

太赫兹时域光谱技术对紫米掺假的检测研究

刘燕德, 杜秀洋, 李斌, 郑艺蕾, 胡军, 李雄, 徐佳

华东交通大学机电与车辆工程学院, 水果智能光电检测技术与装备国家地方联合工程研究中心, 江西 南昌 330013

摘要 紫米是生活中常见的食材,具有丰富的营养价值。由于紫米价格较高导致染色紫米大量流入市场。本文使用太赫兹时域光谱技术结合化学计量学方法探索紫米掺假的快速检测方法。采用太赫兹时域光谱技术(THz-TDS)采集0~7 THz范围内紫米掺假的光谱数据,并选择0.5~2.5 THz波段的吸收系数谱和折射率谱进行分析并采用化学计量学方法对光谱数据进行建模分析。分别采用 Savitzky-Golay 卷积平滑(SG Smoothing, SG 平滑)、基线校正(Baseline)、归一化(Normalization)、多元散射校正(MSC)等方法进行光谱预处理,结合偏最小二乘判别分析(PLS-DA)对紫米、紫米掺染色大米和紫米掺染色黑米进行定性分析。定性分析结果显示,通过主成分分析(PCA)的三种样品平面分布存在明显差异;经过基线校正的光谱数据建立的 PLS-DA 模型效果最佳,误判率为0。接着使用偏最小二乘法(PLS)结合 SG 平滑、Baseline、Normalization、MSC 等预处理方法分别对紫米中掺染色大米和紫米中掺染色黑米的光谱数据建立 PLS 定量模型。结果显示,采用基线校正预处理方法的 PLS 建模效果最佳,紫米掺染色大米的预测集相关系数为0.936,预测集均方根误差(RMSEP)为0.095。紫米掺染色黑米的预测集相关系数为0.914,预测集均方根误差为0.096。为对比分析线性(PLS)与非线性(LS-SVM)两种定量模型方法的预测精度,采用相同预处理方法后的紫米掺假含量光谱数据建立最小二乘支持向量机(LS-SVM)预测模型,选用径向基函数(RBF)作为核函数。结果表明采用基线校正处理后 LS-SVM 模型效果最佳,紫米中掺染色大米的预测集均方根误差(RMSEP)为0.092,预测集相关系数(R_p)为0.979;紫米中掺染色黑米的预测集均方根误差(RMSEP)为0.093,预测集相关系数(R_p)为0.948。对比发现对紫米掺假的含量建立 LS-SVM 预测模型较 PLS 模型的稳定性更好、精确度更高。研究表明,太赫兹时域光谱结合化学计量学方法可为紫米掺假的定性定量分析提供快速精确的分析方法。

关键词 太赫兹时域光谱;紫米;偏最小二乘;最小二乘支持向量机

中图分类号: O433.4 **文献标识码:** R **DOI:** 10.3964/j.issn.1000-0593(2020)08-2382-06

引言

紫米是珍贵的水稻品种,有补血益气、暖脾胃的功效;且糯性紫米粒大饱满,紫米饭清香、可口,营养价值和药用价值都比较高。黑米营养丰富,食、药用价值也很高,但消化不良者或者消化能力弱的老人小孩不宜吃黑米,适合吃紫米来调养;且黑米含铜量较高不适合肝豆状核变性患者、高铜患者食用。大米是全球人民的主要粮食之一,其营养成分远低于紫米。由于大米和黑米的价格都低于紫米,一些不法商家将大米和黑米进行染色充当紫米进行销售,从中获取暴利。这样的行为不仅降低了人们的生活质量,甚至还会危害人们的身体健康。因此对紫米掺假的鉴别及含量的检测研究

