

利用同步荧光光谱快速鉴别泔水油

孙艳辉, 安海洋, 贾小丽, 王娟

滁州学院生物与食品工程学院, 安徽 滁州 239012

摘要 为快速鉴别泔水油, 采用三维同步荧光光谱结合平行因子法解析泔水油的特征波长差($\Delta\lambda$), 并利用支持向量机建立泔水油鉴别模型。结果表明, 泔水油的特征 $\Delta\lambda$ 为 60 nm; 特征 $\Delta\lambda$ 下的样品原始同步荧光光谱经过主成分分析提取 5 个主成分, 以径向基函数(RBF)为核函数, 利用网格搜索和 6-fold 交叉验证优化建模参数, 得到惩罚因子 $C=512$ 、核参数 $g=0.5$, 该条件下建立的模型对训练集和预测集的判别率均达到 100%。采用同步荧光光谱可以快速、准确地鉴别泔水油。

关键词 泔水油; 同步荧光光谱; 平行因子法; 支持向量机; 鉴别

中图分类号: O657; TS207.3 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2012)10-2726-04

引言

泔水油含有大量饱和及不饱和的醛、酮、内酯甚至丙烯酰胺、多环芳烃类致癌化合物, 这些物质对人体酶系统有严重的破坏作用。另一方面, 反复的加热精炼使得营养物质如维生素 E、植物甾醇大量减少^[1]。不法商贩为追求高额利润在食用油中掺杂泔水油, 严重威胁人民群众的身体健康。因此, 需要加强泔水油检测技术的开发, 保障食用油安全。

国内学者对泔水油的检测进行了积极探索^[2, 3]。然而, 泔水油产业逐渐做大做强, 精炼程度高, 工业化水平发达, 原有针对酸价、水分含量、胆固醇、重金属、表面活性剂等指标的检测方法无法反映真实状况。因此, 需要建立适应新情况的泔水油快速检测技术。

食用油中有很多能发出荧光的物质, 如生育酚、生育三烯酚、叶绿素、脱镁叶绿素、黄酮类物质, 另外污染物黄曲霉毒素、多环芳烃类物质等也具有荧光, 这些物质的存在使得采用荧光光谱法检测食用油成为可能。王艳等利用三维荧光光谱法测定了八种植物油脂的荧光信息^[4], 方惠敏认为植物油的荧光光谱可能是其含有的叶绿素、酚类等荧光物质共同作用的结果^[5]。贾艳华等利用同步扫描荧光光谱法, 并结合红外吸收光谱法对经过高温煎炸的几种食用油进行了分析和研究, 结果表明荧光光谱方法可以用来对煎炸食用油进行检测^[6]。章汝平采用恒能量同步荧光光谱法测定食品中的苯并[a]芘的含量, 该方法可以测定苯并[a]芘的含量。而采用

荧光光谱法鉴别泔水油的研究未见报道。

本工作首先绘制泔水油和食用油的三维同步荧光光谱, 结合平行因子分析法解析得到泔水油同步荧光光谱的特征 $\Delta\lambda$; 然后, 特征 $\Delta\lambda$ 下扫描食用油、泔水油及其混合油样品得到特征同步荧光光谱, 通过支持向量机算法, 建立稳定、可靠的泔水油鉴别模型, 为泔水油的快速鉴别提供技术支持。

1 实验部分

1.1 样品与仪器

12 个食用油(米糠油、调和油、大豆油、菜籽油)样品和 4 个泔水油样品由滁州市技术监督检测中心提供。以泔水油的质量百分比为: 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10% 分别与米糠油、调和油、大豆油和菜籽油混合, 搅拌均匀, 得到混合油样品 72 个。

Cary Eclipse 荧光分光光度计, 1 cm 石英比色皿, 瓦里安有限公司。

1.2 光谱测量

对米糠油、调和油、泔水油、大豆油和菜籽油进行三维同步荧光光谱扫描, 条件为: 激发波长 240~700 nm, 波长差 $\Delta\lambda$ 为 10~120 nm, 步长为 10 nm, 激发狭缝 5 nm, 发射狭缝 10 nm, 扫描速度 1 200 nm·min⁻¹, 数据采集间隔 2 nm。

对食用油、泔水油及其混合油样品进行同步荧光光谱, 扫描条件为: 激发波长 240~700 nm, 波长差 $\Delta\lambda$ 为 60 nm,

收稿日期: 2012-02-21, 修订日期: 2012-05-25

基金项目: 国家自然科学基金项目(30901129), 安徽省自然科学基金项目(1208085QC69)和安徽省农产品质量与安全特色专业基金项目(20101032)资助

