

基于近红外光谱和化学计量学的李果实成熟度鉴别方法研究

牛晓颖¹, 贡东军¹, 王艳伟¹, 陆文卿¹, 梁贺¹, 赵志磊¹, 任瑞²

(1. 河北大学质量技术监督学院, 河北保定 071002)

(2. 河北省林果桑花质量监督检验管理中心, 河北石家庄 050081)

摘要: 采收成熟度是影响李果实贮藏质量的重要因素, 为实现快速无损判别李果实的成熟度, 本文根据开花后发育时间的不同, 将 163 个李果实样品分为早期 (n=53)、中期 (n=55) 和晚期 (n=55) 三个成熟度, 利用近红外光谱分析技术, 对不同成熟度的李果实进行了分类。通过对马氏距离判别法、聚类独立软模式分类法、最小二乘-支持向量机的分类模型结果进行比较, 发现原始光谱前 20 个主成分得分作为输入时的马氏距离判别法模型结果最优, 校正集和预测集判别正确率分别为 96.33% 和 96.30%。对不同成熟度样品的可溶性固形物、可滴定酸及坚实度进行单因素方差分析发现, 各指标均存在显著差异, 坚实度差异最大。提取品质指标数据的主成分发现, 其聚类趋势与光谱主成分聚类趋势相似。结果表明, 使用近红外光谱技术结合化学计量学方法对李果实成熟度进行鉴别是有效的、可行的, 且其品质指标含量的差异可作为近红外光谱分类结果的理化验证。

关键词: 李果实; 成熟度; 近红外光谱; 马氏距离判别法; 主成分分析; 方差分析

文章编号: 1673-9078(2014)12-230-234

DOI: 10.13982/j.mfst.1673-9078.2014.12.039

Determination of Plum Maturity Levels by Using NIR and Chemometrics

NIU Xiao-ying¹, GONG Dong-jun¹, WANG Yan-wei¹, LU Wen-qing¹, LIANG He¹, ZHAO Zhi-lei¹, REN Rui²

(1. College of Quality and Technical Supervision, Hebei University, Baoding 071002, China)

(2. Center for Quality Supervision and Test of forestry fruit, mulberry and flower, Shijazhuang 050081, China)

Abstract: Harvest maturity is an important factor that can affect plum quality during the storage period. In order to achieve rapid non-destructive determination of plum maturity level, in this study, 163 plum samples were classified into three maturity levels-early stage (n = 53), middle stage (n = 55), and late stage (n = 55)-according to different fruit development times after flowering. Near infrared spectroscopy (NIR) was used to classify plum samples into different maturity levels. The accuracies of the classification models established by Mahalanobis distances analysis, soft independent modeling of class analogy, and least squares-support vector machine were compared. Mahalanobis distance model with the first 20 principal components (PCs) extracted from original spectra as inputs yielded the best result, and the correct classification rates for calibration set and prediction set were 96.33% and 96.30%, respectively. Soluble solid content (SSC), titratable acid, and firmness were analyzed using single factor one-way analysis of variance (ANOVA). There were significant differences in each index, and the greatest difference was found in SSC. The PCs that were extracted from the data of the three quality indices had very similar clustering tendencies as those extracted from spectrum data. These results showed that combining NIR with chemometrics to discriminate plums of different maturity levels was feasible and effective, and the difference in quality indices could be a physicochemical validation for the classification results from NIR.

Key words: plum; maturity; near infrared spectroscopy; Mahalanobis distances analysis; principal component analysis; analysis of variance

收稿日期: 2014-07-03

基金项目: 国家自然科学基金资助项目 (31201430); 河北省自然科学基金项目 (编号: C2013201113); 公益性行业 (农业) 科研专项资助: (编号: 201303075); 河北省科技计划项目 (14225503D); 河北大学大学生创新创业训练项目 (2012059)

作者简介: 牛晓颖 (1980-), 女, 博士, 副教授, 研究方向: 农产品/食品无损检测技术

采收成熟度是影响李果实贮藏质量的重要因素, 适宜的采收成熟度对提高果实的耐贮性和贮藏后的商品价值至关重要^[1]。在果实采收环节, 确定果实成熟度的传统方法是根据经验和目测判断, 预测准确率较低^[2], 因此, 采收的果实往往成熟度不一。在水果采后贮藏、加工等环节, 这种不同成熟度果实相互混杂的情况是导致果实损失率高达 30.45% 的主要原因之