具有重要的现实意义。

近年来仪器分析技术迅速发展,在鉴别食品掺假的行业中广泛应用。顶空分析法是目前最简单的判别食品掺假的方法,但只能通过气体来鉴定是否掺假。毛细管电泳法是利用食品内部不同电子在高压的情况下会发生不同速度的定向移动原理对食品各成分进行分离检测,但其存在制备能力差、灵敏度低的缺点^[1]。太赫兹电磁波一般指的是波长在0.03~3 mm 范围之间的电磁辐射,其波段处于微波和红外辐射之间,属于远红外波段^[2]。有机物分子的转动能级和振动多数分布于此波段范围,因此太赫兹波透过样品可获取大量的物理和化学信息^[3]。太赫兹时域光谱技术是近年来发展起来的新兴技术,在食品掺假方面作出巨大贡献。目前,殷明^[4]等采用太赫兹时域光谱技术结合化学计量学方法对食用油的掺

收稿日期:2019-07-31,修订日期:2019-11-28

基金项目:国家自然科学基金项目(31760344),江西省优势科技创新团队建设计划项目(20153BCB24002),南方山地果园智能化管理技术与装备协同创新中心(赣教高字[2014]60号),江西省研究生创新专项资金项目(YC2018-S249)资助

作者简介:刘燕德,女,1967年生,华东交通大学机电与车辆工程学院教授 e-mail: jxliuyd@163.com

假做定性鉴别和定量分析研究;李斌^[5]等利用太赫兹时域光谱技术对葛粉中掺杂薯粉进行定性定量的检测研究,对葛粉中掺薯粉进行快速准确检测;Liu^[6]等利用太赫兹时域光谱技术(Terahertz time domain spectroscopy, THz-TDS)的ATR模式对蜂蜜掺假进行研究,研究表明 THz-TDS是可以准确快速的无损鉴别蜂蜜掺假的新型光谱技术。

采用 THz-TDS 结合化学计量学方法及光谱预处理方法对紫米中掺入染色大米或黑米进行定性及定量分析的检测研究。通过获取样品的吸收系数谱和折射率谱,结合偏最小二乘判别分析(partial least squares decision analysis, PLS-DA)进行建模定性分析。再建立紫米掺假含量的偏最小二乘法(partial least squares, PLS)和最小二乘支持向量机(least squares support vector machine, LS-SVM)数学模型,对模型预测结果分析比较,获得最优模型效果,为紫米掺假含量的快速检测提供方法。

1 实验部分

1.1 样品制备

实验所用的紫米、黑米和大米采购于某知名品牌,购买的紫米和黑米不存在染色掺假现象;可食用染色剂购买于某生化公司。先将黑米和大米分别加入可食用染色剂染色,染至与紫米颜色相近。将染色后的黑米和大米放入烘干箱中,温度调至 50 °C,烘干 48 h。之后将烘干后的染色黑米、大米和正常紫米分别放入高速粉碎机中进行粉碎研磨,再放入陶瓷研钵充分研磨成粉末状。根据实验设计,染色黑米和大米按不同质量比分别掺入紫米放置在离心管中,总量均为 10 g。对配比好的混合粉末使用涡旋搅拌器进行 3 min 的振荡,确保均匀分布。取适量粉末制备样品,采用压片法,制片压力约为 8 t,时间约为 1 min,压片厚度为 1.00 mm,制备压片的直径为 13 mm,表面光滑无裂痕。

1.2 仪器及光谱采集

实验使用 TAS7500SU 太赫兹时域光谱仪由日本 Advantest 公司生产,该光谱仪测量频谱范围为 0~7 THz,频谱分辨率为 7.6 GHz,采谱扫描速度为 8 ms,扫描次数设置为 8 192 次,透射模式下测量。在采集光谱前将实验室温度控制在 20 °C,湿度控制在 25%以下,开机后,通过无油静音空气压缩机向设备充入气体保证环境干燥,预热约 30 min 待设备内部的空气湿度降低至 5%,开始采集光谱。再通过傅里叶变换将时域信号转换为频域信号,采用 Dorney^[7]和 Duvillaret^[8]等提出的方法进行数据处理得到样品的吸收系数和折射率。