作者简介: 孙艳辉, 1978 年生, 滁州学院生物与食品工程学院副教授 e-mail: syh2004@126.com

激发狭缝 5 nm, 发射狭缝 10 nm, 扫描速度 $1\ 200\ \text{nm} \cdot \text{min}^{-1}$, 数据采集间隔 2 nm。

随机抽取样品, 将 0.8 mL 油注入样品池中进行测量。每份样品进行 3 次装样, 分别进行扫描, 取 3 次扫描的平均光谱作为原始分析光谱。

1.3 数据处理

在 MATLAB7.8 平台进行数据处理, 平行因子分析法为丹麦哥本哈根大学生命科学学院食品科学系提供(<http://www.models.life.ku.dk/nwaytoolbox>)的共享程序包, 支持

向量机算法为台湾大学林智仁提供(<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>)的共享程序包。

2 结果与讨论

2.1 泔水油特征 $\Delta\lambda$ 的获取

对米糠油、调和油、泔水油、大豆油和菜籽油进行三维同步荧光扫描, 得到如图 1 所示的三维同步荧光光谱图。

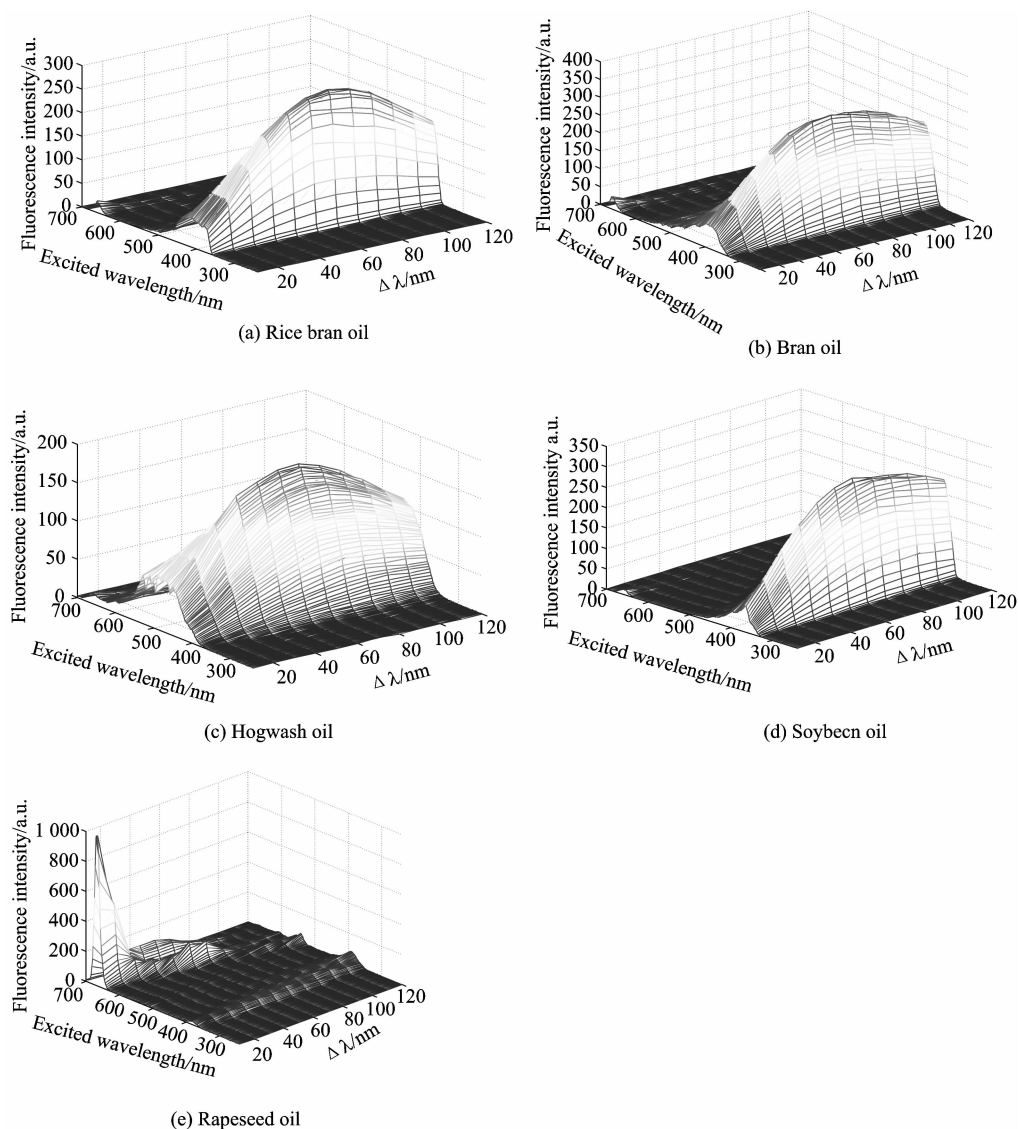


Fig. 1 Three dimensional synchronous fluorescence spectra of feed oil and reused oil

三维同步荧光光谱的三个维度分别是激发-发射波长差 ($\Delta\lambda$)、激发波长 (Excited wavelength) 和荧光强度 (Fluorescence intensity), 通过在不同的 $\Delta\lambda$ 下记录激发光谱获得。其显著特点是可提供样品的特征同步荧光峰。由图 1 可以发现, 米糠油、调和油、泔水油、大豆油和菜籽油在峰位置、峰强、峰形状等均有差异。米糠油在 ($\Delta\lambda$, Excited wavelength) 为 70 和 382 nm 处有同步荧光峰; 而调和油在 ($\Delta\lambda$, Excited

