

小波矩结合三维荧光光谱简单快速鉴别掺伪芝麻油

潘 钊¹, 李瑞航¹, 吴希军^{1*}, 崔耀耀²

1. 燕山大学河北省测试计量技术及仪器重点实验室, 河北 秦皇岛 066004
2. 燕山大学信息科学与工程学院, 河北 秦皇岛 066004

摘 要 芝麻油营养丰富, 因市场价格较高, 掺假现象频出, 严重损害了消费者利益和市场的健康发展。因此, 研发一种简单快速准确鉴别掺伪芝麻油的方法, 对保障消费者权益和市场健康具有重要意义。为此, 提出了一种小波矩结合三维荧光光谱掺伪芝麻油鉴别方法。该方法简单快速, 计算样本的任一有效特征进行谱系聚类, 即可准确鉴别掺伪芝麻油。以43个样本(芝麻油16个, 掺伪菜籽油、掺伪大豆油及掺伪玉米油各9个)为研究对象, 用FS920荧光光谱仪获得样本的三维荧光光谱。用db2小波将光谱进行多尺度分解(MRSD), 用MRSD的一阶离散逼近系数构造小波矩。用前两阶小波矩值 $W_{0,0}$, $W_{1,0}$, $W_{1,1}$, $W_{0,1}$, $W_{2,0}$, $W_{2,1}$, $W_{2,2}$, $W_{1,2}$, $W_{0,2}$ 分别作为特征对样本进行谱系聚类, 观察分析聚类结果。结合邓恩分类指数(DVI)进一步分析, 研究同阶小波矩分类效果及规律。进而研究各阶小波矩的分类效果及规律。最终确定了用于鉴别掺伪芝麻油的最佳小波矩值。结果表明: MRSD一阶逼近重构光谱可以在保留原光谱的有效特征基础上, 大量去除噪声, 减少光谱数据量72.4%, 增强模型的抗噪稳定性和实时性。利用小波矩前两阶矩值 $W_{2,1}$, $W_{2,2}$, $W_{1,2}$, $W_{0,2}$ 其一作为分类特征进行谱系聚类, 即可鉴别掺伪芝麻油。同阶小波矩($W_{p,q}$)随 p 值减小 q 值增大呈现规律性, 确定了同阶小波矩的有效矩值及最佳有效矩值。小波矩随着阶数的增加DVI先增后减, 最后趋于稳定, 确定了各阶小波矩中可用于鉴别掺伪芝麻油的目标矩值 $W_{0,q \geq 2}$ 及最佳目标矩值 $W_{0,6}$ 。小波矩的有效及目标矩值是针对样本分类的有效特征, 计算样本的任一有效特征进行谱系聚类, 即可实现掺伪芝麻油的鉴别。该研究思路及结论为矩值法应用到三维荧光光谱提供参考。该方法简单快速, 可实现在线测量, 为质监部门及生产企业提供油品检测和鉴定手段。

关键词 芝麻油; 三维荧光; 小波矩; 多尺度分析; 谱系聚类; 邓恩指数

中图分类号: O657.3 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2020)05-1547-07

引 言

芝麻油又称香油, 含有丰富的营养。因市场价格较高, 掺假现象严重。违法商贩用菜籽油、大豆油、玉米油等廉价油勾兑芝麻香精降低成本进行销售, 严重损害了消费者权益和市场健康^[1], 因此鉴别掺伪芝麻油具有重要应用意义。

掺假芝麻油与芝麻油气味、颜色等物理特性相近, 无法肉眼区分。目前植物油的鉴别方法主要有红外光谱法^[1-2]、拉曼光谱法^[3]、色谱法^[4]、荧光光谱法^[5]及太赫兹法^[6]等。荧光光谱法以其灵敏度高、操作简单、分析快速等优点在植物油鉴别领域具备一定的优势^[5]。经典荧光分析方法结合主

成分分析^[7]、平行因子^[8]、交替三线性分解^[9]及偏最小二乘^[10]等算法提取或分离荧光光谱的特征进行定性或定量分析取得了很好的效果。近年来, 图像矩法被引用到三维光谱, 被证明是定性或定量分析的强大工具^[11-13]。但目前三维光谱领域, 有关图像矩法的相关报道较少, 且缺乏较为全面的定性或定量的规律性分析。图像矩是从整个图像空间计算得到, 反映目标的全局信息, 包含丰富的光谱信息。通过找到样本分类的有效矩值, 利用有效矩值即可将样本准确分类。矩值法计算量大, 高阶矩严重敏感于噪声是矩值法在应用时的弊端^[11]。

本文将小波矩引入三维荧光光谱的定性分析。将三维荧光光谱视为灰度图, 利用小波矩提取光谱形状的几何信息。

收稿日期: 2019-04-04, 修订日期: 2019-08-12

基金项目: 国家自然科学基金项目(61471312), 河北省自然科学基金项目(F2015203072), 河北高等学校科学技术研究项目(QN2018071)和燕山大学基础研究专项课题(16LGA008)资助