一^[3]。采收时对果实的成熟度进行鉴别和分级对降低贮藏损失、提升果实品质及提高采收的自动化程度均有重要意义。

水果的一些品质指标如坚实度、可溶性固形物、可滴定酸等与其成熟度密切相关,常用来作为判断水果成熟度的间接依据。但这些指标的理化检测均需要破坏果实的组织,不适合在采收时用于果实成熟度的鉴别,同时,由于操作繁琐,也不适用于大批量水果的检测。近红外光谱分析技术具有无损、快速的特点,已在水果成熟度的直接和间接检测中得到较多应用^[4-13]。其中,Pérez-Marín等^[4]使用可见-近红外光谱对李果实的可溶性固形物和坚实度进行了定量;Louw等^[5]使用傅立叶变换近红外光谱仪对李果实的总可溶性固形物、总酸、糖酸比、坚实度以及重量进行了定量;赵志磊等^[9]使用傅立叶变换近红外光谱对李果实坚实度指标进行了定量。而李永玉等^[12]和闫润等^[13]则使用近红外分析技术分别对小型西瓜和草莓的成熟度进行了定性研究。目前,未见使用近红外光谱对李果实成熟度进行鉴别和分类的研究文献。

本文研究了使用近红外光谱分析技术对李果实成熟度进行鉴别和分类的方法,使用主成分分析的方法分析了不同成熟度李果实样品的聚类趋势,比较了马氏距离判别法、簇类独立软模式分类法(Soft Independent Modeling of Class Analogy, SIMCA)及最小二乘-支持向量机(Least Squares-Support Vector Machine, LS-SVM)鉴别模型的结果。对不同成熟度李果实样品的可溶性固形物、可滴定酸和坚实度指标进行了方差分析和主成分分析,从理化的角度验证了李果实成熟度的近红外分类结果。

1 材料与方法

1.1 实验材料

选用163个晚熟品种“安哥诺”李果实作为样品。根据样品采摘时的开花后发育时间长短,将其分为早期、中期和晚期三个成熟度,其中早期样品53个,中期样品55个,晚期样品55个。实验当天从河北省保定市易县独乐村李果园中采摘样品,进行简单清洁并自然风干后即进行近红外光谱的采集。

1.2 光谱采集与建模方法

使用MPA近红外光谱仪(Bruker,德国)进行了李果实样品的近红外漫反射光谱采集。光谱采集的硬件和各项参数为:20 W钨卤光源,Pbs检测器,波段范围4000~8000 cm⁻¹,分辨率8 cm⁻¹,扫描次数32,

背景光谱为仪器内置的金背景,光谱采集软件为仪器自带的OPUS 6.0。在李果实样品赤道圆周每隔120°选取1点采集光谱,共采集3条光谱,以3条光谱的平均光谱作为样品的原始光谱进行后续分析。

使用TQ Analyst 6.0(Thermo Nicolet,美国)软件对李果实样品光谱数据进行了主成分提取,并建立和优化了马氏距离判别法和SIMCA的分类模型,在MATLAB 7.0软件中编制程序,结合LS-SVM 1.5工具箱(Suykens Leuven,比利时)完成了LS-SVM分类模型的建立和各项参数寻优。

1.3 品质指标理化检测与分析方法

光谱采集完毕后,使用酸碱滴定法测定样品的可滴定酸含量,具体步骤如下:取5 g样品研碎后离心,以得到的上清液作为待测溶液,并以酚酞为指示剂,用0.1 mol/L的NaOH溶液滴定至溶液呈微红色;使用PAL-1数显手持式折光计(Atago,日本)测定果实的可溶性固形物含量;使用FHM-5型果实硬度计(Takemura Electric Works.ltd,日本)测定果实坚实度。以上检测点位置与光谱采集位置一致。使用SPSS 16.0(SPSS Inc.,美国)对不同成熟度李果实样品的可溶性固形物、可滴定酸及坚实度指标进行方差分析(Analysis of Variance, ANOVA)。

