

基于高光谱成像的马铃薯环腐病无损检测

郭红艳¹, 刘贵珊^{1,*}, 吴龙国², 王松磊^{1,2}, 康宁波², 陈亚斌¹, 何建国^{1,2}, 贺晓光¹

(1.宁夏大学农学院, 宁夏 银川 750021; 2.宁夏大学土木水利工程学院, 宁夏 银川 750021)

摘要:为探讨高光谱成像技术无损检测马铃薯环腐病的可行性, 采用反射高光谱(980~1 650 nm)成像技术, 以120个马铃薯样本(合格60个, 环腐60个)为研究对象, 对比多元散射校正、标准正态变换、卷积+一阶导数等对建模的影响, 优选出多元散射校正的光谱预处理方法; 然后基于偏最小二乘回归系数法提取9个特征波长(993、1 005、1 009、1 031、1 112、1 162、1 165、1 225、1 636 nm), 建立特征波长下马铃薯环腐病的2类线性判别分析(linear discriminant analysis, LDA)模型和4类支持向量机(support vector machine, SVM)模型, 即Fisher-LDA、马氏距离-LDA、线性核SVM、径向基核SVM、多项式核SVM和S型核SVM。结果表明, LDA模型中马氏距离法最优, SVM模型中S型核SVM最优, LDA模型整体优于SVM模型, 最终确定基于马氏距离LDA的马铃薯环腐病判别模型为最佳模型, 校正集、验证集识别率分别为100%和93.33%。实验结果表明高光谱无损检测马铃薯环腐病具有可行性。

关键词:高光谱成像技术; 马铃薯; 环腐病; 无损检测

Hyper-Spectral Imaging Technology for Nondestructive Detection of Potato Ring Rot

GUO Hongyan¹, LIU Guishan^{1,*}, WU Longguo², WANG Songlei^{1,2}, KANG Ningbo², CHEN Yabin¹, HE Jianguo^{1,2}, HE Xiaoguang¹

(1. School of Agriculture, Ningxia University, Yinchuan 750021, China;

2. School of Civil Engineering and Water Conservancy, Ningxia University, Yinchuan 750021, China)

Abstract: To investigate the feasibility of using hyper-spectral imaging technique to detect potato ring rot, hyper-spectral imaging operated in reflectance mode in the wavelength range of 980–1 650 nm was applied to 120 potato samples (60 qualified and 60 ring rot). The effects of multiple scattering correction, standard normal transformation and savitzky-golay + first derivative on model performance were compared. Multiple scattering correction was chosen as the best spectral pre-processing method. Then, 9 characteristic wavelengths (993, 1 005, 1 009, 1 031, 1 112, 1 162, 1 165, 1 225, and 1 636 nm) were extracted based on the partial least squares method. Two linear discriminant analysis (LDA) models, Fisher-LDA and Mahalanobis distance-LDA, and four support vector machine (SVM) models, linear kernel SVMs, SVM with radial basis kernel, polynomial kernel SVM and SVM with Sigmoid kernel, were built for ring rot potato at characteristic wavelengths. The results showed that Mahalanobis distance-LDA was better than Fisher-LDA while Sigmoid kernel performed best among all the SVM models. The LDA models were overall better than the SVM models. Thus, the LDA model of potato ring rot based on Mahalanobis distance was the best model. Its recognition rate was 100% for calibration set and 93.33% for validation set. This study indicated that hyper-spectral imaging technology can be used to identify potato ring rot.

Key words: hyper-spectral imaging; potato; ring rot; non-destructive detection

DOI:10.7506/spkx1002-6630-201612036

中图分类号: S532; TP391.4

文献标志码: A

文章编号: 1002-6630(2016)12-0203-05

引文格式:

郭红艳, 刘贵珊, 吴龙国, 等. 基于高光谱成像的马铃薯环腐病无损检测[J]. 食品科学, 2016, 37(12): 203-207.