作者简介: 潘 钊, 1982年生, 燕山大学仪器科学与工程系副教授 e-mail: panzh_zach@hotmail.com

* 通讯联系人 e-mail: wuxijun@ysu.edu.cn

利用多尺度分解 (multiresolution signal decomposition, MRSD) 一阶逼近重构光谱构造小波矩, 在一定程度上克服了矩值法计算量大、高阶矩受噪声影响严重的弊端, 实现了掺伪芝麻油简单快速鉴别。研究小波前两阶矩值的分类效果, 结果表明, 前两阶小波矩即可鉴别掺伪芝麻油; 结合邓恩分类指数 (Dunn's cluster validity index, DVI), 研究了同阶小波矩分类效果及规律, 确定了同阶矩中掺假芝麻油定性分析的有效及最佳矩值; 进而研究各阶小波矩的分类效果及规律, 确定了可以用于掺伪麻油鉴别的各阶矩中的目标矩值及最佳目标矩值。所确定的有效矩值是针对掺伪芝麻油定性分析的有效特征, 计算任一有效特征, 进行谱系聚类, 即可鉴别掺伪芝麻油。该方法简单、快速, 可实现在线测量。文中研究思路及结论, 为植物油特征提取和分类鉴别提供参

考, 为保障消费者权益和市场健康提供手段。

1 实验部分

从市场随机购买不同品牌、同品牌不同批次芝麻油样本 Z_1 — Z_{16} 、菜籽油样本 C_1 、玉米油样本 Y_1 、大豆油样本 D_1 和芝麻香精样本 J_1 。样品质量均由品牌商保证。在实验室将菜籽油、玉米油、大豆油样本分别与芝麻香精样本按不同体积比例 (10%~90%) 混合, 得到菜籽油掺伪样本 JC_1 — JC_9 、玉米油掺伪样本 JY_1 — JY_9 和大豆油掺伪样本 JD_1 — JD_9 。将所有样本混合后放在搅拌机上搅拌 10 min, 而后立刻进行荧光光谱采集。样本信息如表 1 所列。

表 1 样本信息表

Table 1 Reagents and samples

		Label	Brand	Label Origin					
Z_1	Huarui	Z_{11}	Lihong	JC_1	$J_1 : C_1 = 9 : 1$	JD_1	$J_1 : D_1 = 9 : 1$	JY_1	$J_1 : Y_1 = 9 : 1$
Z_2	Fulinmen	Z_{12}	Mingxuan	JC_2	$J_1 : C_1 = 8 : 2$	JD_2	$J_1 : D_1 = 8 : 2$	JY_2	$J_1 : Y_1 = 8 : 2$
Z_3	Damingyongzhen	Z_{13}	Damingyongzhen	JC_3	$J_1 : C_1 = 7 : 3$	JD_3	$J_1 : D_1 = 7 : 3$	JY_3	$J_1 : Y_1 = 7 : 3$
Z_4	Lihong	Z_{14}	Huarui	JC_4	$J_1 : C_1 = 6 : 4$	JD_4	$J_1 : D_1 = 6 : 4$	JY_4	$J_1 : Y_1 = 6 : 4$
Z_5	Mingxuan	Z_{15}	Damingyongzhen	JC_5	$J_1 : C_1 = 5 : 5$	JD_5	$J_1 : D_1 = 5 : 5$	JY_5	$J_1 : Y_1 = 5 : 5$
Z_6	Fulaiwei	Z_{16}	Luhua	JC_6	$J_1 : C_1 = 4 : 6$	JD_6	$J_1 : D_1 = 4 : 6$	JY_6	$J_1 : Y_1 = 4 : 6$
Z_7	Damingyongzhen	J_1	Shangwei	JC_7	$J_1 : C_1 = 3 : 7$	JD_7	$J_1 : D_1 = 3 : 7$	JY_7	$J_1 : Y_1 = 3 : 7$
Z_8	Fulaiwei	C_1	Jinding	JC_8	$J_1 : C_1 = 2 : 8$	JD_8	$J_1 : D_1 = 2 : 8$	JY_8	$J_1 : Y_1 = 2 : 8$
Z_9	Flinmen	Y_1	Wannian	JC_9	$J_1 : C_1 = 1 : 9$	JD_9	$J_1 : D_1 = 1 : 9$	JY_9	$J_1 : Y_1 = 1 : 9$
Z_{10}	Jinlongyu	D_1	Fulinmen						