wavelength) 为 70 和 370 nm 处有同步荧光峰, 大豆油在 ($\Delta\lambda$, Excited wavelength) 为 70 和 364 nm 处有同步荧光峰, 菜籽油在 ($\Delta\lambda$, Excited wavelength) 为 10, 666, 70, 606, 70, 372 nm 处有同步荧光峰, 这与方惠敏研究结果一致^[5], 这些特征峰可能是植物油中所含叶绿素、酚类等荧光物质共同作用的结果。而泔水油在 ($\Delta\lambda$, Excited wavelength) 为 70, 448, 60, 466, 60, 512, 60, 530 nm 有同步荧光峰, 这些特征峰可

能是高温作用下产生二聚物、三聚物等氧化物共同作用的结果^[6]。

三维同步荧光光谱虽然能提供精细特征峰,但存在扫描时间长的缺点。同步荧光光谱法具有选择性提高、谱图简化、光散射干扰减少等特点,并且操作简便、节省分析成本,与三维荧光光谱相比能缩短分析时间,尤其适合对多组分混合物的分析。如能找到样品间的一个特征差异 $\Delta\lambda$, 单独扫描得到特征同步荧光光谱, 就可以解决获取样本光谱的时间问题。平行因子分析法(PARAFAC)可以对三维荧光数据进行分解, 依据是三线性分解理论, 它的最大优点是分解具有唯一性。刘海龙等已采用平行因子法和三维荧光光谱技术对绿茶进行了分析和分类^[7]。吕桂才等采用平行因子法应用于赤潮藻滤液三维荧光光谱特征的提取^[8]。

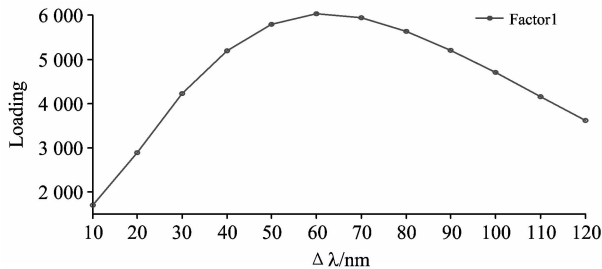


Fig. 2 Single factor loading of $\Delta\lambda$

对米糠油、调和油、泔水油、大豆油和菜籽油的 10 个三维数据矩阵进行 1 组分的平行因子解析, 得到如图 2 所示的 $\Delta\lambda$ 载荷得分。图 2 中的曲线是不同 $\Delta\lambda$ 的载荷得分, 得分越高, 说明该 $\Delta\lambda$ 代表的样本间差异信息越丰富。由图 2 可以发现, $\Delta\lambda$ 为 60 nm 时其载荷得分最高, 所以, 60 nm 为泔水油与食用油差异的特征 $\Delta\lambda$ 。

在 $\Delta\lambda$ 为 60 nm 时绘制食用油、泔水油及其混合样品的同步荧光图谱, 图 3 为部分样品的原始同步荧光图。由图 3 可以看出, 泔水油的同步荧光特征与米糠油、调和油、大豆油和菜籽油等食用油的具有显著性差异。