1.3 模型的建立

1.3.1 建模方法

偏最小二乘法是光谱技术检测分析中一种基于因子分析的多变量校正算法,是一种多因变量对多自变量的回归建模方法^[9]。以光谱矩阵 \mathbf{X} 与浓度矩阵 \mathbf{Y} 相关联数学模型为

$$\mathbf{Y} = \mathbf{bX} + e \quad (1)$$

式(1)中, \mathbf{b} 所表征的即为 \mathbf{X} 和 \mathbf{Y} 间的回归系数向量, e 标示模型残差。

最小二乘支持向量机采用最小二乘线性系统作为损失函数,通过解一组线性方程组代替传统 SVM 采用的较复杂的二次规划方法,降低计算复杂性,加强了求解速度^[10]。本文主要选用 RBF 作为核函数,其表达式为

$$K(X_k, X_i) = \exp\left(-\frac{\|X_k - X_i\|^2}{(2\sigma^2)}\right) \quad (2)$$

式(2)中, X_k 是样本点, X_i 是核函数中心点, σ^2 是内核参数。

1.3.2 模型评价

通过校正集均方根误差(root mean square error of calibration, RMSEC)、预测集均方根误差(root mean square error of prediction, RMSEP)、校正相关系数(R_c)、预测相关系数(R_p)等参数来评价模型的优劣。若 R_c 和 R_p 数值越接近 1,说明其校正集和预测集相关程度越高。当 RMSEC 和 RMSEP 的值越接近 0,表明模型的建模效果好和预测能力越强^[11]。

2 结果与讨论

2.1 太赫兹时域光谱结合化学计量学的定性分析

2.1.1 定性光谱分析

本太赫兹时域光谱仪的测量范围为 0~7 THz,由于 0~0.5 THz 及 2.5~7 THz 波段的信噪比低,故选取 0.5~2.5

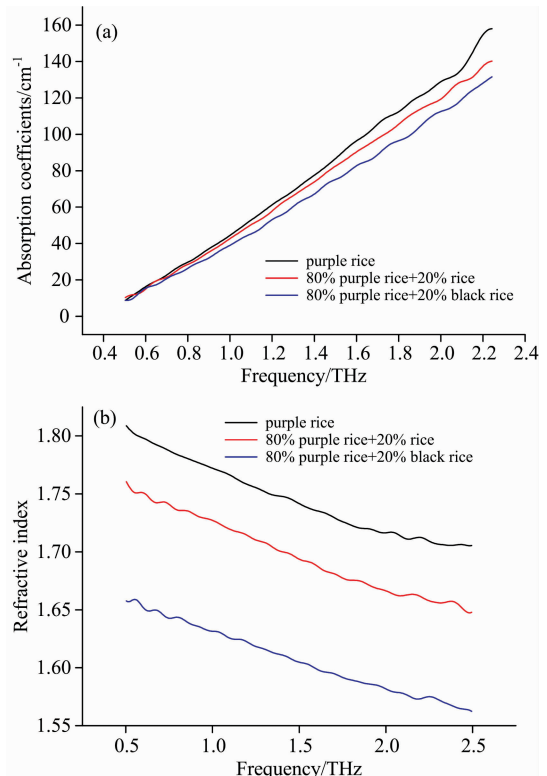


图 1 紫米掺假的吸收谱和折射谱图

(a): 吸收系数图; (b): 折射率图

Fig. 1 Absorption spectrum and refraction spectrum of purple rice adulteration

(a): Absorption coefficient diagram; (b): Refractive index

THz 范围为本实验的研究波段。实验准备纯紫米、80%紫米+20%大米和 80%紫米+20%黑米 3 类样品, 每类样品制备 20 个, 共 60 个样品。对每个样品采集 3 条太赫兹光谱, 取其平均光谱数据进行处理分析。