2 结果与讨论

2.1 光谱测量结果

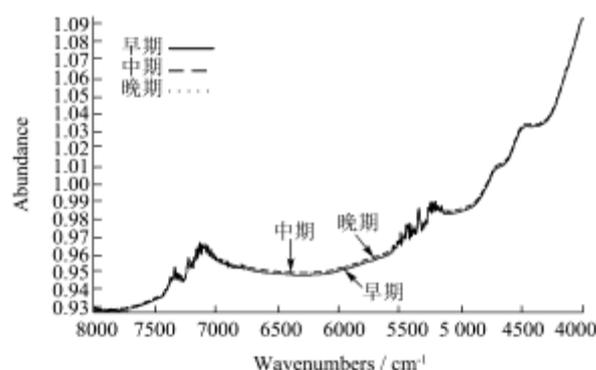


图1 不同成熟度李果实样品平均光谱

Fig. 1 Average spectra for plum samples of different maturity levels

早期、中期和晚期三个成熟度李果实样品的近红外漫反射平均光谱如图1。从图1中可以看出,三个成熟度的样品光谱曲线非常相似,但某些波段吸光度的大小略有差异,中期和晚期的样品光谱基本重合在一起,而早期样品光谱的吸光度比中晚期样品吸光度略低。随着李果实成熟程度的加深,其中的糖、酸等

化学指标含量以及坚实度等物理指标均发生了变化，从而在其近红外光谱的吸收上出现了差异，这是使用近红外光谱鉴别不同成熟度水果的理论基础。将不同成熟度的李果实样品按照大概为 2:1 的比例随机分为校正集和预测集，各样品集的具体数据见表 1。

表 1 不同成熟度李果实样品的校正集和预测集

Table 1 Calibration and prediction sets for plum samples of different maturity levels

样品类型	样品数	校正集样品数	预测集样品数
早期	53	35	18
中期	55	37	18
晚期	55	37	18
总计	163	109	54

2.2 主成分分析

为将高维的光谱数据降维以方便下文分类模型的建立和优化，提取早期、中期和晚期三个成熟度李果实样品光谱数据的主成分。前 20 个主成分的累积贡献率达到 95.974%，表达了原始光谱绝大部分的差异特征，这说明使用降维后的主成分得分作为分类模型的输入变量是可靠的。前 3 个主成分对光谱差异解释的累积贡献率为 93.085%，其三维图可以较为可靠地表现出不同成熟度李果实样品光谱的原始聚类趋势，见图 2。从图 2 中可以看出，早期、中期和晚期李果实样品有各自聚类的趋势，早期和晚期的样品可以完全分开，但中期与早期和晚期样品均有较为严重的重叠，这有可能是因为，李果实的成熟是一个渐变的过程，相应的各种成分变化也是在逐渐的积累中产生，而中期样品的成熟程度处于早期和晚期之间，某些中期样品的成熟度可能比较接近早期或晚期，具有相似的光谱特征，从而产生了重叠，而早期和晚期样品的成熟程度相差较大，各种成分含量的较大差异使得其近红

外吸收光谱的特征差异较大，从而在主成分得分图上呈现出完全分开的状态。

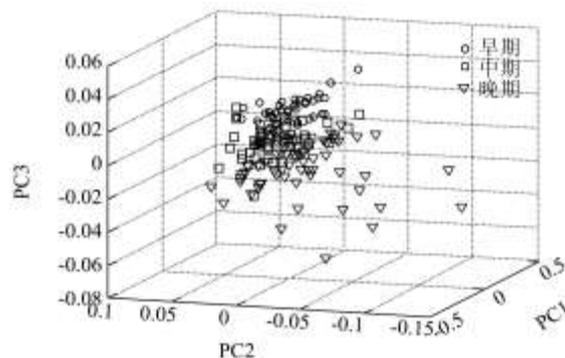


图 2 不同成熟度李果实样品光谱前 3 个主成分得分图

Fig.2 Plot of first three principal component scores for plum sample spectra of different maturity levels