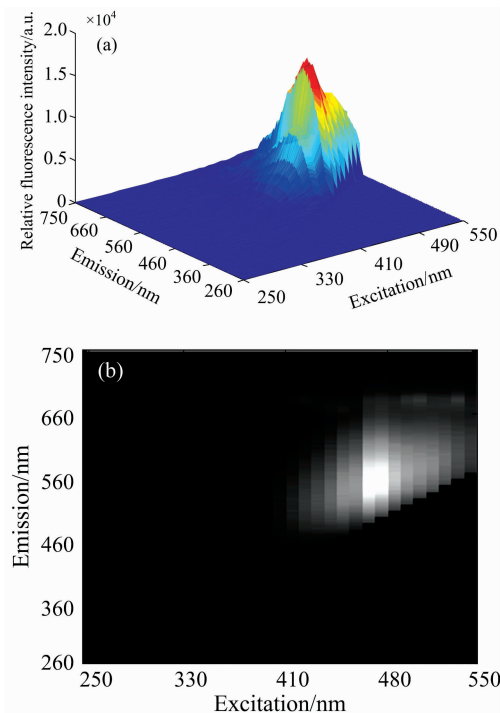


图 1 Z_1 的三维荧光光谱 (a) 及灰度图 (b)

Fig. 1 The 3D fluorescence spectrum (a) and grayscale map (b) of Z_1

设定 FS920 荧光光谱仪激发区间 250~550 nm, 步长为 10 nm, 发射波长 260~750 nm, 步长为 2 nm, 狭缝宽度 1.1 mm, 对应光谱分辨率为 2 nm, 采集得到 43 个样本三维光谱数据 (246×31×43)。每个样本的荧光光谱采集三次平均得到最终结果。将获得的荧光光谱进行激发发射校正预处理。数据操作皆在 MATLAB R2014a 环境下完成。芝麻油 Z_1 三维荧光光谱图及对应灰度图如图 1 所示。

1.1 方法

1.1.1 小波多尺度分析

小波多尺度分析的严格和完整的介绍参见文献[14]。以下作理论简述:

设 $\{V\}_{j \in Z}$ 是 $L^2(R^2)$ 的一个二维多尺度分析 (multi-resolution analysis, MRA), 尺度函数为 φ , 二维尺度空间为 $V_j = V_j \otimes V_j$, 定义小波空间 W_j , $V_{j+1} = V_j \otimes W_j$ 。定义三个小波函数:

$$\Psi_{j,k,m}^1(x,y) = \phi_{j,k}(x)\Psi(j,m)(y) \quad (k,m) \in Z^2 \quad (1)$$

$$\Psi_{j,k,m}^2(x,y) = \phi_{j,k}(x)\Psi(j,m)(y) \quad (k,m) \in Z^2 \quad (2)$$

$$\Psi_{j,k,m}^3(x,y) = \phi_{j,k}(x)\Psi(j,m)(y) \quad (k,m) \in Z^2 \quad (3)$$

构成小波子空间 W_j 规范正交基, 其收缩平移是 $L^2(R^2)$ 的一个标准正交基。将二维图像 $f(x,y) \in L^2(R^2)$ 多尺度小波分解得到

$$f(x,y) = A_j + \sum_{k \geq j} (D_k^1 + D_k^2 + D_k^3) \quad (4)$$

式中

$$A_j = \sum_{(m,n)} a_{m,n}(j) \phi_{m,n}(j), a_{m,n}(j) \leq f(x,y), \phi_{m,n}(j) \geq 0$$

$$D_k^1 = \sum_{(m,n)} d_{m,n}^1(k) \Psi_{m,n}^1(k), d_{m,n}^1(k) \leq f(x,y), \Psi_{m,n}^1(k) > 0$$

$$D_k^2 = \sum_{(m,n)} d_{m,n}^2(k) \Psi_{m,n}^2(k), d_{m,n}^2(k) \leq f(x,y), \Psi_{m,n}^2(k) > 0$$

$$D_k^3 = \sum_{(m,n)} d_{m,n}^3(k) \Psi_{m,n}^3(k), d_{m,n}^3(k) \leq f(x,y), \Psi_{m,n}^3(k) > 0$$

A_j 是小波变换尺度 j 下的离散逼近, $D_k^i, i=1,2,3$ 是小波变换水平、垂直和对角三个方向的离散细节。

1.2 小波矩

将样本光谱矩阵的相对荧光强度看作灰度值, 则荧光光谱数据矩阵可表示为一幅灰度图。对于二维平面上的灰度图像 $f(x,y)$, 它的 $p+q$ 阶几何矩定义为

$$M_{pq} = \sum_{x,y \in z} x^p y^q f(x,y), p,q \in N \quad (5)$$

如果共轭滤波器 $H(\omega)$ 在零点处有 $R+1$ 次平滑性。将图像 $f(x,y)$ 在尺度函数 $\phi(x,y)$, 小波函数 $\Psi(x,y)$ 上展开, 其逼近系数为 $\{a_{m,n}(j)\}$, 则当 $p+q \leq R$ 时, 以小波变换逼近系数所表达的 $p+q$ 阶小波矩成立^[15]

$$W_{pq} = 2^{-(p+q+1)j} \sum_{m,n \in z} m^p n^q a_{m,n}(j), p,q \in N \quad (6)$$