DOI:10.7506/spkx1002-6630-201612036. <http://www.spkx.net.cn>

GUO Hongyan, LIU Guishan, WU Longguo, et al. Hyper-spectral imaging technology for nondestructive detection of potato ring rot[J]. Food Science, 2016, 37(12): 203-207. (in Chinese with English abstract) DOI:10.7506/spkx1002-6630-201612036. <http://www.spkx.net.cn>

收稿日期: 2015-09-29

基金项目: 2013年度宁夏自然科学基金项目(NZ13005)

作者简介: 郭红艳(1991—), 女, 硕士研究生, 研究方向为农产品无损检测。E-mail: 1525490213@qq.com

*通信作者: 刘贵珊(1979—), 男, 副教授, 博士, 研究方向为农产品无损检测和肉制品加工。E-mail: liugs@nxu.edu.cn

我国已成为世界上马铃薯生产和消费的第一大国。2015年,第一代马铃薯馒头在北京成功上市,马铃薯主食产业化的不断推进使马铃薯有望成为与三大主粮协调发展的第四大主粮,其产业发展前景良好,经济效益、营养效益、生态效益、社会效益可观,对保障国家粮食安全、加快农业发展方式转变具有重要支撑作用,马铃薯的产业发展也受到世界各国的高度关注^[1]。

马铃薯环腐病是马铃薯生产上一种重要的细菌性病害,具有高度的传染性,常造成马铃薯的烂种、死苗和死株,一般减产10%,严重者减产达60%以上^[2]。目前的检测方法为人工目测症状检测或实验室生化检测,或效率低,或操作复杂,因此对马铃薯环腐病进行快速无损检测进行研究具有重要意义。

高光谱成像技术融合了传统的数字成像技术和光谱分析技术,不仅可以获取被测物的空间图像信息,还能获取每个像素的光谱信息^[3]。在农产品研究领域,高光谱可以进行定性分析或定量检测,目前国内外学者的研究报道涵盖了对尺寸、表面缺陷损伤、排泄物污染、农药残留等外部指标的分析,也包括对糖度、水分、固形物、硬度等内部指标的检测。苏文浩等^[4]运用高光谱成像技术检测了马铃薯的机械损伤、孔洞、疮痂、表面碰伤等问题;周竹^[5]、韩广涛^[6]等运用高光谱成像技术研究了马铃薯黑心病;Polder等^[7]运用高光谱成像技术判别分析了西红柿的成熟度;Sun Dawen等^[8]运用高光谱反射成像技术检测了鲈鱼的新鲜度(6 d内);有研究^[9-17]分别利用高光谱图像技术检测柑橘果锈、脐橙溃疡、水果硬度、牛肉嫩度、茶叶质量等级、柑橘果锈、咖啡豆品种及脐橙糖度。然而,目前利用高光谱成像技术检测马铃薯环腐病的研究鲜见报道。

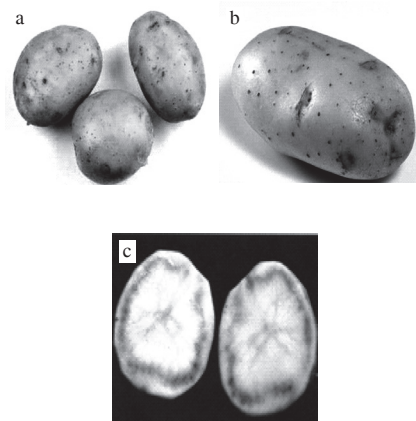
针对马铃薯环腐病的高光谱无损检测研究尚少的现状,利用近红外高光谱成像系统(900~1 700 nm)采集马铃薯内部环腐病与正常马铃薯的高光谱图像,选择感兴趣区域,提取平均光谱,优化光谱预处理方法,采用偏最小二乘回归系数(partial least squares regression, PLSR)法提取特征波长,对比分析特征波长下建立的线性判别分析(linear discriminant analysis, LDA)模型和支持向量机(support vector machine, SVM)模型,优选模型,实现马铃薯内部环腐病的无损检测,以期的高光谱成像技术应用于马铃薯品质的在线分级分选提供参考。

1 材料与方法

1.1 材料

实验所用马铃薯样本为宁薯六号,产于宁夏吴忠市红寺堡区。按马铃薯等级规格选取120个马铃薯作为样本,所有样本购回后做简单清理,室温条件下放置24 h,

使其自身温度与室温相同。切开得到60个环腐马铃薯,其余60个正常马铃薯作为对照组,如图1所示。



a.完好马铃薯; b.环腐马铃薯; c.环腐马铃薯剖面。

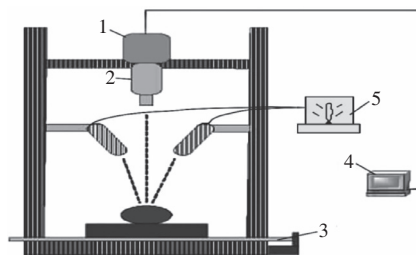
图1 各类马铃薯样本

Fig. 1 Different potato samples

1.2 仪器与设备

N17E-NIR近红外高光谱成像系统、4个35 W的HSIA-LS-TDIF卤钨灯线光源、PSA200-11-X推扫式输送装置 北京卓立汉光仪器有限公司; Inspector N17E近红外成像光谱仪 芬兰奥卢光谱成像有限公司; Zelos-285GV系列CCD相机 德国Kappa Opto-electronics GmbH公司; 采集暗箱和高性能的计算机 联想集团有限公司。