2.3 分类模型的建立和优化

为得到较佳的算法组合，本文比较了马氏距离判别法、SIMCA 和 LS-SVM 三种算法所建立分类模型的结果。使用高斯核函数作为 LS-SVM 算法的核函数将输入变量的特征映射到高维空间以进行分类，建模过程中使用二次网格搜索法 (Two Step Grid Searching) 以及十等分交叉验证 (Robust 10-fold Cross Validation) 对回归误差权重 γ 和核参数 σ^2 进行了全局寻优，两个参数对于控制模型的复杂度和逼近误差以及模型的测量精度有重要影响^[4]。

分别使用前 1~20 个主成分得分作为三种算法分类模型的输入变量，发现马氏距离判别法、SIMCA 和 LS-SVM 模型分别在输入前 20、前 8 以及前 9 个主成分得分时得到的模型结果较优。分别对各模型进行平滑 (5 点、15 点及 25 点)、多元散射校正以及标准归一化的光谱预处理，发现未对各模型分类结果有改善，原始光谱提取主成分得分作为输入的模型仍较优。

表 2 不同建模算法对不同成熟度李果实样品的分类结果

Table 2 Classification results for plum samples of different maturity levels by different modeling algorithms

建模方法	校正集		预测集		主成分数
	正确率/%	错误样本数	正确率/%	错误样本数	
马氏距离判别	96.33	4	96.30%	2	20
SIMCA	94.50	6	87.04%	7	8
LS-SVM	100.00	0	83.33%	9	11

表 2 为三种不同建模算法对不同成熟度李果实样品的较优分类结果。从表 2 中知，马氏距离判别法方法所建立的分类模型结果最优，校正集有 4 个样品产生误判，判别正确率为 96.33%，预测集有 2 个样品产生误判，判别正确率为 96.30%。早期、中期和晚期样品的单独判别正确率分别为 94.34%、96.36% 以及

98.18%。其中，早期样品 3 个误判为中期；中期样品 2 个误判为早期；晚期样品 1 个误判为中期。图 3 为不同成熟度李果实样品分类的三维马氏距离图。经过马氏距离判别法分类，早期、中期和晚期样品的区分更加明确，光谱主成分得分图 (图 2) 中中期样品与早期和晚期样品的重叠在这里得到了较大的改善，个

别样品的重叠可能是由于样品的特征较为相似所致。

2.4 品质指标差异分析

分别对早期、中期和晚期三个不同成熟度李果实样品中的可溶性固形物、可滴定酸及坚实度数据进行单因素方差分析。分析结果及各品质指标的平均值/标准偏差见表 3。从表 3 中可以看出,不同成熟度样品的三个指标均具有显著差异,比较三指标的 F 值可以发现,坚实度指标的 F 值远大于其他两指标,说明不同成熟度李果实样品的坚实度差异比其可溶性固形物和可滴定酸含量差异大,相比其他两指标,使用坚实度指标可以更容易地区分李果实样品的成熟度。

表 3 不同成熟度李果实样品可溶性固形物、可滴定酸及坚实度的平均值/标准偏差、 F 值及差异显著性

Table 3 Mean/standard deviation, F value, and significance of differences for soluble solid content, titratable acid, and firmness of plum samples of different maturity levels

指标	早期	中期	晚期	F	差异显著性 $p < 0.05$
可溶性固形物/%	7.77/1.20	7.19/1.26	7.01/1.48	4.86	显著
可滴定酸/%	10.03/0.89	9.51/0.98	9.43/0.79	7.09	显著
坚实度/(kg/cm ²)	1.56/0.29	1.56/0.32	1.00/0.27	66.72	显著

提取不同成熟度李果实样品三个品质指标数据的主成分,前 3 个主成分的累积贡献率达到 100%,完全解释了原数据的差异特征。图 4 为三个成熟度李果实样品两两组合的前 3 个主成分得分图,从图 4 中可以看出,早期与晚期、中期与晚期样品聚类趋势明显,彼此区分较好,但早期与中期重叠严重,说明早期和中期样品的三品质指标特征较为相似,而早期与晚期及中期与晚期的指标特征差异较大。这样的聚类趋势与光谱主成分的聚类趋势(图 2)极为相似,从三个主要品质指标的角度验证了不同成熟度李果实样品的近红外分类结果。同时,李果实作为一种农产品,其成分极为复杂,近红外光谱所表达的样品特征信息较为全面,而图 4 仅表达了李果实样品在三个主要品质指标方面的特征信息,因此,其区分效果没有近红外光谱的分类结果好。

