

基于近红外光谱技术的猪肉新鲜度等级研究

郭培源 林岩 付妍 王昕琨 袁芳 许冠男

(北京工商大学计算机与信息工程学院, 北京 100048)

摘要 利用近红外光谱技术检测猪肉在腐败过程中不同时刻的光谱,研究了猪肉新鲜度等级划分的可行性,并运用近红外 OPUS 软件建立了分析模型。为了更好地反映猪肉新鲜度等级,用 SOM 神经网络聚类方法重新划分了总挥发性盐基氮(TVBN)国家标准等级,由原来的 3 个等级划分成 5 个等级标准。为了提高模型预测准确度,在选用一阶导数+矢量归一化(平滑点数为 13)预处理方法基础上,在聚类分析前用主成分分析方法进行降维,使预测偏差减小,使样品预测正确率得到进一步提高,预测级别偏差减少,提高了模型预测能力。

关键词 光谱学;近红外光谱技术;主成分分析;聚类分析;猪肉新鲜度

中图分类号 TS207.3;TP29 **文献标识码** A **doi**: 10.3788/LOP50.033002

Research on Freshness Level of Meat Based on Near-Infrared Spectroscopic Technique

Guo Peiyuan Lin Yan Fu Yan Wang Xinkun Yuan Fang Xu Guannan

(College of Computer and Information Engineering, Beijing Technology and Business University, Beijing 100048, China)

Abstract We study the feasibility of pork freshness grading using the spectra detected at different time with near infrared spectroscopy and establish the analysis model by using near infrared OPUS software. In order to reflect the pork freshness level better, we use SOM neural network clustering method to differentiate afresh the volatile base nitrogen national standard (Total Volatile Basic Nitrogen, TVBN) level, from 3 levels to 5 levels. In order to improve the prediction accuracy, based on the pretreatment by selecting a derivative+vector normalization method (13 smooth points), principal component analysis is used for dimension reduction before clustering analysis. It can reduce the prediction deviation, further improve the prediction accuracy of sample, reduce the prediction level deviation, and improve the prediction ability.

Key words spectroscopy; near-infrared spectroscopic technique; principal component analysis (PCA); cluster analysis; meat freshness

OCIS codes 300.6364; 300.6170; 070.4790

1 引言

随着人们生活水平的提高,近年来我国猪肉消耗量持续增加,而各种问题猪肉层出不穷。国家对肉食品安全给予了极大重视,如何快速、准确、简便地检测肉类食品质量更是人们关注的热点。近年来近红外(NIR)技术已在豆类、牛奶食品检测中得到了应用^[1,2]。杨燕等^[3]运用近红外技术检测蜜源,为近红外光谱技术的发展提供了参考;毕松等^[4]运用多信息处理和人工神经网络(ANN)方法为猪肉检测提供了新思路。

近红外光谱是介于可见光区和中红外区间的电磁波,波长为 700~3000 nm^[5,6]。近红外光谱是由物质吸收光能使分子振动从基态向高能级跃迁时产生的,记录的是分子中单个化学键的基频振动的倍频及和频信息,它受含氢基团 X-H(X=C,N,O)的倍频与和频的重叠主导。由于猪肉品质内在物理结构具有一定的光谱特性,猪肉成分大多由这些基团构成,基团的吸收频谱表征了这些成分的化学结构。当猪肉由于微生物

收稿日期: 2012-08-28; **收到修改稿日期**: 2012-12-03; **网络出版日期**: 2013-02-28

基金项目: 北京市自然科学基金(4122020)资助课题。

作者简介: 郭培源(1958—),男,博士,教授,主要从事农产品智能检测技术和多数据融合、嵌入式技术等方面的研究。

E-mail: ggppyy@126.com

物的污染引起变质或由于自身酶分解作用而引起发酵时,其内在物理结构、成分组成特征信息会发生变化,含氢基团 X-H 或多键振动的倍频与和频信息也在发生变化^[7]。当近红外光谱和未知样品的化学组分之间的定量模型建立后,再向模型输入未知样品的近红外光谱信息,就能够确定未知样品的品质和等级^[8]。