高光谱成像系统(图2)可采集的波长范围为900~1 700 nm,光谱分辨率为5 nm,共有256个波长,像素为320×300。



1.成像光谱仪; 2. CCD相机; 3.载物平台; 4.光源; 5.配套软件的计算机。

图2 高光谱成像系统

Fig. 2 Hyper-spectral imaging system

1.3 方法

1.3.1 成像系统采集参数的确定

为避免环境光线和噪音的干扰,整个图像的扫描过程在采集暗箱中进行。预先根据光源的照度设定好摄像头曝光时间,以保证图像清晰;并调整好输送装置的速度,以避免图像空间分辨率失真。采集图像时,线阵探测器作垂直于光学焦平面的横向移动,从而获取所扫描空间中每个像素在整个光谱区域的图像信息;同时,样

本在输送装置的作用下作垂直于摄像机的纵向移动, 最终完成整个样本图像的采集。预实验进行参数的调整和优化, 确定曝光时间为10 ms, 图像采集速率为13 mm/s, 物距为2 cm。

采集样本图像前, 为了减弱或消除CCD相机暗电流和光源不均匀亮度的产生的影响, 需要进行黑白校正^[18], 校准图像(T)的计算如公式(1)所示:

$$T = \frac{T_0 - T_d}{T_w - T_d} \quad (1)$$

式中: T_0 为样本的原始漫反射高光谱图像; T_d 为全黑图像; T_w 为白板图像; T 为校正后的漫反射高光谱图像。

1.3.2 高光谱图像数据的采集

将黑色背景板置于载物台上, 每次取1个马铃薯样本规整放于该板上, 采集样本的高光谱图像。为减少背景图像中多余信息的影响, 以马铃薯样本的实际尺寸为参照, 把各个样本图像的像素尺寸由320×300调整到200×160。然后, 利用ENVI 4.6软件选择合适的感兴趣区域, 提取平均光谱值。

1.3.3 数据处理方法

PLSR融合了多种多元统计数据分析方法的优点, 是一种新型的分析方法, 有效地解决了变量相关度高的问题, 并且比较适用于处理变量多而样本数少的问题^[19]。

SVM^[20]通过引入核函数把基于内积运算的线性算法非线性化, 即将样本非线性映射到新的高维空间, 在高维空间中进行相应的线性操作^[21]。根据核函数类型可以分为: 线性核函数(linear kernel, LK)、多项式核函数(polynomial kernel, PK)、径向基核函数(radial basis function, RBF)、S型核函数(sigmoid kernel, S)等^[22]。

LDA通过特征提取把同类样本聚集在一起, 不同类样本分开^[23]。按照判别函数的判别准则, LDA分为Fisher判别分析、距离判别分析、Bayes判别分析^[24]等。经典LDA使用的是Fisher LDA^[25], 而马氏距离是距离判别分析中常用的一种判别函数^[25]。

实验中应用到的数据软件为ENVI 4.6 (Research System Inc., Boulder, Colo, USA)、MATLABR2011a (The Mathworks Inc., Natick, MA, USA) 以及Unscrambler X 10.2 (CAMO Software AS, OSLO, Norway)。

1.3.4 模型评价指标

识别率是评价模型的指标, 正确识别率越高, 表示该模型性能越好, 识别马铃薯环腐病的能力越强。环腐病马铃薯的正确识别率的计算如公式(2)所示:

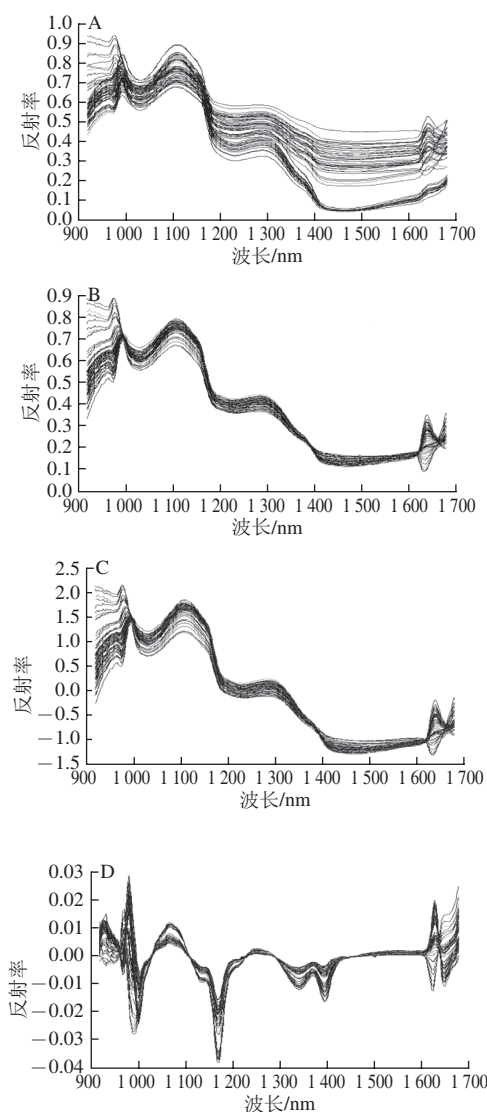
$$a/\% = \frac{b}{c} \times 100 \quad (2)$$

式中: a 为环腐样本的正确识别率/%; b 为正确识别环腐样本的数目; c 为被识别样本的总数目。

2 结果与分析

2.1 光谱预处理

为了使样本的光谱曲线更具代表性, 分别计算120个马铃薯样本表面10×10个像素点下的平均光谱值, 作为相应样本的原始光谱。由于原始光谱数据受样品表面不均匀性、形态多样性、仪器噪音和暗电流等因素的影响, 使光谱曲线产生不重复和基线漂移的现象, 因此必须对原始光谱进行预处理。本实验选用多元散射校正(multiplicative scatter correction, MSC)、标准正态变换(standard normal variate, SNV)、卷积+一阶导数(savitzky-golay+first derivative, S-G+FD)等预处理方法处理原始光谱, 原始光谱和不同预处理方法的处理效果, 如图3所示。



A. 原始光谱; B. MSC; C. SNV; D. S-G+FD。

图3 原始光谱和预处理后的光谱曲线图

Fig. 3 Original and pretreated spectra

从图3可以看出, 980 nm以下和1 650 nm以上的波长部分噪声明显, 使该波长的数据失去分析的价值。因此, 选取980~1 650 nm的光谱区域进行后续分析。

在980~1 650 nm波长范围内, 以2/3样本为校正集、1/3样本为验证集, 对原始光谱和不同预处理光谱, 分别建立基于马氏距离的马铃薯环腐病LDA判别模型, 结果如表1所示。

表 1 原始光谱与预处理光谱的LDA模型比较
Table 1 Comparison of the LDA models established with original and pretreated spectra

函数类型	类别	正确识别率/%	
		校正集	验证集
LDA	原始光谱	92.78	93.33
	MSC	97.78	95.36
	SNV	96.89	94.53
	S-G+FD	85.35	82.36

不同预处理光谱在同一建模方法下建模, 模型的正确识别率越高, 则证明光谱预处理方法越好, 对环腐病马铃薯的判别越准确。由表1可知, 基于MSC的LDA模型对样本校正集、验证集的正确识别最高, 分别达到97.78%和95.36%, 判别效果最好; 与原始光谱相比较, 基于S-G+FD的LDA模型性能反而下降, 校正集和验证集的正确识别率均降低, 这可能是由于在卷积求导的过程中噪音等无用信息无形中被放大的结果。因此, 选取MSC法为光谱预处理方法。

2.2 特征波长的提取

对980~1 650 nm有效波长内的全部高光谱数据进行分析处理不仅耗时, 而且计算量大^[26]。此外, 这些数据本身具有一定的冗余性, 部分信噪比低、噪音大的波长会对模型的识别性能产生影响, 因此, 必须对光谱数据进行降维。实验对MSC预处理的光谱进行PLS回归, 基于局部最大值和最小值的原则, 在980~1 650 nm波长范围内选取了9个理想可用特征波长, 分别为993、1 005、1 009、1 031、1 112、1 162、1 165、1 225、1 636 nm。