图 1 为紫米、紫米掺入染色大米和紫米掺入染色黑米的吸收谱和折射谱。由于紫米、大米和黑米都是由碳水化合物、蛋白质和脂肪等物质构成的, 三种样品的吸收图谱和折射率谱的形状相似且没有明显的特征峰。由图 1 可知, 随着频率的提高, 吸收系数不断升高, 折射率不断降低。纯紫米光谱的吸收系数和折射率均高于掺假紫米, 紫米掺染色大米的吸收系数和折射率均高于紫米掺染色黑米。根据光谱曲线的高低不能准确地鉴定出紫米是否存在掺假现象, 需要结合化学计量学方法鉴定出紫米是否存在掺假。

2.1.2 结合化学计量学的建模定性分析

本文为了对比三种样品的光谱差异, 对光谱进行主成分分析(principal component analysis, PCA), 以光谱的前 2 个主成分(前两个 PC 累计方差贡献率为 92%)的得分计算样本的平面距离, 三种样品的二维得分分布图如图 2 所示。三种样品的平面分布存在明显的差别, 反映了三种样品由于碳水化合物、蛋白质、脂肪等物质的含量不同在光谱信息中存在差异。

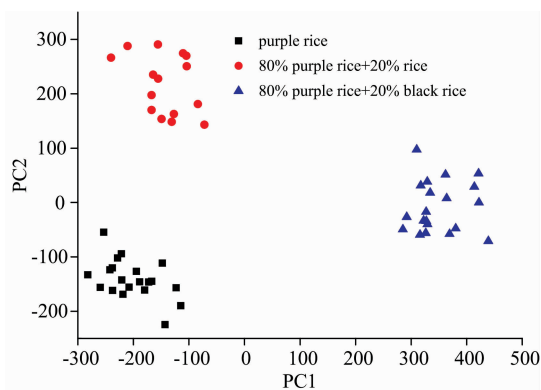


图 2 三种样品的前 2 个主成分得分分布

Fig. 2 The first two PCs score plots of spectra of the three samples

使用偏最小二乘判别分析(PLS-DA)对三种样品进行定性建模分析。校正集和预测集按照 3:1 的比例进行分配。为了减少采集光谱过程中的误差干扰, 对光谱采用 SG 平滑、归一化、多元散射校正(multiple scattering correction, MSC)、基线校正等多种预处理方法后建立 PLS-DA 模型^[12]。表 1 为经过不同预处理之后建立模型的效果, 使用基线校正预处理的数据进行 PLS-DA 建立模型效果最佳, 其预测相关系数(R_p)为 0.971, 预测均方根误差(RMSEP)为 0.142。

图 3(a)为在基线校正之后的 PLS-DA 建模所得的模型散点图, 其中 T1 为第一类样品和第二类样品的分界线, 其阈值为 1.5; T2 为第二类样品和第三类样品的分界线, 其阈值为 2.5。从散点图可知: 所有预测集样品未出现误判的情况, 预测误判率为 0。即可以利用 PLS-DA 建模, 以误判率和

相关系数为指标, 实现紫米中是否掺假大米或黑米的鉴别。

表 1 不同预处理方法的 PLS-DA 模型结果

Table 1 PLS-DA model results of different pretreatment methods

预处理方法	R_c	R_p	RMSEC	RMSEP
原始	0.972	0.961	0.142	0.173
SG 平滑	0.968	0.958	0.151	0.175
基线校正	0.981	0.971	0.115	0.142
归一化	0.984	0.935	0.112	0.213
多元散射校正	0.983	0.925	0.113	0.232