近红外光谱检测样品无需使用化学试剂或高温、高压等测试条件,测试速度快,穿透肉品的能力强(最深可达 5 cm)^[9]。由于猪肉由蛋白质、水、多层生物组织和蛋白质分解组分等组成,红外光谱不是各组分单独光谱的叠加,可多组分同时检测,因此一次测试可以测定多种成分和指标,操作简单,无浪费,无污染,具有很高的精度,分析结果的统计准确度可接近标准化学检验方法。

本文运用近红外光谱技术对猪肉新鲜度等级进行分析研究。限于实验条件和测量环境,在通用检测标准允许的范围内,实验样品采用在市场中购买的尽量新鲜的猪肉。每隔 2 h 对猪肉样品进行 1 次测量,采集到了不同变质程度的猪肉样品光谱图,并对其进行了聚类分析和建模研究,研究了 SOM 神经网络聚类法应用于检测猪肉新鲜度等级的可行性。

2 实验材料和方法

采用德国 Bruker 公司的 MATRIX-F 近红外光谱分析仪,光谱范围为 700~2500 nm,分辨率为 2 cm^{-1} ,该仪器采用铟镓砷(InGaAs)检测器和固体测量容器。透反射测量有效光程为 2 mm,采用漫反射积分球和相关固体测量柱形杯具,杯具直径为 6 cm。实验材料在不同猪肉中选取,分成具有代表性的 35 份里脊猪肉和脂肪组织样品并编号,分别切成厚 0.5 cm 的肉片,大小正好能全部盖住杯底。同时制备 35 份相同部位的样品猪肉留作化学实验样品,并保存在 25 °C 恒温箱中。

总挥发性盐基氮 TVBN 等级(GB2722-81)是检验肉类新鲜度的国家标准。在每隔 2 h 测量一份样品的近红外光谱同时,取相同的一份样品,用凯氏定氮法测定其 TVBN 值。这样,可同时测得 35 幅近红外光谱图以及 35 个 TVBN 值。

3 猪肉新鲜度近红外光谱分析

3.1 近红外光谱定性分析流程

近红外光谱定性分析流程如图 1 所示。先将测定的近红外光谱样品分成学习集和检验集,对学习集样品谱图进行校正、预处理和信息提取等,然后建立定性模型,并不断用检验集进行模型验证,直到输出模型效果最好为止。