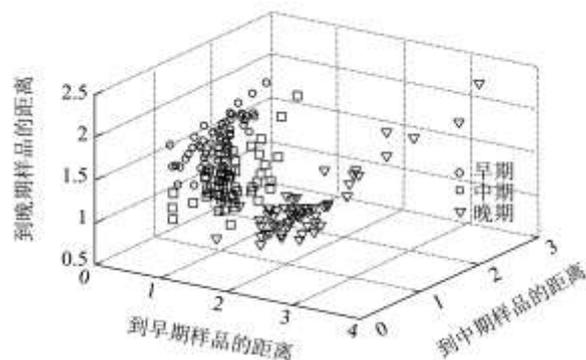


图 3 不同成熟度李果实样品分类的三维马氏距离图

Fig.3 Dimensional plots of Mahalanobis distances for classification of plum samples of different maturity levels

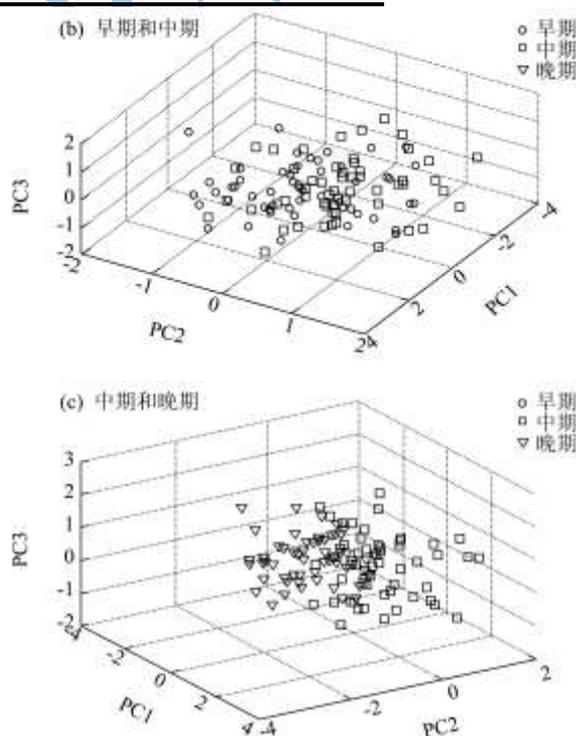


图 4 (a) 早期与晚期、(b) 早期与中期、(c) 中期与晚期李果实样品可溶性固形物、可滴定酸及坚实度数据前 3 主成分得分图
Fig.4 First three principal component scores for soluble solid content, titratable acid, and firmness of plum samples of different maturity levels (a) early and late stage, (b) early and middle stage, and (c) middle and late stage

3 结论

本文对早期、中期和晚期三个成熟度的李果实样品进行了近红外分类模型的建立和优化。对不同成熟度李果实样品的近红外漫反射光谱提取了主成分并进行分析发现,不同成熟度样品有聚类趋势,且成熟度相差越大区分越明显。使用马氏距离判别法、SIMCA和LS-SVM三种算法建立了分类模型,并比较了前1-20主成分得分的不同输入变量及平滑(5点、15点及25点)、多元散射校正和标准归一化的不同光谱预处理方法对模型分类结果的影响发现,原始光谱前20主成分得分作为输入时的马氏距离判别法模型分类结果最优,校正集和预测集判别正确率分别为96.33%和96.30%。分别对早期、中期和晚期三个不同成熟度李果实样品中的可溶性固形物、可滴定酸及坚实度数据进行单因素方差分析发现,各指标均存在显著差异,且坚实度指标差异最大。对可溶性固形物、可滴定酸及坚实度数据提取主成分并进行分析发现,不同成熟度样品有聚类趋势,且其聚类趋势与光谱主成分极为相似,从理化指标的角度验证了不同成熟度李果实样品的近红外分类结果。本文研究表明,使用近红外光谱分析技术结合化学计量学方法对李果实成熟度进行鉴别是有效的、可行的。

