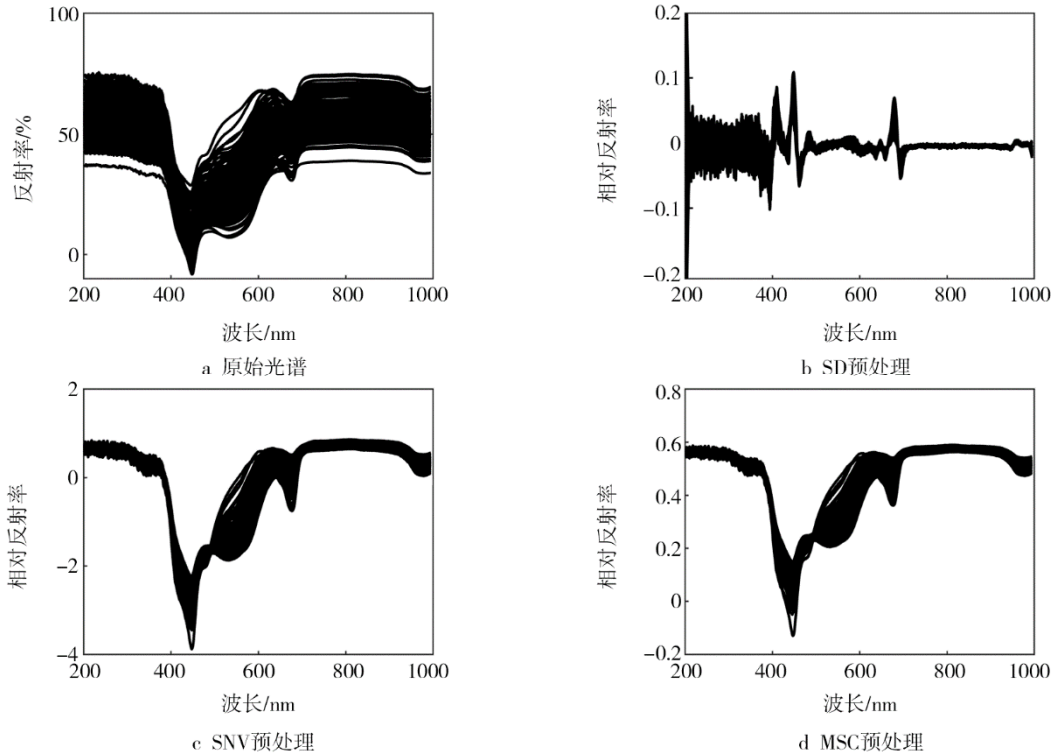


图 3 光谱反射率曲线
Fig.3 Spectral reflectance curves

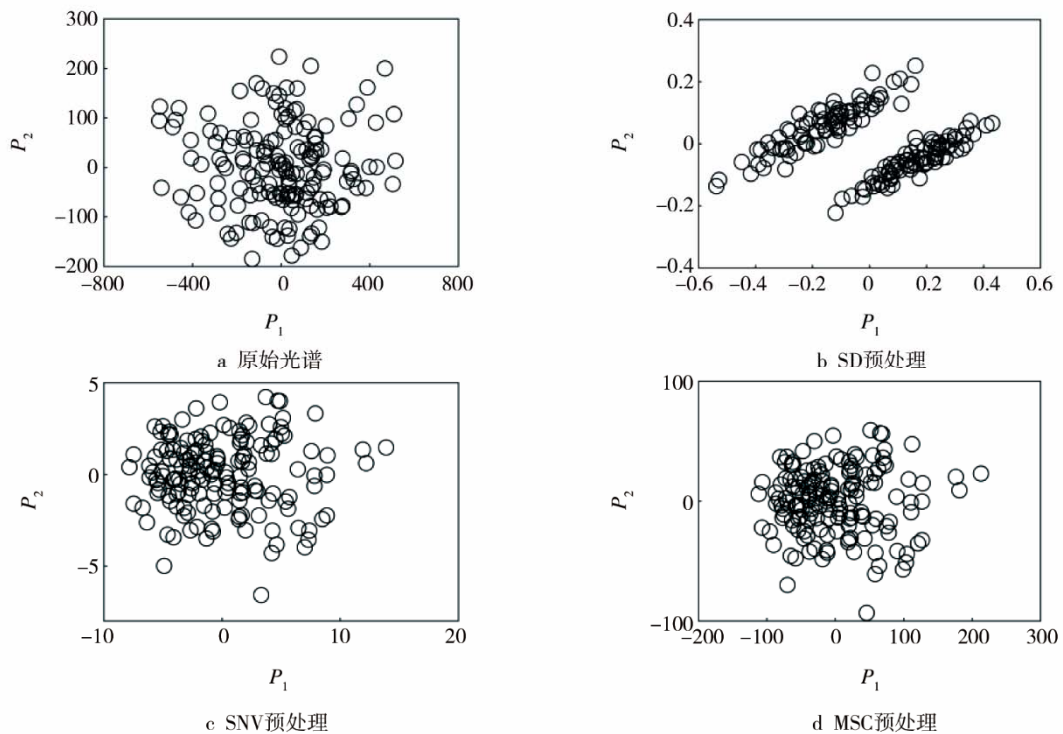


图 4 不同光谱数据的第 1 主成分和第 2 主成分的得分分布
Fig.4 Scores scatter plot of P_1 and P_2 of different spectra

另外, 进一步从主成分得分的累计贡献可知前7个主成分(第1主成分 P_1 , 第2主成分 P_2 ...第7主成分 P_7)可表达99%以上的原始信息, 前7个主成分(P_1 — P_7)的得分贡献率及其累计贡献率见表1。为了提高所建识别模型的正确率, 将前7个主成分(新变量 P_1 — P_7)作为样本集特征光谱数据。

表1 主成分得分贡献率

主成分	贡献率/%	累计贡献率/%
P_1	73.91	73.91
P_2	14.81	88.72
P_3	5.65	94.37
P_4	2.28	96.65
P_5	1.4	98.05
P_6	0.72	98.77
P_7	0.29	99.06

表2 Fisher, KNN 和 PLS-DA 识别模型对预测集样本识别结果

Tab.2 Recognition results of samples in prediction set using discriminant models of Fisher, KNN and PLS-DA

识别模型	100 N 挤压苹果(80个)		125 N 挤压苹果(80个)		完好无损苹果(80个)		总正确率/%