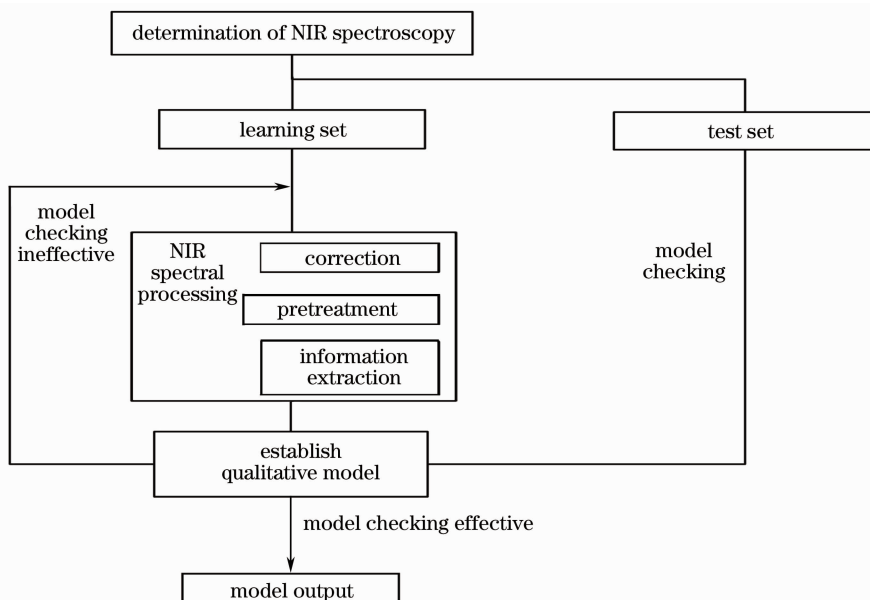


图 1 近红外光谱定性分析流程图

Fig. 1 NIR spectroscopy qualitative analysis flow chart

3.2 等级划分和集合划分

TVBN 含量是我国检测肉类新鲜度的国家标准,它能有规律地反映新鲜肉、次鲜肉和变质肉之间的差异。简单地将肉品划分成 3 个等级不能满足企业生产和人们生活中对肉品多级分级的要求。本文在猪肉腐败过程中采集了多个 TVBN 值序列,利用 SOM 神经网络(NN)对这些序列进行聚类。SOM 算法采取“胜者为王”的竞争性学习方法,将数据以定量的方式进行定性分析,在没有目标分类结果的情况下仅依据输入样本就可聚类区分,突显出不同于反向传播神经网络(BPNN)等算法的优势。再根据聚类中心以及聚类中心之间的距离,将猪肉新鲜度等级重新划分成 5 个等级标准,既可满足实际应用需要,又可验证模型的准确度级别。利用 Matlab 软件对 SOM 神经网络进行训练和仿真。SOM 神经网络结构设计如图 2 所示。

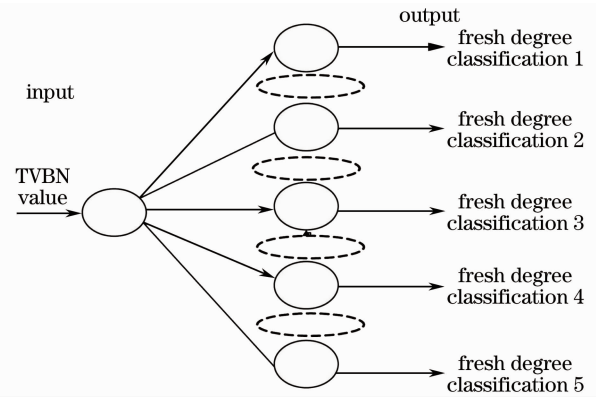


图 2 猪肉新鲜度等级划分的 SOM 神经网络结构图
Fig. 2 SOM neural network structure of pork freshness grading

取补偿 120 步后,经过训练得到各类聚类中心对应的 TVBN 值分别为 9.248、13.654、20.220、29.031、36.641 mg/100 g。新鲜度等级重新划分前后 TVBN 值范围对比如表 1 所示。

表 1 新鲜度等级重新划分前后 TVBN 值范围对比

Table 1 TVBN value range comparison before and after the re-division of the freshness grade

Before grade classification			After grade classification		
Grade name	Category	TVBN value /(mg/100 g)	Grade name	Category	TVBN value /(mg/100 g)
Fresh meat	1	<15	Fresh meat	1	<10.951
Defective meat	2	15~25	Middle fresh meat	2	10.951~16.937
			Defective meat	3	16.937~24.626
Putrid meat	3	≥25	Middle putrid meat	4	24.626~32.836
			Putrid meat	5	≥32.836

按照上面重新划分后的 5 个等级标准,将与之相同时刻测量的近红外光谱划分成 5 级并进行编号。TVBN 值小于 10.951 的对应光谱标识为 1 级,表示新鲜肉,其中样品编号为“类别·编号”,如 1.1,1.2,1.3,⋯,依此类推;TVBN 值为 10.951~16.937 对应光谱标识为 2 级,表示中鲜肉;TVBN 值为 16.937~24.626 对应光谱标识为 3 级,表示次鲜肉;TVBN 值为 24.626~32.836 对应光谱标识为 4 级,表示中腐肉;TVBN 值大于等于 32.836 的对应光谱标识为 5 级,表示腐败肉。在 35 个光谱样本中抽取了编号为 1.2,1.6,2.7,4.1,5.6 共 5 个检验集样本,其余 30 个光谱构成学习集样本。

3.3 近红外光谱预处理

采用近红外光谱对猪肉新鲜度等级研究,能够满足猪肉测量光谱波段的选择、光谱测量的定标、猪肉含氢基团 X-H 振动的倍频与和频各项检测物质的光谱“指纹特征”谱线测量及其精度分析要求。