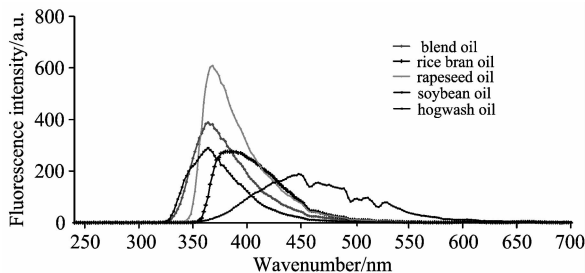


Fig. 3 Synchronous fluorescence spectra of hogwash oil and edible oil ($\Delta\lambda = 60$ nm)

2.2 原始光谱的主成分分析

将泔水油及其混合液样品定义为 0, 食用油样品定义为 1。为提高模型稳定性, 采用 SPXY 算法按 3 : 1 比例将试验样本分为训练集和预测集, 用于模型的建立和验证。SPXY 算法以光谱与理化值的共生距离为依据, 对样本集进行选择, 进而最大程度的表征样本分布^[9]。

由于同步荧光光谱中包含的波长点数较多, 对原始光谱数据直接进行建模会导致建模时间长, 并且光谱中相互重叠的信息无法消除。因此, 在建立模型前, 需要对数据降维。主成分分析(PCA)是常用的方法之一。合适的主成分既能避免过多地丢失光谱信息, 又能去除建模中的信息冗余^[10]。对所研究的全部 88 份油样品原始同步荧光光谱数据进行主成分分析, 结果表明前 5 个主成分的累积方差贡献率为 99.33%, 几乎能全部反映油样品原始同步荧光光谱矩阵的总体信息。因此, 选用原始光谱的前 5 个主成分建立模型。

2.3 支持向量机(SVM)模型建立

支持向量机是通过结构风险最小化原理来提高泛化能力的一种通用学习方法, 能较好地解决小样本、高维数、非线性、局部极小点等问题。支持向量机包含适用在线性可分情况下的线性核函数、适用在非线性可分情况下的多项式核函数、二层神经网络核函数和高斯径向基函数(RBF)^[11]。由于泔水油样本集属于非线性可分问题, 并且 RBF 有利于减少模型训练过程中的复杂性, 因此选用 RBF 为核函数。

惩罚因子 C 和核宽度参数 g 的选择对采用 RBF 核函数的 SVM 模型非常重要。 C 用于衡量误差罚分的大小, g 则控制函数回归误差, 并且会直接影响初始的特征值和特征向量^[12]。研究中以训练集为对象, 采用网格搜索法和 6-fold 交叉验证相结合的方法来选择参数。参数优化步骤如下: 首先设定核参数 C 和 g 的步长、起始值、终止值均为 $[-10, 2, 10]$; 然后独立并行地逐格搜索每一参数对 (C, g) 的值, 最后利用 6-fold 交叉验证法对目标函数寻优计算, 直到找出最优参数对, 确保交叉验证准确率最高。搜索结果如图 4 所示, 图中所示曲线为参数组合 $\log_2 C$ 和 $\log_2 g$, SVM 分类模型的交叉验证预测准确率的等高线图, 等高线上标注的为预测率。由图 4 可以看出, 在 $C=512$ 和 $g=0.5$ 时交叉验证预测准确率最优, 为 93.84%。

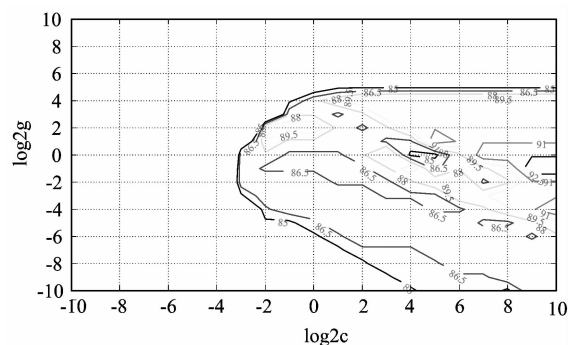


Fig. 4 Contour map of cross validation prediction rate of SVM model

选用惩罚因子 $C=512$ 、核宽度参数 $g=0.5$ 作为 RBF 的参数, 对测试集进行训练以建立模型。该模型对训练集样本的判别率和预测集样本的预测鉴别率均达到 100%。