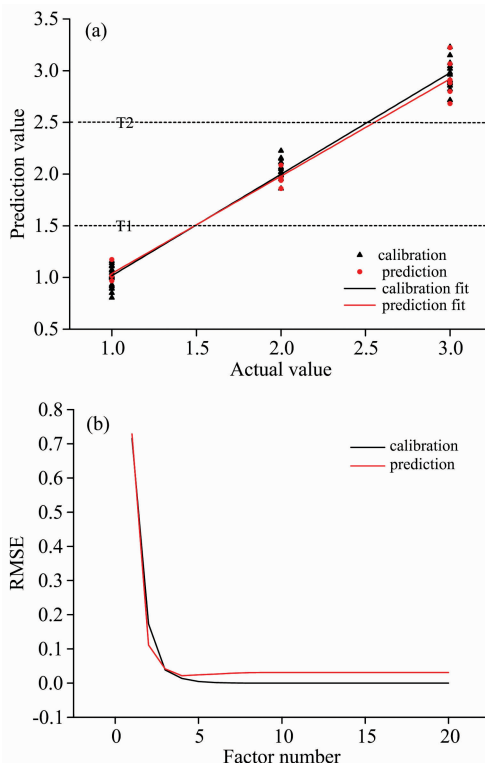


图 3 PLS-DA 建模结果图

(a): 建模和预测散点图; (b): 均方根误差随主成分因子变化图

Fig. 3 PLS-DA modeling result graph

(a): Modeling and predicting scatter plots;

(b): RMSE varies with principal component factor

图 3(b)为主成分因子决定图, 校正集及预测集均方根误差随主成分数的增加而下降, 当主成分数为 3 时, 预测均方根误差降到最低。之后随着主成分数的增加, 预测均方根误差保持稳定。故最佳主成分因子数为 3。

2.2 太赫兹时域光谱结合化学计量学的定量分析

2.2.1 定量光谱及偏最小二乘法建模分析

在对三种类型的样品进行定性分析之后, 分别对紫米中掺染色大米和紫米中掺染色黑米进行定量建模分析。两个模型的建立需根据不同的质量分数制备 21 个质量分数梯度, 每个样品制备 4 或 8 个样品, 各样品质量分数和个数如表 2 所示。每个样品采集 3 条太赫兹光谱, 取其平均光谱数据进

表 2 21 组样品质量分数(g · g⁻¹, %)

Table 2 Mass fraction of 21 samples

样品编号	质量分数	样品个数	样品编号	质量分数	样品个数	样品编号	质量分数	样品个数
1	0	8	8	35	4	15	70	4
2	5	4	9	40	8	16	75	4
3	10	4	10	45	4	17	80	8
4	15	4	11	50	4	18	85	4
5	20	8	12	55	4	19	90	4
6	25	4	13	60	8	20	95	4
7	30	4	14	65	4	21	100	8

行处理分析。

图 4(a)为 6 种不同质量分数的紫米掺染色大米样品的光谱吸收系数图,由图可知,纯紫米的吸收系数最高,随着染色大米掺入的质量分数越大,光谱的吸收系数降低越多。每个不同质量分数的吸收曲线都会有一些不规则的波动,不能证明这些波动是由于紫米和 大米中碳水化合物、蛋白质或者脂肪含量的不同造成的。要将太赫兹时域光谱与化学计量学结合起来进行紫米掺入染色大米的定量分析。