	识别数	正确率/%	识别数	正确率/%	识别数	正确率/%	
SD+Fisher	80	100	80	100	80	100	100
SD+KNN	73	91.3	74	92.5	80	100	94.6
SD+PLS-DA	80	100	80	100	80	100	100
SNV+Fisher	78	97.5	79	98.8	80	100	98.8
SNV+KNN	74	92.5	78	97.5	80	100	96.7
SNV+PLS-DA	79	98.8	80	100	80	100	99.6
MSC+Fisher	78	97.5	80	100	80	100	99.2
MSC+KNN	75	93.8	78	97.5	80	100	97.1
MSC+PLS-DA	79	98.8	80	100	80	100	99.6

3 结语

文中以“红富士”苹果为研究对象, 利用光谱采集系统收集了完好无损和挤压损伤苹果的光谱数据, 通过比较不同光谱预处理方法(SD, SNV 和 MSC)对各模型识别效果的影响, 得出SD对光谱反射率预处理的效果最好。采用主成分分析法选择了累计贡献率超过99%的前7个主成分(P_1 — P_7)作为样本集特征光谱数据, 很好地实现了光谱数据的降维(从全波段1024个波长范围内选择了前7个主成分作为新变量)。基于紫外/可见光谱技术结合模式识别算法建立的3种判别模型均能基本满足实际要求, 且SD+Fisher和SD+PLS-DA识别模型的识别性能最优, 对校正集和预测集样本的总正确识别率均高达100%。结果表明, 采用光谱采集系统获取的紫外/可见光谱信息结合模式识别算法可以实现苹果的轻微挤压损伤检测, 这为开发损伤水果在线快速检测系统