1.3 谱系聚类

谱系聚类又称层次聚类, 是效果较好、经常使用的聚类算法之一。其基本原理如下^[16]:

设待分类的模式特征矢量为 $\{x_1, x_2, \dots, x_N\}$, G_N 表示第 k 次合并时的第 i 类。则谱系聚类算法基本步骤为:

① 初始分类, 令 $k=0$, 每个模式自成一类, 即 $G_i^{(0)} = \{x_i\}, i=1,2,\dots,N$ 。

② 计算各类间距离 D_{ij} , 由此生成一个对称的距离矩阵 $D^{(k)} = (D_{ij})_{m \times m}$, 其中 m 为类的个数(初始时 $m=N$)。

③ 找出前一步求得的矩阵 $D^{(k)}$ 中最小的元素, 设它是 $G_i^{(k)}$ 和 $G_j^{(k)}$ 间的距离, 则将两类合并为一类, $G_i^{(k+1)}, G_j^{(k+1)}$, ..., 于是产生新的聚类, 令 $k=k+1, m=m-1$ 。

④ 检查类的个数, 如果 m 大于 2, 转至步骤②; 否则停止。

1.4 模型评价

结合邓恩分类指数(Dunn's cluster validity index, DVI) 评价聚类结果^[17]。设样本被分为 K 类, 分别用 $Y_1, Y_2, \dots, Y_m, \dots, Y_n, \dots, Y_K$ 各表示所划分的 K 个样本空间, 那么 DVI 可表示为

$$DVI = \frac{\min_{0 < m \neq n < K} \{ \min_{\forall x_i \in Y_m} \{ \|x_i - x_j\| \} \}}{\max_{0 < m < K} \max_{\forall x_i, x_j \in Y_m} \{ \|x_i - x_j\| \}} \quad (7)$$

即, DVI 越大, 则聚类样本的类间距越大, 同时类内距越小, 整体分类效果越好。

2 结果与讨论

用 db2 小波将荧光光谱进行 MRSD 分解。利用任一尺度逼近皆可用来构造小波矩, 但由于存在吉布斯效应, 随着分解尺度的增加会产生积分误差^[15]。因此, 选择第一尺度逼近系数重构光谱。为了增强实时性, 将重构光谱归一化到 0~1 范围内。以芝麻油 Z_1 为例, 重构光谱及原光谱的指纹图如图 2(a)和(b)所示, 发射图如图 2(c)和(d)所示。