近红外猪肉新鲜度的分析采用的是近红外光分析软件 OPUS6.5。OPUS6.5 给用户提供了 5 种预处理方式^[10]:一阶导数、二阶导数、矢量归一化、一阶导数+矢量归一化、二阶导数+矢量归一化。实验中通过各种比较,确定选用一阶导数+矢量归一化(平滑点数为 13)和二阶导数+矢量归一化(平滑点数为 9)预处理后的光谱图。选择处理光谱范围为 700~2500 nm,分辨率为 2 cm⁻¹。图 3(a)为矢量归一化预处理后的光谱图,图 3(b)和(c)分别是一阶导数+矢量归一化(平滑点数为 13)和二阶导数+矢量归一化(平滑点数为 9)预处理后的光谱图。

实验结果证明选用一阶导数+矢量归一化(平滑点数为 13)能够很好地解决基线的偏移,有效消除基线和背景干扰,提高灵敏度和分辨率,同时又能减少引入的噪声^[11],放大某些波峰和波谷对应的光谱特征,提高模型预测的准确度,能为后面的聚类分析提供准确数据。

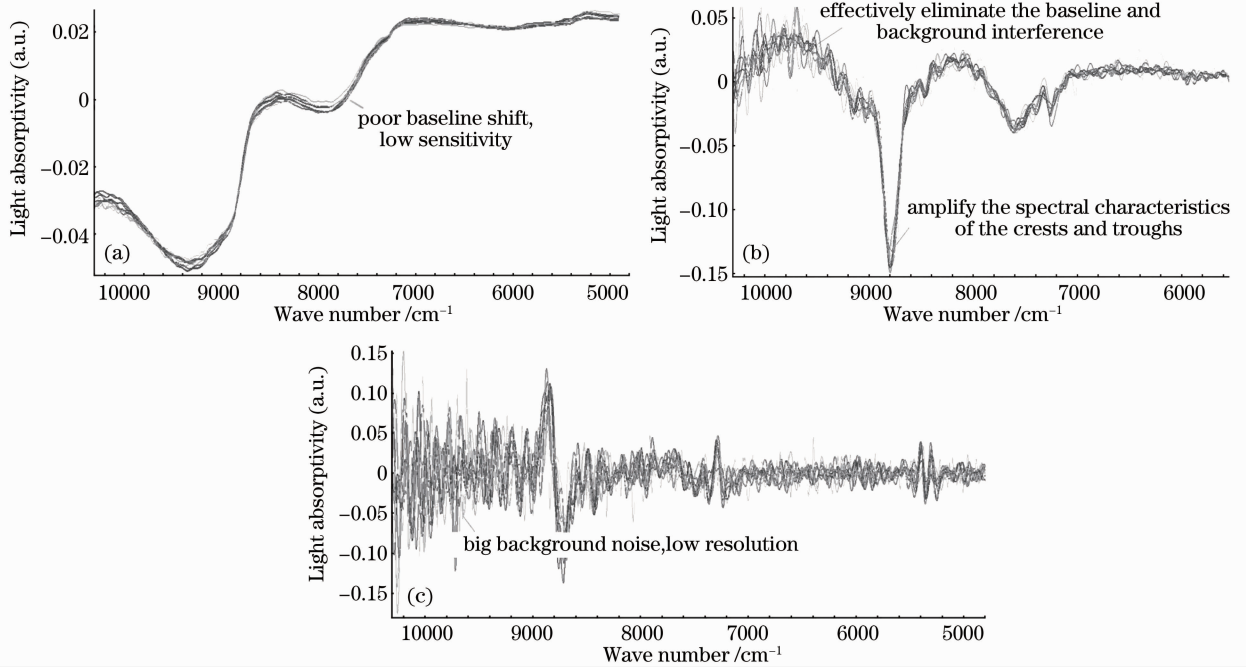


图3 预处理后的光谱图。(a)矢量归一化;(b)一阶导数+矢量归一化;(c)二阶导数+矢量归一化

Fig. 3 Spectra after pretreatment. (a) Vector normalization; (b) the first derivative and vector normalization; (c) the second derivative and vector normalization

4 猪肉新鲜度近红外光谱聚类分析

为了有效提高模型的预测能力,在聚类分析前选用主成分分析(PCA)方法进行降维处理。经PCA降维后可以排除众多化学信息共存下的相互重叠的信息。将原变量进行转换,用少数几个新变量作为原变量的线性组合,而这些新变量能尽可能好地表征原变量的数据结构特征而不丢失信息。降维过程中采用偏最小(二乘)回归方法,对变量Y中因子进行降维的同时也对变量X所含信息进行降维。