参考文献

- [1] 及华,关军锋,孙玉龙,等.不同采收成熟度黑宝石李冷藏期间品质变化的研究[J].保鲜与加工,2010,10(3):22-25
JI Hua, GUAN Jun-feng, SUN Yu-long, et al. Research of quality change for black diamond plums of different maturity during cold storage [J]. Storage and Process, 2010, 10(3): 22-25
- [2] 毛莎莎,曾明,何绍兰,等.近红外光谱技术在水果成熟期预测中的应用[J].亚热带植物科学,2010,39(1):82-87
MAO Sha-sha, ZENG Ming, HE Shao-lan, et al. Application of near infrared spectra (NIRS) technology in prediction of maturity stage of fruit [J]. Subtropical Plant Science, 2010, 39(1): 82-87
- [3] 贺艳楠,魏永胜,郑颖.水果成熟度无损检测技术研究进展[J].北方园艺,2010,(3):208-212
HE Yan-nan, WEI Yong-sheng, ZHENG Ying. Research advanced of nondestructive detection of maturity in fruits [J]. Northern Horticulture, 2010, 39(1): 82-87
- [4] Pérez-Marín D, Paz P, José-Emilio G, et al. Miniature handheld NIR sensor for the on-site non-destructive assessment of post-harvest quality and refrigerated storage behavior in plums [J]. Journal of Food Engineering, 2010, 99(3): 294-302
- [5] Louw D E, Theron I K. Robust prediction models for quality parameters in Japanese plums (*Prunus Salicina* L.) using NIR spectroscopy [J]. Postharvest Biology and Technology, 2010, 58(3): 176-184
- [6] Bobelyn E, Serban A, Nicu M, et al. Postharvest quality of apple predicted by nir spectroscopy: study of the effect of biological variability on spectra and model performance [J]. Postharvest Biology and Technology, 2010, 55(2): 133-143
- [7] Jha S N, Ruchi G. Non-destructive prediction of quality of intact apple using near infrared spectroscopy [J]. Journal of Food Science and Technology, 2010, 47(2): 207-213
- [8] Sánchez M, Habaa M, Guerrero J, et al. Testing of a local approach for the prediction of quality parameters in intact nectarines using a portable NIRS instrument [J]. Postharvest Biology and Technology, 2011, 60(2): 130-135
- [9] 赵志磊,代旭静,郝清,等.近红外光谱无损检测李果实硬度的研究[J].安徽农业科学,2011,39(24):14552-14554
ZHAO Zhi-lei, DAI Xu-jing, HAO Qing, et al. Research on the non-destructive test of plum hardness with the method of near-infrared spectroscopy [J]. Journal of Anhui Agricultural Sciences, 39(24): 14552-14554
- [10] Liu Y, Gao R, Sun X, et al. Improvement of near-infrared spectral calibration models for brix prediction in 'Gannan' navel oranges by a portable near-infrared device [J]. Food and Bioprocess Technology, 2012, 5(3): 1106-1112
- [11] Ruslan R, Ehsani R, Lee W S. Quantification of total soluble solids and titratable acidity for citrus maturity using portable Vis-NIR spectroradiometer [J]. Applied Engineering in Agriculture, 2012, 28(5): 735-743
- [12] 李永玉,赵洪卫,常冬,等.小型西瓜果实成熟度的无损定性判别[J].光谱学与光谱分析,2012,32(6):1526-1530
LI Yong-yu, ZHAO Hong-wei, CHANG Dong, et al. Maturity qualitative discrimination of small watermelon fruit [J]. Spectroscopy and Spectral Analysis, 2012, 32(6): 1526-1530
- [13] 闫润,王新忠,张莹莹.基于神经网络及近红外光谱的草莓成熟度快速识别方法[J].安徽农业科学,2012, 40(10): 6292- 6294
YAN Run, WANG Xin-zhong, ZHANG Ying-ying. Fast recognition method of strawberries' maturity level based on neural network and near infrared spectra [J]. Journal of Anhui Agricultural Sciences, 2012, 40(10): 6292-6294
- [14] 李晓丽,程术希,何勇.基于漫反射光谱的初制绿茶含水率

无损检测方法[J].农业工程学报,2010,26(5):195-201
LI Xiao-li, CHENG Shu-xi, HE Yong. Nondestructive
measurement of moisture content of green tea in primary

processing based on diffuse reflectance spectroscopy [J].
Transactions of the CSAE, 2010, 26(5): 195-201

现代食品科技