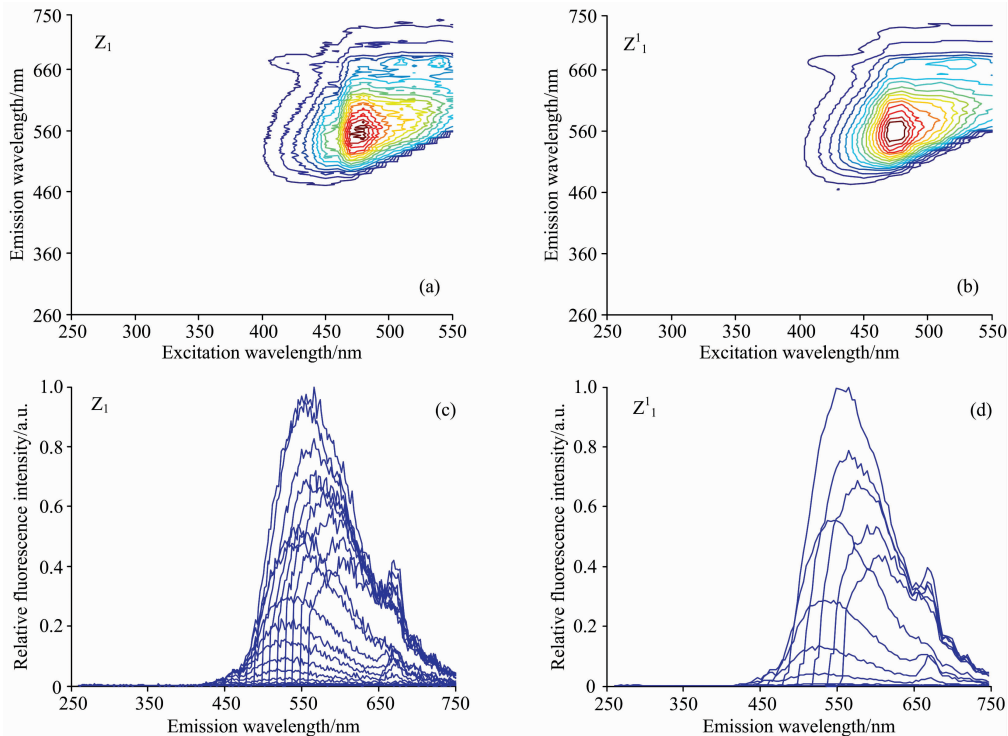


图 2 Z_1 的光谱图和第一尺度重构光谱图

Fig. 2 The original spectrum and the reconstructed spectrum of Z_1

如图 2(a)–(d)所示: Z_1' 与 Z_1 的主峰值个数、峰值中心等形状特征高度一致; Z_1' (124×17) 较 Z_1 (246×31) 数据量减少 72.4%; Z_1' 发射图曲线较 Z_1 平滑, 含噪量少。MRSD 的一阶逼近重构光谱可以在保留原光谱的有效形状特征基础上去

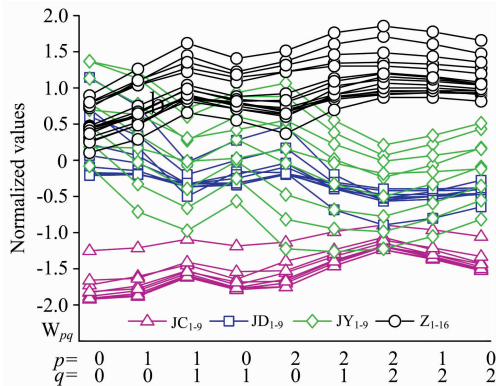


图 3 样本矩集图示 ($W_{p \leq 2, q \leq 2}$)

Fig. 3 The graphical of the moments ($W_{p \leq 2, q \leq 2}$)

除噪声, 大量减少计算量, 增强模型的抗噪稳定性及实时性。

小波矩提取光谱形状的几何信息, 不同的小波矩值表征光谱不同层面的信息。构造样本的前两阶矩值小波矩: $W_{0,0}$, $W_{1,0}$, $W_{1,1}$, $W_{0,1}$, $W_{2,0}$, $W_{2,1}$, $W_{3,2}$, $W_{1,2}$, $W_{0,2}$ 。将矩值 Z-score 标准化, 图示化如图 3 所示: 芝麻油和掺伪芝麻油样本矩值特征 $W_{0,0}$, $W_{1,0}$, $W_{1,1}$, $W_{0,1}$, $W_{2,0}$ 存在重叠, 无法通过谱系聚类法分类; 而矩值 $W_{2,1}$, $W_{2,2}$, $W_{1,2}$, $W_{0,2}$ 相互分离, 存在鉴别的可能性; 随着矩阶数增加, 芝麻油与掺伪样本类间距有增加趋势。

定义样本间距欧式距离, 类间距离差平方和。分别用九个矩值将样本进行谱系聚类。结果表明: 矩值 $W_{0,0}$, $W_{1,0}$, $W_{1,1}$, $W_{0,1}$, $W_{2,0}$ 无法鉴别掺伪样本; 矩值 $W_{2,1}$, $W_{2,2}$, $W_{1,2}$, $W_{0,2}$ 皆可鉴别掺伪样本。以 $W_{2,2}$ 为例, 其聚类树形图如图 4 所示: 芝麻油在 1.26 上聚成一类, 三类掺伪芝麻油在 2.78 上聚成一类, 合并两类则需 8.45 以上。结果表明, 小波矩可以提取样本的有效分类特征, 可实现掺伪芝麻油的分类鉴别。计算前两阶小波矩 $W_{2,1}$, $W_{2,2}$, $W_{1,2}$, $W_{0,2}$ 中任一矩值作为特征进行谱系聚类, 即可鉴别掺伪芝麻油。

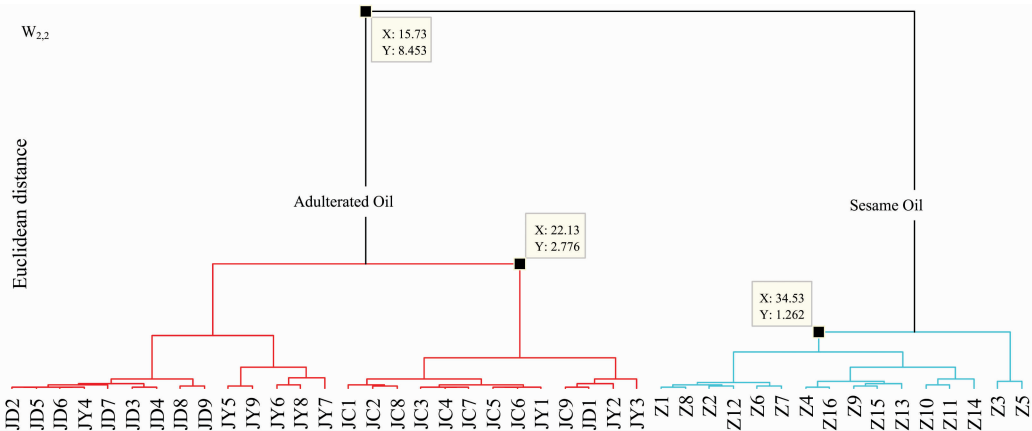


图 4 样本的聚类树形图 ($W_{2,2}$)