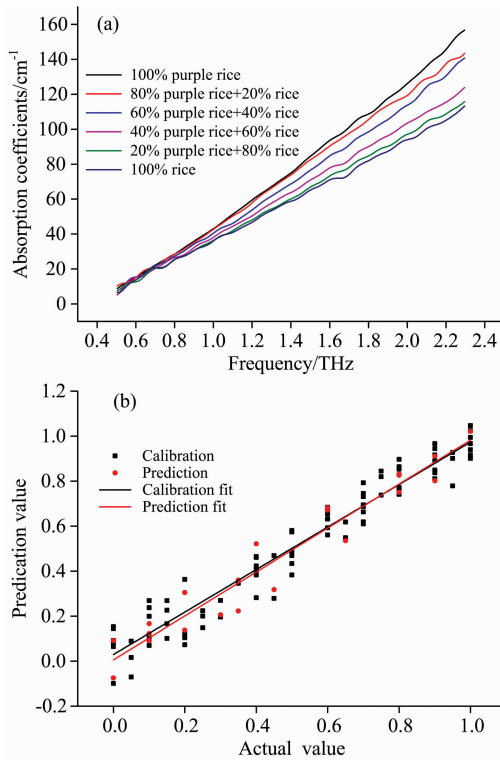


图 4 紫米掺染色大米光谱分析及建模结果图

(a): 不同含量样品光谱图; (b): 建模和预测含量散点图

Fig. 4 Spectral analysis and modeling results of purple rice-stained rice

(a): Spectral map of different content samples;

(b): Modeling and prediction content scatter plot

使用偏最小二乘法(PLS)对紫米掺入染色大米的实验进行定量建模分析,将 108 个样品数据按照校正集和预测集 3 : 1 的比例进行配比,校正集 81 个光谱数据,预测集 27 个

光谱数据。对光谱采用 SG 平滑、归一化、多元散射校正、基线校正的预处理方法。结果如表 3 所示,使用基线校正的预处理方法结果最好;预测集相关系数 R_p 为 0.936,预测集均方根误差 RMSEP 为 0.095。图 4(b)为经过基线校正的校正集与预测集的散点图及拟合线,表示其真实值与预测值的相关性。由图可知,模型精度和预测准确度均达到较高水平。

表 3 不同预处理方法的 PLS 模型结果

Table 3 PLS model results of different pretreatment methods

预处理方法	R_c	RMSEC	R_p	RMSEP
原始	0.959	0.061	0.926	0.099
SG 平滑	0.954	0.064	0.926	0.099
多元散射校正	0.912	0.096	0.832	0.149
归一化	0.961	0.058	0.821	0.153
基线校正	0.968	0.061	0.936	0.095

采用同样的预处理方法和建模方法对紫米中掺染色黑米的数据进行处理。结果显示,校正集相关系数 R_c 为 0.916,校正集均方根误差 RMSEC 为 0.078;预测集相关系数 R_p 为 0.914,预测集均方根误差 RMSEP 为 0.096。使用 PLS 建模能够准确预测紫米中掺入染色黑米的含量。

2.2.2 最小二乘支持向量机建模分析

由于建立的 PLS 模型为线性模型,即在 LS-SVM 中选择 RBF_kernel 核函数来建立非线性模型进行对比。其中需确定正则化参数(γ)和内核参数(σ^2),这两个参数会影响模型精确度。采用 4 种预处理方法处理的数据作为最小二乘支持向量机(LS-SVM)模型的输入值。表 4 为 LS-SVM 预测紫米中掺入染色大米的预测建模结果,经基线校正处理的建模预测精度最高,预测均方根误差(RMSEP)为 0.0917,相关系数(R_p)高达 0.979 6,其对应的参数 $\gamma=8 658, \sigma^2=1 845$ 。

表 4 LS-SVM(RBF_kernel)模型的预测结果

Table 4 Prediction results of LS-SVM (RBF_kernel) model

预处理方法	参数		R_p	RMSEP
	γ	σ^2		
原始	497 434	3 088	0.979 4	0.091 9
SG 平滑	1 049 499	3 286	0.979 5	0.092 2
多元散射校正	688 010	2 365	0.962 4	0.121 6
归一化	455 182	2 430	0.968 7	0.111 4
基线校正	8 658	1 845	0.979 6	0.091 7