2.3 模型的建立与分析

将MSC预处理的9个特征波长分别输入SVM和LDA, 建立马铃薯环腐病的Fisher-LDA、马氏距离-LDA、LK-SVM、PK-SVM、RBF-SVM和S-SVM共6类判别模型, 结果如表2所示。

SVM的主要优势在于解决线性不可分问题。构造出性能良好的SVM, 核函数的选择是关键, 这主要包括核函数类型的选择和后续后相关参数的选择两部分工作。实验采用Cross-Validation方法, 分别试用不同的核函数, 所得误差最小的核函数, 即为最好的核函数。采用S函数时, SVM实现的是一种多层感知器神经网络, SVM的理论基础决定它最终求得的是全局最优值, 也保证了它对于未知

样本的良好泛化能力。表2显示, 在相同数据条件下识别环腐马铃薯时, 采用S型核函数的SVM要比采用其他核函数的SVM误差小, 其验证集正确识别率最高, 为91.33%。

表 2 SVM与LDA模型识别结果比较
Table 2 Comparison of recognition results between SVM and LDA models

函数类型	类别	正确识别率/%	
		校正集	验证集
SVM	LK	100	73.33
	PK	93.33	88.89
	RBF	100	73.33
	S	93.33	91.33
LDA	Fisher	100.0	91.78
	马氏距离	100.0	93.33

LDA适合于模式识别问题, 传统的LDA最大的缺点是难以找到足够多的训练样本以保证类内离散度矩阵可逆^[25]。马氏距离算法简单, 具有平移不变性、旋转不变性和仿射不变性^[27], 较好地解决了小样本问题。表2显示, LDA模型识别马铃薯环腐病的效果良好, 其中基于马氏距离的LDA算法识别率最高, 验证集正确识别率为93.33%。

综合分析, 基于LDA法的马铃薯环腐病判别模型的正确识别率整体优于SVM法, 分析SVM模型识别效果差的原因可能是受样本数量少的影响。因此, 6种模型中, 采用马氏距离-LDA法建立的马铃薯环腐病判别模型为最佳模型, 其校正集正确识别率为100%, 验证集正确识别率为93.33%。

3 结 论

采用高光谱成像技术在900~1 700 nm波长范围内, 对宁夏马铃薯环腐病进行了无损检测。全波长内, MSC、SNV、S-G+FD和原始光谱相比较, 优选出MSC的光谱预处理法, 预测集正确识别率为95.36%; PLSR可从有效波长中优选出9个特征波长(993、1 005、1 009、1 031、1 112、1 162、1 165、1 225、1 636 nm), 简化运算, 提高模型性能; 整体上, 特征波长下的LDA马铃薯环腐病判别模型较SVM马铃薯环腐病判别模型的识别性能更好; LDA的2类模型中, 马氏距离-LDA模型最优; SVM的4类模型中, S-SVM模型最优。最佳模型为基于马氏距离的LDA马铃薯环腐病判别模型, 其校正集识别率为100%, 验证集识别率为93.33%。反射高光谱成像技术作为一种快速、高效的无损检测技术对马铃薯环腐病进行检测具有可行性。

参考文献:

- [1] 周康才, 林玉清, 檀时龙, 等. 马铃薯发展前景分析及其优势利用[J]. 安徽农学通报, 2011, 17(17): 67-68. DOI:10.3969/j.issn.1007-7731.2011.17.033.