Fig. 4 The cluster tree of the samples ($W_{2,2}$)

随着矩阶数的增加, 芝麻油和掺伪芝麻油矩值特征的类间距有增加趋势(图 3)。低阶小波矩提取光谱的整体性特征, 高阶矩提取细节特征。阶数低, 提取的细节信息量不足, 但阶数过高, 也会因高阶矩敏感于噪声和微小误差造成分类效果差。

2.1 同阶小波矩值分类效果及规律

同阶小波矩组成的矩集表征着光谱的同一几何信息, 其所包含的各矩值为该几何信息的分量。计算样本各阶小波矩集进行谱系聚类, 观察聚类结果, 并计算 DVI。以十阶矩各矩值 DVI 为例, 十阶小波矩各矩值图示化如图 5(a)所示, 其 DVI 值如图 5(b)所示。

结果表明, 当 p , q 如图 5 中方式排列时, 同阶小波矩分类效果呈现规律性, 但低阶(阶数 $n \leq 4$)和高阶(阶数 $n \geq 5$)规律不一致。高阶: 如图 5(b), 同阶矩 DVI 随着 q 的增大, p 的减小有增大趋势, 即相同阶矩 q 值越大分类效果越好, p 值越小分类效果越好。 p 值过大同时 q 值小会造成分类失败, 失败案例如图 5(b)中所示。其他高阶小波矩在谱系聚类

时有相同规律。矩值 $W_{p \leq 10, q \geq 2}$ 可以鉴别掺伪芝麻油称有效矩值, 其中最佳有效矩值为 $W_{0,10}$ 。 $W_{0,n}$ ($n \geq 5$) 为高阶小波矩定性分析的最佳表征。低阶: 同阶矩 DVI 随着 q 的增大, p 的减小呈先增后减的趋势。0 阶和 1 阶无有效矩值。2, 3 和 4 阶的最佳有效矩值分别为 $W_{2,2}$, $W_{3,3}$, $W_{2,4}$ 。

2.2 不同阶矩值分类效果及规律

不同阶矩值表征着光谱不同层面的几何信息。计算各阶小波矩的最佳有效矩值, 以各阶最佳有效矩值为特征将样本进行谱系聚类, 并计算 DVI。矩值图示化如图 6(a)所示, DVI 值如图 6(b)所示。

如图 6(a)所示, 随着矩阶数的增加, 样本 Z_3 和 Z_5 的矩值特征与其他芝麻油样本差异逐渐增大, 导致高阶小波矩分类效果不佳。高阶矩提取细节特征, 对噪声和微小差异存在放大作用。如图 6(b)所示, DVI 随着阶数的增加先增加后减小, 最后趋于稳定。随着阶数的增加, 分类效果渐佳, 但阶数过高会造成样本 Z_3 和 Z_5 与其他芝麻油样本差异化, 造成分类效果变差。 $W_{0,q} \geq 2$ 可以鉴别掺伪芝麻油, 称为目标矩

值,其中 $W_{0.6}$ 为最佳目标矩值。 $W_{0.6}$ 聚类树形图如图 7 所示:芝麻油 1.24 以上聚成一类,三类掺伪芝麻油 1.13 以上聚成一类,合并两类则需 8.94 以上。文中所确定的小波矩的