图 5 为基线校正后 LS-SVM 预测紫米掺假染色大米含量的散点图。

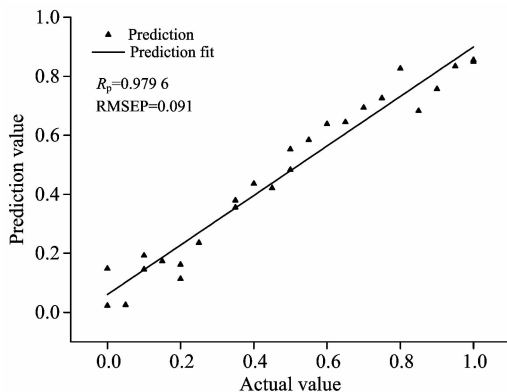


图 5 LS-SVM 紫米掺染色大米含量预测模型散点图

Fig. 5 LS-SVM purple rice blended dye rice prediction model scatter plot

采用基线校正的预处理和 RBF_kernel 核函数对紫米掺染色黑米的数据建立 LS-SVM 模型。结果显示, 预测集均方根误差 (RMSEP) 为 0.093 6, 预测集相关系数 (R_p) 为 0.947 5。

2.2.3 模型对比分析

结果显示 PLS 和 LS-SVM 均可对紫米掺假含量进行准

确检测, 但采用基线校正处理建立 LS-SVM 模型精度最佳。LS-SVM 建立紫米掺入染色大米含量模型的 RMSEP 为 0.091、 R_p 为 0.979; LS-SVM 建立紫米掺入染色大米含量模型的 RMSEP 为 0.093、 R_p 为 0.947。LS-SVM 比 PLS 预测紫米掺假含量效果更佳、模型更为稳定, 原因可能是紫米及掺假品中含有更多的非线性信息。

3 结 论

基于太赫兹时域光谱技术结合化学计量学方法对紫米掺假进行定性和定量分析。首先进行对紫米、紫米掺染色大米和紫米掺染色黑米进行定性分析研究。使用基线校正预处理之后的数据进行 PLS-DA 建立模型的误判率为 0, 能够准确地区分三种样品。对紫米掺染色大米和紫米掺染色黑米分别进行定量分析研究, 使用偏最小二乘法 (PLS) 与最小二乘支持向量机法 (LS-SVM) 分别建立定量模型。结果显示, 使用基线校正预处理之后的数据结合最小二乘支持向量机法 (LS-SVM) 进行定量建模效果最佳, 紫米掺染色大米的预测相关系数 (R_p) 为 0.979, 预测均方根误差 (RMSEP) 为 0.091; 紫米掺染色黑米的预测相关系数 (R_p) 为 0.948, 预测均方根误差 (RMSEP) 为 0.093。研究表明, 太赫兹时域光谱技术可用于紫米掺假定性定量研究。本研究为太赫兹时域光谱技术在食品领域的快速检测提供方法参考。

References

- [1] ZHAO Jia(赵 佳). China Food Safety Magazine(食品安全导刊), 2018, (12): 105.
- [2] Josette El Haddad, Frederick de Miollis, Joyce Bou Sleiman, et al. Analytical Chemistry, 2014, 86: 4927.
- [3] HUANG Li-juan, ZHANG Xin, WANG Guo, et al(黄丽娟, 张 欣, 王 果, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2016, 36(S1): 79.
- [4] YIN Ming(殷 明). Identification and Adulteration Detection of Edible Oils Based on Broad Band Terahertz Spectrum and Stoichiometry (基于宽频太赫兹光谱和化学计量学的食用油脂鉴别及掺假检测研究). China University of Mining and Technology(中国矿业大学), 2017.
- [5] LI Bin, DU Xiu-yang, LIU Yan-de, et al(李 斌, 杜秀洋, 刘燕德, 等). Laser and Optoelectronics Progress(激光与光电子学进展), 2019, (20): 322.
- [6] Liu W, Zhang Y, Yang S, et al. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2018, 196: 123.
- [7] Dorney T D, Baraniuk R G, Mittleman D M. Journal of the Optical Society of America A-Optics Image Science and Vision, 2001, 18(7): 1562.
- [8] Duvillaret L, Garet F, Coutaz J L. Applied Optics, 1999, 38(2): 409.
- [9] YAN Yan-lu(严衍禄). Basic and Application of Near Infrared Spectroscopy(近红外光谱分析基础与应用) Beijing: China Light Industry Press(北京: 中国轻工业出版社), 2005. 8.
- [10] CHU Xiao-li(褚小立). Molecular Spectroscopy Analytical Technology Combined with Chemometrics and Its Applications(化学计量学方法与分子光谱分析技术). Beijing: Chemical Industry Press(北京: 化学工业出版社), 2011. 2.
- [11] HAO Yong, SUN Xu-dong, GENG Xiang(郝 勇, 孙旭东, 耿 响). Food Science(食品科学), 2013, 34(18): 137.
- [12] OUYANG Ai-guo, ZHANG Yu, CHENG Meng-jie, et al(欧阳爱国, 张 宇, 程梦杰, 等). Chinese Optics(中国光学), 2017, 10(6): 752.