4.1 主成分分析

PCA^[12,13]是一种数学变换方法,它把给定的一组相关变量通过线性变换转成另一组不相关的变量,这些新的变量按照方差依次递减顺序排列^[14],第一主成分具有最大的方差。此处应用PCA方法旨在利用降维的思想,把多指标转化为少数几个综合指标^[15]。猪肉新鲜度检测的PCA过程如图4所示。

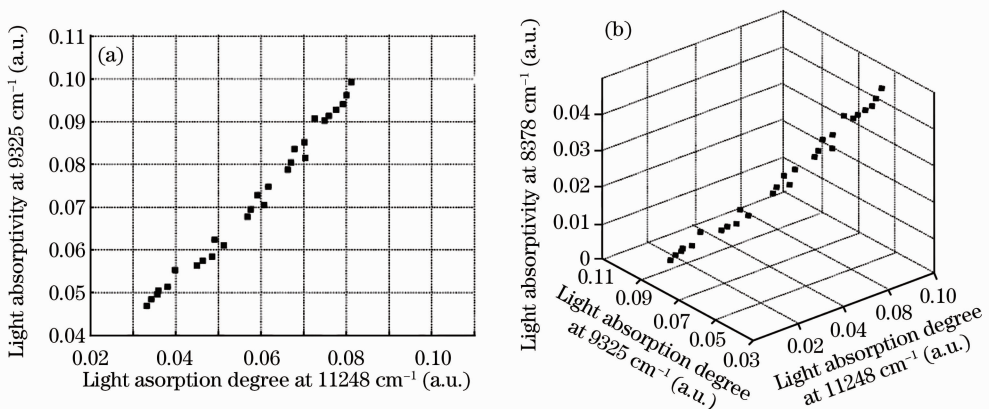


图4 猪肉的主成分分析图。(a)二维图;(b)三维图

Fig. 4 Principal component analysis of meat. (a) Two-dimensional spatial data information; (b) three-dimensional spatial data information

数学上存在 1 个 2048 维度的变量空间,可以完整地表达上述 30 个学习集样本之间的相对变化信息或者数据结构信息。

PCA 可以用几个(如 3 个或 5 个)向量建立低维光谱向量空间,能够近似完整地表达这 30 个学习集样本光谱组成的数据结构信息,与 2048 维度的变量空间相比^[16],不会丢失数据结构信息。

用 PCA 法法进行降维处理将变量浓缩成一个或者几个因子,如果校正模型的变量数(主因子数)过少,模型达不到精度;变量数过多,模型不稳定。为了提高光谱信号的有效信息利用率,并且同时避免出现“过拟合”现象,需要对主因子的个数进行合理选择。图 5 为交叉验证均方差(PRESS)值随主因子数变化的曲线图,考虑校正模型,采用主因子数为 4^[17,18]。

经过 PCA 降维处理,能将 30 个学习集样本完全聚类,说明 PCA 并没有降低模型预测能力。为了说明经过 PCA 降维处理的聚类分析方法能很好地提高模型的预测能力,进行了相应的实验,实验中仍然取 5 个在不同时刻的近红外光谱检验集样本和 TVBN 值,用经过和未经过 PCA 降维处理的聚类分析方法对 30 个学习集样本进行预测,如表 2 所示。预测结果表明,经 PCA 降维处理后,样品预测正确率由 50% 提高到 70%,预测级别偏差为 3 的样品也由 1 个减少到 0,提高了模型预测能力。

表 2 预测结果对比

Table 2 Comparison of prediction results

Whether the clustering method using PCA	Number of correct prediction samples	Number of samples whose level deviation is 1	Number of samples whose level deviation is 2	Number of samples whose level deviation is 3	Number of samples whose level deviation is 4
No	15	10	4	1	0
Yes	21	6	3	0	0

4.2 聚类分析定量模型分析

采用 OPUS 软件提供的系统聚类方法建立模型。该软件提供了 5 种不同的聚类方法:单耦合(single linkage)、全耦合(complete linkage)、平均值耦合(average linkage)、中值滤波算法(median algorithm)和质心算法(centroid algorithm)。通过比较分析发现,采用 complete linkage 聚类方法^[19]得到的聚类结果最为理想,聚类结果如图 6 所示。