- [2] 韩学俭. 马铃薯环腐病的危害及其防治[J]. 蔬菜, 2003(2): 22-23. DOI:10.3969/j.issn.1001-8336.2003.02.018.
- [3] 马本学, 应义斌, 饶秀勤, 等. 高光谱成像在水果内部品质无损检测中的研究进展[J]. 光谱学与光谱分析, 2009, 29(6): 1611-1615. DOI:10.3964/j.issn.1000-0593(2009)06-1611-05.
- [4] 苏文浩, 刘贵珊, 何建国, 等. 近红外高光谱图像技术在马铃薯外部缺陷中的应用[J]. 食品与机械, 2013(5): 128-134.
- [5] 周竹, 李小昱, 高海龙, 等. 漫反射和投射光谱检测马铃薯黑心病的比较[J]. 农业工程学报, 2012, 28(11): 237-242.
- [6] 韩广涛, 杨志辉, 朱杰华, 等. 双重PCR技术检测马铃薯环腐病菌和黑胫病菌的方法建立[J]. 中国农业科学, 2011, 44(20): 4199-4206.
- [7] POLDER G, van der HEIJDEN G, YOUNG I T. Spectral image analysis for measuring ripeness of tomatoes[J]. Trans of ASAE, 2002, 45(4): 1155-1162.
- [8] SUN Dawen. Hyperspectral imaging for food quality analysis and control[M]. Elsevier: Academic Press of Elsevier, 2010.
- [9] 蔡健荣, 王建黑, 黄星奕, 等. 高光谱图像技术检测柑橘果锈[J]. 光电工程, 2009, 36(6): 26-30. DOI:1003-501X(2009)06-0026-05.
- [10] 李江波, 饶秀勤, 应义斌, 等. 基于高光谱成像技术检测脐橙溃疡[J]. 农业工程学报, 2010, 26(8): 222-228. DOI:10.3969/j.issn.1002-6819.2010.08.038.
- [11] JIAN Weiqin, THOMAS F B, MARK A R, et al. Detection of citrus canker using hyperspectral reflectance imaging with spectral information divergence[J]. Journal of Food Engineering, 2009, 93(2): 183-191.
- [12] LU R F, PENG Y K. Hyperspectral scattering for assessing peach fruit firmness[J]. Biosystems Engineering, 2006, 93(2): 161-171.
- [13] XIA J J, WEAVER A, GERRARD D E, et al. Heating induced optical property changes in beef muscle[J]. Journal of Food Engineering, 2008, 84(1): 75-81. DOI:10.1016/j.jfoodeng.2007.04.023.
- [14] 陈全胜, 赵杰文, 蔡健荣, 等. 利用高光谱图像技术评判茶叶的质量等级[J]. 光学学报, 2008, 28(4): 669-674. DOI:10.3321/j.issn:0253-2239.2008.04.011.
- [15] 蔡健荣, 王建黑, 陈全胜, 等. 波段比算法结合高光谱图像技术检测柑橘果锈[J]. 农业工程学报, 2009, 25(1): 127-131.
- [16] IVANIRA M, IEDA S S. Chemometric discrimination of genetically modified coffea arabica cultivars using spectroscopic and chromatographic fingerprints[J]. Talanta, 2013, 107: 416-422.
- [17] 郭恩有, 刘木华, 赵杰文, 等. 脐橙糖度的高光谱图像无损检测技术[J]. 农业机械学报, 2008, 39(5): 91-93.
- [18] 吴龙国, 何建国, 刘贵珊, 等. 基于NIR高光谱成像技术的长枣虫眼无损检测[J]. 发光学报, 2013, 34(11): 1527-1532.
- [19] XIAOBO Z, JIEWEN Z, POVEY M J W, et al. Variables selection methods in near-infrared spectroscopy[J]. Analytica Chimica Acta, 2010, 667(1): 14-32. DOI:10.1016/j.aca.2010.03.048.
- [20] 许禄. 化学计量学方法[M]. 北京: 科学出版社, 1995: 58-64.
- [21] 杨国鹏, 余旭初, 刘伟, 等. 基于支持向量机的高光谱影像分类研究[J]. 计算机工程与设计, 2008, 29(8): 2029-2032.
- [22] 吴龙国. 基于高光谱成像技术的灵武长枣常见缺陷无损检测研究[D]. 银川: 宁夏大学, 2014.
- [23] ZHANG Xiaoxun, JIA Yunde. A linear discriminant analysis framework based on random subspace for face recognition[J]. Pattern Recognition, 2007, 40(9): 2585-2591. DOI:10.1016/j.patcog.2006.12.002.
- [24] 李玉榕, 项国波. 一种基于马氏距离的线性判别分析分类算法[J]. 计算机仿真, 2006, 23(8): 86-88. DOI:10.3969/j.issn.1006-9348.2006.08.025.
- [25] 刘忠宝, 王士同. 一种改进的线性判别分析算法MLDA[J]. 计算机科学, 2010, 11(11): 239-242. DOI:10.3969/j.issn.1002-137X.2010.11.057.
- [26] WANG Nannan, YANG Yichao, SUN Dawen. Self-life prediction of 'Gros Michel' bananas with different browning levels using hyperspectral reflectance imaging[J]. Food Analytical Methods, 2014(7): 1381-1386. DOI:10.1007/s12161-013-9760-z.
- [27] 陈昕, 向健勇. 不变形理论用于空中目标的识别[J]. 红外与毫米波学报, 1997, 16(1): 39-44.