Detection of Purple Rice Adulteration by Terahertz Time Domain Spectroscopy

LIU Yan-de, DU Xiu-yang, LI Bin, ZHENG Yi-lei, HU Jun, LI Xiong, XU Jia

School of Mechatronics Engineering, East China Jiaotong University, National and Local Joint Engineering Research Center of Fruit Intelligent Photoelectric Detection Technology and Equipment, Nanchang 330013, China

Abstract Purple rice is a common ingredient in life and has rich nutritional value. Due to the high price of purple rice, the dyed purple rice has entered the market in large quantities. In this paper, terahertz time-domain spectroscopy combined with chemometric methods is used to explore the rapid detection method of purple rice adulteration. The spectral data of purple rice adulteration in the range of 0~7 THz was collected by Terahertz Time domain Spectroscopy (THz-TDS), and the absorption coefficient spectrum and refractive index spectrum of 0.5~2.5 THz band were selected for analysis and adopted. The chemometric method models and analyzes the spectral data. Savitzky-Golay convolution smoothing (SG smoothing), baseline correction (Baseline), normalization (Normalization), multiple scattering correction (MSC) and other methods are used for spectral pre-processing. Qualitative analysis of purple rice, purple rice mixed with rice and purple rice mixed with black rice was carried out by partial least squares decision analysis (PLS-DA). Qualitative analysis showed that there were significant differences in the plane distribution of the three samples by Principal Component Analysis (PCA); the PLS-DA model established by baseline corrected spectral data had the best effect, and the false positive rate was 0. Then using partial least squares (PLS) combined with SG smoothing, Baseline, Normalization, MSC and other pretreatment methods to establish a PLS quantitative model for the spectral data of the black rice mixed with dyed rice and purple rice. The results showed that the PLS model with baseline correction pretreatment method had the best effect. The correlation coefficient of the prediction set of purple rice-doped rice was 0.936, and the root means square error of prediction (RMSEP) was 0.095. The correlation coefficient of the prediction set of purple rice blended black rice was 0.914, and the root mean square error of the prediction set was 0.096. In order to compare and analyze the prediction accuracy of linear (PLS) and nonlinear (LS-SVM) quantitative model methods, the least squares support vector machine (least squares support vector) is established by using the same pretreatment method. Machine, LS-SVM) predictive model, using radial basis function (RBF) as the kernel function. The results showed that the LS-SVM model had the best effect after baseline correction. RMSEP of the predicted rice with purple rice was 0.092, and the correlation coefficient (R_p) of the prediction set was 0.979. RMSEP of the meter is 0.093, and the prediction set correlation coefficient (R_p) is 0.948. The comparison found that the LS-SVM prediction model for the content of purple rice adulteration is better and more accurate than the PLS model. Studies have shown that terahertz time-domain spectroscopy combined with chemometric methods can provide a fast and accurate analytical method for qualitative and quantitative analysis of purple rice adulteration.

Keywords Terahertz time-domain spectroscopy; Purple rice; Partial least squares; Least squares support vector machine

(Received Jul. 31, 2019; accepted Nov. 28, 2019)