光谱的距离表明了谱图的相似度^[20~22],分析图 6 可知,只要光谱的距离大于 0.924,小于 2.186,则 30 个学习集样本可以聚类为新鲜肉、中鲜肉、次鲜肉、中腐肉、腐败肉 5 级。图 7(a)和(b)分别是样品编号为 1.2 和 4.1 的预测结果。结果表明:编号为 1.2 的样品为新鲜肉,编号为 4.1 的样品为中腐肉。

5 结论

研究了近红外光谱分析技术用于检测猪肉新鲜度分级的可行性。为了提高分级模型预测的准确度,选用一阶导数+矢量归一化预处理算法,为了有效提高模型的预测能力,选用主成分分析方法进行降维,通过主成分分析,建立了猪肉新鲜度等级的预测模型。实验证明,近红外技术用于测定猪肉新鲜度等级所采用的 SOM 神经网络聚类方法等级划分方案是有效、可行的。研究表明,分级增加到 5 个等级,既可满足实际应用

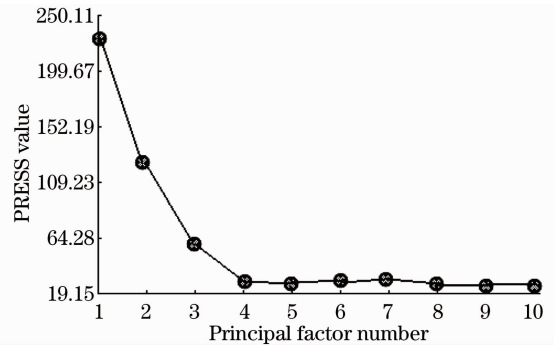


图 5 PRESS 值随主因子数变化曲线图
Fig. 5 PRESS value versus the number of principal factors

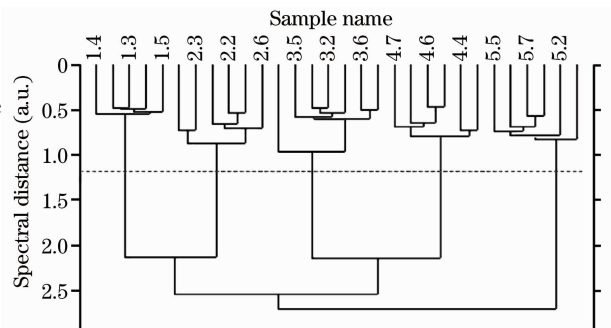


图 6 聚类树形图

Fig. 6 Clustering tree view

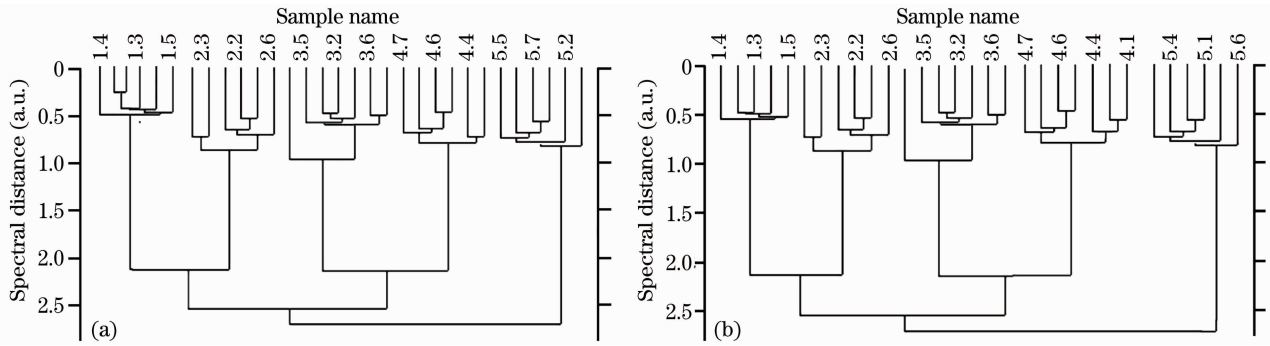


图7 样品(a) 1.2和(b) 4.1的预测结果

Fig.7 Prediction result of sample (a) 1.2 and (b) 4.1

需要,又可验证模型的准确度级别,而经过主成分分析降维处理的聚类分析方法能很好地提高模型的预测能力。这为近红外技术在工业化生产中用于分析和检测猪肉新鲜度的可行性及其预测能力的提高与改进提供了一种有效的方法和思路。