2.3 识别模型的建立与预测分析

基于光谱预处理以及主成分分析后的光谱数据, 建立了不同的PLS-DA, Fisher, KNN识别模型, 分别对完好无损和不同程度挤压损伤苹果样本进行识别检测, 识别结果见表2。从表2可以看出, KNN判别分析中MSC+KNN模型对预测集样本的总正确识别率为97.1%, 略好于SNV+KNN模型(96.7%)和SD+KNN模型(94.6%)。Fisher和PLS-DA判别分析中SD+Fisher和SD+PLS-DA模型的识别性能最优, 对校正集和预测集样本的总正确识别率均高达100%; PLS-DA判别分析中, SNV+PLS-DA和MSC+PLS-DA模型对预测集样本的识别效果也非常好, 总体正确识别率均为99.6%。总体而言, 3种识别模型均可用于识别不同程度的损伤苹果和无损苹果, SD+Fisher和SD+PLS-DA模型识别效果最好, SD+KNN模型识别效果相对最差。

提供了基础。

参考文献:

- [1] 冯迪, 纪建伟, 张莉, 等. 苹果品质高光谱成像检测技术研究进展[J]. 食品工业科技, 2017, 38(10): 389—394.
FENG Di, JI Jian-wei, ZHANG Li, et al. Progress of Hyperspectral Imaging Detection Technology in Apple Quality[J]. Science and Technology of Food Industry, 2017, 38(10): 389—394.
- [2] 郭文川, 王铭海, 谷静思, 等. 近红外光谱结合极限学习机识别贮藏期的损伤猕猴桃[J]. 光学精密工程, 2013, 21(10): 2720—2727.
GUO Wen-chuan, WANG Ming-hai, GU Jing-si, et al. Identification of Bruised Kiwifruits During Storage by Near Infrared Spectroscopy and Extreme Learning Machine[J]. Optics and Precision Engineering, 2013, 21(10): 2720—2727.