有效及目标矩值是鉴别掺伪芝麻油样本的有效特征,直接计算样本的任一有效特征进行谱系聚类,即可准确鉴别掺伪芝麻油。

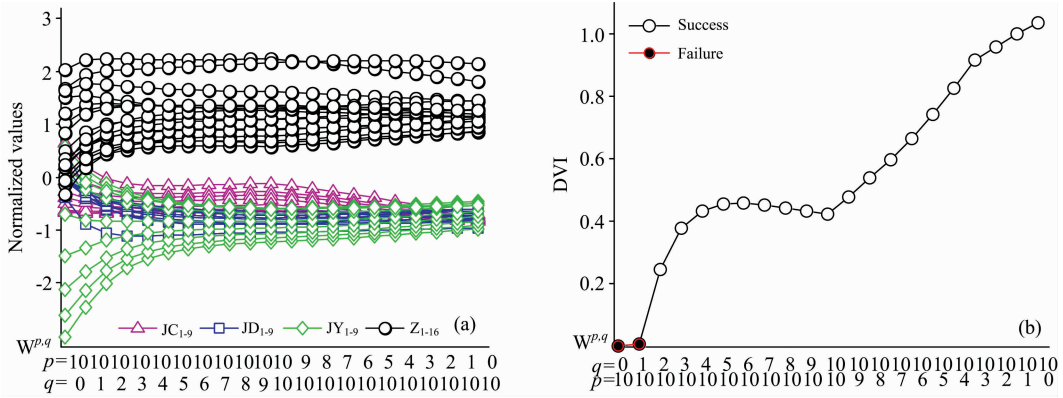


图 5 十阶小波矩各矩值分类效果及规律

Fig. 5 The DVI of the tenth order wavelet moments

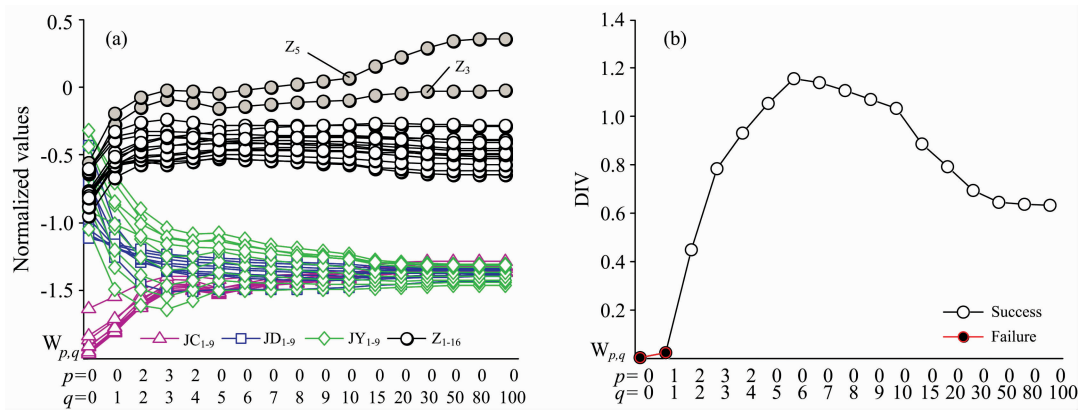


图 6 不同阶小波矩分类效果及规律

Fig. 6 The DVI of different order wavelet moments

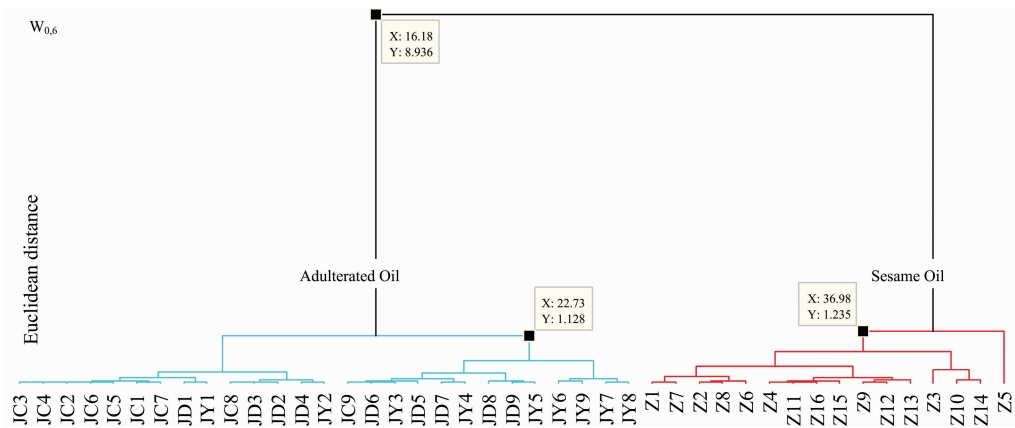


图 7 各阶小波矩中最佳有效矩值 $W_{0.6}$ 聚类树形图

Fig. 7 The cluster tree of optimal target WMs ($W_{0.6}$)

3 结 论

提出将小波矩应用到三维荧光光谱,以有效矩值为特征

进行谱系聚类,实现了掺伪芝麻油样本的鉴别。以 MRSD 的一阶逼近重构光谱,去除了噪声,减少光谱数据量 72.4%,增加了矩值法的抗噪稳定性和实时性。通过研究同阶矩及各阶矩分类效果及规律,确定了可以鉴别掺伪芝麻油的小波矩