参 考 文 献

- 1 Lu Wanzen. Modern Near Infrared Spectroscopy Analytical Technology[M]. Beijing: Sinopec Press, 2007
陆婉珍. 现代近红外光谱分析技术[M]. 北京: 中国石化出版社, 2007
- 2 Zhou Zili, Zhang Yu, He Yong *et al.*. Method for rapid discrimination of varieties of rice using visible NIR spectroscopy [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2009, **25**(8): 131~135
周子立, 张 瑜, 何 勇 等. 基于近红外光谱技术的大米品种快速鉴别方法[J]. 农业工程学报, 2009, **25**(8): 131~135
- 3 Yang Yan, Nie Pengcheng, Yang Haiqing *et al.*. Rapid recognition method of nectar plant based on visible-near infrared spectroscopy[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2010, **26**(3): 238~242
杨 燕, 聂鹏程, 杨海清 等. 基于可见-近红外光谱技术的蜜源快速识别方法[J]. 农业工程学报, 2010, **26**(3): 238~242
- 4 Bi Song, Guo Peiyuan. Pork Freshness detecting method based on the change of germ area[J]. *Agricultural Mechanization Research*, 2009, (5): 88~92
毕 松, 郭培源. 基于细菌菌斑变化的猪肉新鲜度检测方法研究[J]. 农机化研究, 2009, (5): 88~92
- 5 Qiu Baijing, Liu Baoling, Wu Chundu *et al.*. Applications of near-infrared image processing in agricultural engineering[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2005, **21**(5): 102~106
邱白晶, 刘保玲, 吴春笃 等. 近红外图像处理技术在农业工程中的应用[J]. 农业工程学报, 2005, **21**(5): 102~106
- 6 Zhan Hui, Li Xiaoyu, Zhou Zhu *et al.*. Detection of chestnut defect based on data fusion of near-infrared spectroscopy and machine vision[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2011, **27**(2): 345~349
展 慧, 李小昱, 周 竹 等. 基于近红外光谱和机器视觉融合技术的板栗缺陷检测[J]. 农业工程学报, 2011, **27**(2): 345~349
- 7 Liu Jie, Li Xiaoyu, Li Peiwu *et al.*. Determination of moisture in chestnuts using near infrared spectroscopy [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2010, **26**(2): 338~341
刘 洁, 李小昱, 李培武 等. 基于近红外光谱的板栗水分检测方法[J]. 农业工程学报, 2010, **26**(2): 338~341
- 8 S. D. Shackelford. Color image analysis technique for measuring of meat[J]. *Meat Science*, 2007, **65**(3): 125~138
- 9 Lu Jun, Xiang Bingren, Liu Hao *et al.*. Application of two-dimensional near-infrared correlation spectroscopy to the discrimination of Chinese herbal medicine of different geographic regions[J]. *Spectrochim. Acta Part A*, 2008, **69**(2): 580~586
- 10 Guo Weiliang, Wang Dan, Song Jia *et al.*. Simultaneous and rapid quantitative analysis of four components in cordyceps militaris mycelium powder using near infrared spectroscopy[J]. *Acta Optica Sinica*, 2011, **31**(2): 0230002
郭伟良, 王 丹, 宋 佳 等. 近红外光谱法同时快速定量分析蛹虫草菌丝体中4种有效成分[J]. 光学学报, 2011, **31**(2): 0230002
- 11 Chen Cong, Lu Qipeng, Peng Zhongqi. Preprocessing methods of near-infrared spectrum based on NLMS adaptive filtering [J]. *Acta Optica Sinica*, 2012, **32**(5): 0530001
陈 丛, 卢启鹏, 彭忠琦. 基于NLMS自适应滤波的近红外光谱去噪处理方法研究[J]. 光学学报, 2012, **32**(5): 0530001
- 12 Yang Xiukun, Zhong Mingliang, Jing Xiaojun *et al.*. FTIR microscopic image analysis based on principal component