- [3] 迟茜, 王转卫, 杨婷婷, 等. 基于近红外高光谱成像的猕猴桃早期隐性损伤识别[J]. 农业机械学报, 2015, 46(3): 235—241.
CHI Qian, WANG Zhuan-wei, YANG Ting-ting, et al. Recognition of Early Hidden Bruises on Kiwifruits Based on Near-infrared Hyperspectral Imaging Technology[J]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(3): 235—241.
- [4] PAN Lei-qian, ZHANG Qiang, ZHANG Wei, et al. Detection of Cold Injury in Peaches by Hyperspectral Reflectance Imaging and Artificial Neural Network[J]. Food Chemistry, 2016, 192: 134—141.
- [5] 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 & Technology, 2016, 111: 352—361.
- [6] ERKINBAEV C, HENDERSON K, PALIWAL J. Discrimination of Gluten-free Oats from Contaminants Using Near Infrared Hyperspectral Imaging Technique[J]. Food Control, 2017, 80: 197—203.
- [7] GUO Wen-chuan, ZHAO Fan, DONG Jin-lei. Nondestructive Measurement of Soluble Solids Content of Kiwifruits Using Near-infrared Hyperspectral Imaging[J]. Food Analytical Methods, 2016, 9(1): 38—47.
- [8] PU Hong-bin, LIU Dan, WANG Lu, et al. Soluble Solids Content and pH Prediction and Maturity Discrimination of Lychee Fruits Using Visible and Near Infrared Hyperspectral Imaging[J]. Food Analytical Methods, 2016, 9: 235—244.
- [9] FAN Shu-xiang, ZHANG Bao-hua, LI Jiang-bo, et al. Prediction of Soluble Solids Content of Apple Using the Combination of Spectra and Textural Features of Hyperspectral Reflectance Imaging Data[J]. Postharvest Biology & Technology, 2016, 121: 51—61.
- [10] 何嘉琳, 乔春燕, 李冬冬, 等. 可见-近红外高光谱成像技术对灵武长枣 Vc 含量的无损检测方法[J]. 食品科学, 2018, 39(6): 194—199.
HE Jia-lin, QIAO Chun-yan, LI Dong-dong, et al. Non-destructive Detection of Vitamin c Content in "Lingwu Changzao" Jujubes (*Zizyphus jujuba* Mill cv Lingwu Changzao) Using Visible Near Infrared Hyperspectral Imaging[J]. Food Science, 2018, 39(6): 194—199.
- [11] KERESZTES J C, GOODARZI M, SAEYS W. Real-time Pixel Based Early Apple Bruise Detection Using Short Wave Infrared Hyperspectral Imaging in Combination with Calibration and Glare Correction Techniques[J]. Food Control, 2016, 66: 215—226.
- [12] TAN W, SUN L, YANG F, et al. The Feasibility of Early Detection and Grading of Apple Bruises Using Hyperspectral Imaging[J]. Journal of Chemometrics, 2018, e3067: 1—14.
- [13] PIECZYWEK P M, CYBULSKA J, SZYMAŃSKA-CHARGOT M, et al. Early Detection of Fungal Infection of Stored Apple Fruit with Optical Sensors—comparison of Biospeckle, Hyperspectral Imaging and Chlorophyll Fluorescence[J]. Food Control, 2018, 85: 327—338.
- [14] LEFCOUT A M, KIM M S, CHEN Y, et al. Systematic Approach for Using Hyperspectral Imaging Data to Develop Multispectral Imaging Systems: Detection of Feces on Apples[J]. Computers and Electronics in Agriculture, 2006, 54: 22—35.
- [15] 赵杰文, 刘剑华, 陈全胜. 利用高光谱图像技术检测水果轻微损伤[J]. 农业机械学报, 2008, 39(1): 106—109.
ZHAO Jie-wen, LIU Jian-hua, CHEN Quan-sheng, et al. Detecting Subtle Bruises on Fruits with Hyperspectral Imaging[J]. Transactions of the Chinese Society for Agricultural Machinery, 2008, 39(1): 106—109.
- [16] 程国首, 肉孜·阿木提, 郭先俊, 等. 基于高光谱图像的新疆红富士苹果颜色分级研究[J]. 新疆农业科学, 2012, 49(9): 1616—1623.
CHENG Guo-shou, ROUZI Atimu, GUO Xian-jun, et al. Study on Color Sorting for Xinjiang Fuji Apples by Using Hyperspectral Imaging Technology[J]. Xinjiang Agricultural Sciences, 2012, 49(9): 1616—1623.
- [17] ZHU Qi-bing, GUAN Ji-yu, HUANG Min, et al. Predicting Bruise Susceptibility of 'Golden Delicious' Apples Using Hyperspectral Scattering Technique[J]. Postharvest Biology and Technology, 2016, 114: 86—94.
- [18] 刘晶晶, 刘付龙, 史铁, 等. 高光谱图像技术检测苹果外部损伤[J]. 中国食品学报, 2018, 18(1): 278—284.
LIU Jing-jing, LIU Fu-long, SHI Tie, et al. Detection of External Damage of Apple by Hyperspectral Image Technique[J]. Journal of Chinese Institute of Food Science and Technology, 2018, 18(1): 278—284.
- [19] 倪力军, 张立国. 基础化学计量学及其应用[M]. 上海: 华东理工大学出版社, 2011: 64—69.
NI Li-jun, ZHANG Li-guo. Basic Chemometrics and Its Application[J]. Shanghai: East China University of Science and Technology Press, 2011: 64—69.