值。该方法简单快速有效,利用任一有效特征即可鉴别掺伪麻油。研究思路及结论为矩值法应用到三维荧光光谱提供参

考。为质监部门及生产企业提供油品检测和鉴定手段。

References

- [1] Gorkem Ozulku, Rusen Metin Yildirim, Omer Said Toker, et al. *Food Control*, 2017, 82: 212.
- [2] Chen H, Lin Z, Tan C. *Vibrational Spectroscopy*, 2018, 99: 178.
- [3] Ana M Jiménez-Carvelo, María Teresa Osorio, et al. *LWT*, 2017, 86: 174.
- [4] Jiménez-Carvelo, Ana M, González-Casado A, et al. *Talanta*, 2017, 164: 540.
- [5] Xu J, Liu X F, Wang Y T. *Food Chemistry*, 2016, 212: 72.
- [6] Liu J, Fan L, Liu Y, et al. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2019, 206: 165.
- [7] Ranamukhaarachchi S A, Peiris R H, Moresoli C. *Food Chemistry*, 2017, 217: 469.
- [8] Spagnuolo M L, Marini F, Sarabia L A, et al. *Talanta*, 2017, 167: 367.
- [9] Wang L, Wu H L, Yin X L, et al. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2017, 170: 104.
- [10] Martín Tornero Elisabet, Espinosa-Mansilla Anunciación, Arsenio M D L P, et al. *Talanta*, 2018, 188: 299.
- [11] Chen J, Li B Q, Zhai H L, et al. *Journal of Chromatography A*, 2014, 1352: 55.
- [12] Chen J, Li B Q, Xu M L, et al. *Talanta*, 2016, 161: 99.
- [13] WU Xi-jun, CUI Yao-yao, PAN Zhao, et al(吴希军, 崔耀耀, 潘 钊, 等). *Spectroscopy and Spectral Analysis(光谱学与光谱分析)*, 2018, 38(8): 2456.
- [14] Mallat S. *IEEE Trans PAMI*, 1989, 11(7): 674.
- [15] JIN Qi, DAI Ru-wei(金 琪, 戴汝为). *Pattern Recognition and Artificial Intelligence(模式识别与人工智能)*, 1995, (3): 179.
- [16] SUN Ji-xiang(孙即祥). *Modern Pattern Recognition(现代模式识别)*. Beijing: Higher Education Press(北京: 高等教育出版社), 2008. 40.
- [17] Dunn J C. *J. Cybernetics*, 1973, 3(3): 32.

Simple and Quick Identification of Adulterated Sesame Oil by Wavelet Moments and Three-Dimensional Fluorescence Spectra

PAN Zhao¹, LI Rui-hang¹, WU Xi-jun^{1*}, CUI Yao-yao²

1. Key Lab of Measurement Technology and Instrumentation of Hebei Province, Yanshan University, Qinhuangdao 066004, China

2. School of Information Science and Engineering, Yanshan University, Qinhuangdao 066004, China

Abstract Sesame oil is rich in nutrients. Due to the high market price, adulteration is frequent, which seriously damages the interests of consumers and the healthy development of the market. Therefore, the development of a fast and accurate method for the identification of adulterated sesame oil (ASO) is of great significance for protecting consumer rights and the market health. To this end, this paper proposed a method for identifying ASO with wavelet moments (WMs) combined with three dimensional fluorescence spectra (3DFS). This method is simple and rapid, and can effectively identify ASO. In the article, Taking 43 samples (16 sesame oil, 9 kinds of rapeseed adulteration sesame oil, soybean adulteration sesame oil and corn adulteration sesame oil, respectively) as research objects. The main research contents and results are as follows: The 3DFS of the samples were obtained using a FS920 fluorescence spectrometer. Multiresolution signal decomposition (MRSD) was performed on the spectra using db2 wavelets, and then the 3DFS was reconstructed using the first-order discrete approximation coefficients of MRSD. The first two orders of WMs: $W_{0,0}$, $W_{1,0}$, $W_{1,1}$, $W_{0,1}$, $W_{2,0}$, $W_{2,1}$, $W_{2,2}$, $W_{1,2}$, $W_{0,2}$, were separately used as feature to perform hierarchical clustering (HC) on the samples. Next, combined with Dunn's cluster validity index (DVI), the classification quality and laws of the same-order and different-order WMs were studied, and the optimal WMs for identifying ASO were determined. Results: MRSD can remove noise and reduce the amount of spectral data by 72.4% on the basis of retaining the effective characteristics of the original spectra. To a certain extent, it can overcome the disadvantages of moment methods that large computational complexity and high-order moments are seriously affected by noise. Using one of $W_{2,1}$, $W_{2,2}$, $W_{1,2}$, $W_{0,2}$ to perform HC as a feature, the ASO can be easily and quickly identified. The same-order WMs ($W_{p,q}$) exhibit regularity as the p decreases q increases, and the effective WMs (EWMs) of the same order were determined. The target moments (TMs) $W_{0,q} \geq 2$ and the

optimal target moment $W_{0.6}$ which can be used to identify ASO were determined. Simple and efficient identification of ASO can be achieved by computing HC with any determined WMs. This method can be extended to online measurements. The research ideas and conclusions provide a reference for the identification and classification of vegetable oils, and provide a means to protect consumer rights and market health.

Keywords Sesame oil; Three-dimensional fluorescence spectra; Wavelet moments; Multiresolution signal decomposition; Hierarchical clustering; Dunn's cluster validity index

(Received Apr. 4, 2019; accepted Aug. 12, 2019)

* Corresponding author