2.4 模型性能分析

为比较支持向量机模型的性能, 同时考察了主成分回归分析(PCR)、偏最小二乘(PLSR)、ANN 神经网络等方法对测试集的预测正确率如表 1 所示。从表 1 中可以看出, 支持

向量机对应的模型性能最好。

Table 1 Effect of different modeling methods on the prediction accuracy of the prediction set

建模方法	PCR	PLSR	PCA-ANN	PLS-ANN	PCA-SVM
正确率/%	31.56	75.42	91.65	97.31	100

的特征 $\Delta\lambda$ 为 60 nm, 特征 $\Delta\lambda$ 下的样品原始同步荧光光谱经过主成分分析后, 结合支持向量机算法建立了泔水油的同步荧光光谱鉴别模型。该方法简单、准确, 为泔水油的快速鉴别与食用油质量控制提供了新的思路。

3 结 论

采用三维同步荧光光谱结合平行因子法解析得到泔水油

References

- [1] DAI Xiao-rong, LIU Bo-yun(代小容, 刘伯云). Cereals & Oils(粮食与油脂), 2008, (4): 36.
- [2] ZHANG Rui, ZU Li-ya, FAN Tie(张蕊, 祖丽亚, 樊铁). China Oils and Fats(中国油脂), 2006, 31(5): 65.
- [3] HUANG Dao-ping, PENG Jin, XIE Yan-xiang, et al(黄道平, 彭进, 谢燕湘, 等). Chinese Journal of Health Laboratory Technology(中国卫生检验杂志), 2006, 16(2): 151.
- [4] WANG Yan, FAN Lu, HUO Quan-gong, et al(王艳, 范璐, 霍权恭, 等). Journal of Henan University of Technology • Natural Science Edition(河南工业大学学报 • 自然科学版), 2010, (6): 36.
- [5] FANG Hui-min(方惠敏). Journal of Biology(生物学杂志), 2009, 26(6): 83.
- [6] JIA Yan-hua, XU Xiao-xuan(贾艳华, 徐晓轩). Acta Photonica Sinica(光子学报), 2006, 35(11): 1717.
- [7] LIU Hai-long, WU Xi-jun, TIAN Guang-jun(刘海龙, 吴希军, 田广军). Chinese Journal of Lasers(中国激光), 2008, 35(5): 685.
- [8] LÜ Gui-cai, ZHAO Wei-hong, WANG Jiang-tao(吕桂才, 赵卫红, 王江涛). Chinese J. Anal. Chem.(分析化学), 2010, 38(8): 1144.
- [9] Tan C, Li M L, Qin X. Analytical Sciences, 2008, 24(5): 647.
- [10] XU Lu, SHAO Xue-guang(许禄, 邵学广). Methods of Chemometrics(化学计量学方法). Beijing: Science Press(北京: 科学出版社), 2004. 130.
- [11] YAO Wei-feng, HU Yu-zhu, MOU Ling-li(姚卫峰, 胡育筑, 牟玲丽). Chinese J. Anal. Chem.(分析化学), 2009, 37(3): 383.
- [12] Hsu C W, Chang C C, Lin C J. A Practical Guide to Support Vector Classification. <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>, 2012. 2.

Rapid Identification of Hogwash Oil by Using Synchronous Fluorescence Spectroscopy

SUN Yan-hui, AN Hai-yang, JIA Xiao-li, WANG Juan

School of Bio & Food Engineering, Chuzhou University, Chuzhou 239012, China

Abstract To identify hogwash oil quickly, the characteristic $\Delta\lambda$ of hogwash oil was analyzed by three dimensional fluorescence spectroscopy with parallel factor analysis, and the model was built up by using synchronous fluorescence spectroscopy with support vector machines (SVM). The results showed that the characteristic $\Delta\lambda$ of hogwash oil was 60 nm. Collecting original spectrum of different samples under the condition of characteristic $\Delta\lambda$ 60 nm, the best model was established while 5 principal components were selected from original spectrum and the radial basis function (RBF) was used as the kernel function, and the optimal penalty factor C and kernel function g were 512 and 0.5 respectively obtained by the grid searching and 6-fold cross validation. The discrimination rate of the model was 100 % for both training sets and prediction sets. Thus, it is quick and accurate to apply synchronous fluorescence spectroscopy to identification of hogwash oil.

Keywords Hogwash oil; Synchronous fluorescence spectroscopy; Parallel factor analysis; Support vector machines; Identification