- analysis-2nd derivative spectral imaging[J]. *Acta Optica Sinica*, 2012, **32**(7): 0711001
- 杨秀坤, 钟明亮, 景晓军 等. 基于主成分分析-二阶导数光谱成像的红外显微图像分析[J]. *光学学报*, 2012, **32**(7): 0711001
- 13 Li Shuo, Wang Shanqin, Zhang Meiqin. Comparison among principal component regression, partial least squares regression and back propagation neural network for prediction of soil nitrogen with visible-near infrared spectroscopy[J]. *Acta Optica Sinica*, 2012, **32**(8): 0830001
- 李 硕, 汪善勤, 张美琴. 基于可见-近红外光谱比较主成分回归、偏最小二乘回归和反向传播神经网络对土壤氮的预测研究[J]. *光学学报*, 2012, **32**(8): 0830001
- 14 A. M. Mouazen, R. Karoui, J. D. Baerdemaecker *et al.*. Classification of soils into different moisture content levels based on VIS-NIR spectral[C]. Proceedings of the 2006 ASABE Annual International Meeting, 2006. 9~12
- 15 Wang Duojia, Zhou Xiangyang, Jin Tongming *et al.*. Application of near-infrared spectroscopy in agriculture and food analysis[J]. *Spectroscopy and Spectral Analysis*, 2004, **24**(4): 447~450
- 王多加, 周向阳, 金同铭 等. 红外光谱检测技术在农业和食品分析上的应用[J]. *光谱学与光谱分析*, 2004, **24**(4): 447~450
- 16 严衍禄, 赵龙莲, 韩东海. 近红外光谱分析基础与应用[M]. 北京: 中国轻工业出版社, 2005
- 17 Tian Haiqing, Ying Yibing, Lu Huishan *et al.*. Study on predicting firmness of watermelon by Vis/NIR diffuse transmittance technique[J]. *Spectroscopy and Spectral Analysis*, 2006, **27**(6): 1113~1117
- 田海清, 应义斌, 陆辉山 等. 可见/近红外光谱漫透射技术检测西瓜坚实度的研究[J]. *光谱学与光谱学分析*, 2006, **27**(6): 1113~1117
- 18 Chen Quansheng, Zhao Jiewen, Cai Jianrong *et al.*. Inspection of tea quality by using multi-sensor information fusion based on NIR spectroscopy and machine vision[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2008, **24**(3): 5~10
- 陈全胜, 赵杰文, 蔡健荣 等. 基于近红外光谱和机器视觉的多信息融合技术评判茶叶品质[J]. *农业工程学报*, 2008, **24**(3): 5~10
- 19 Xu Yun, Wu Jingzhu, Wang Yiming *et al.*. Clustering method of unknown sort samples based on near infrared spectroscopy [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2011, **27**(8): 345~349
- 徐 云, 吴静珠, 王一鸣 等. 基于近红外光谱的未知类别样品聚类方法[J]. *农业工程学报*, 2011, **27**(8): 345~349
- 20 Xiao Wu, Li Xiaoyu, Li Peiwu *et al.*. Adaptability of the model for soil moisture measurement based on near-infrared spectroscopy[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2009, **25**(3): 33~36
- 肖 武, 李小昱, 李培武 等. 基于近红外光谱土壤水分检测模型的适应性[J]. *农业工程学报*, 2009, **25**(3): 33~36
- 21 Tian Yongchao, Zhang Juanjuan, Yao Xia *et al.*. Quantitative modeling method of soil organic matter content based on near-infrared photo acoustic spectroscopy[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2012, **28**(1): 145~152
- 田永超, 张娟娟, 姚 霞 等. 基于近红外光声光谱的土壤有机质含量定量建模方法[J]. *农业工程学报*, 2012, **28**(1): 145~152
- 22 Zhang Peng, Zhang Zhihui. Rapidly changing phase error correction of Fourier transform spectrometer[J]. *Chinese J. Lasers*, 2012, **39**(s1): s115002
- 张 鹏, 张志辉. 傅里叶变换光谱仪非渐变相位误差的校正方法[J]. *中国激光*, 2012, **39**(s1): s115002