

# 基于多源光谱分析技术的鱼油品牌判别方法研究

张瑜<sup>1,2</sup>, 谈黎虹<sup>1</sup>, 曹芳<sup>2</sup>, 何勇<sup>2</sup>

(1. 浙江经济职业技术学院, 浙江杭州 310018) (2. 浙江大学生物系统工程与食品科学学院, 浙江杭州 310058)

**摘要:** 多源光谱分析技术被用于鱼油品牌快速无损鉴别。采用可见光谱分析技术、短波近红外光谱分析技术、长波近红外光谱分析技术、中红外光谱分析技术和核磁共振光谱分析技术采集了 7 种不同品牌的鱼油的光谱特征, 并应用偏最小二乘判别分析法 (partial least squares discrimination analysis, PLS-DA) 和最小二乘支持向量机 (least-squares support vector machine, LS-SVM) 建立判别模型并比较判别结果。基于长波近红外光谱的 PLS-DA 模型和 LS-SVM 模型取得了最高识别正确率, 建模集和预测集识别正确率均达到 100%。采用中红外光谱和核磁共振谱分别建立的 LS-SVM 模型, 也可以获得 100% 的判别正确率。而可见光谱和短波近红外光谱则判别准确率较差。且 LS-SVM 算法较 PLS-DA 更加适合用于建立光谱数据和鱼油品牌之间的判别模型。研究结果表面长波近红外光谱技术能够有效判别不同鱼油的品牌, 为将来鱼油品质鉴定便携式仪器的开发提供了技术支持和理论依据。

**关键词:** 鱼油; 品牌判别; 可见/近红外光谱; 核磁共振; 偏最小二乘判别分析; 最小二乘支持向量机

文章编号: 1673-9078(2014)10-263-267

DOI: 10.13982/j.mfst.1673-9078.2014.10.044

## Study on Brand Discrimination of Fish Oil Based on Multiple Spectroscopy Techniques

ZHANG Yu<sup>1,2</sup>, TAN Li-hong<sup>1</sup>, CAO Fang<sup>2</sup>, HE Yong<sup>2</sup>

(1. Zhejiang Technical Institute of Economics, Hangzhou 310018, China)

(2. College of Biosystems Engineering & Food Science, Zhejiang University, Hangzhou 310058, China)

**Abstract:** In this paper, multiple spectroscopy techniques were used to distinguish different brands of fish oil in a rapid and non-invasive manner. Spectral characteristics of seven brands of fish oil, collected by visible spectroscopy, short wave near infrared spectroscopy (SNIR), long-wave near infrared spectroscopy (LNIR), mid-infrared spectroscopy (MIR), and nuclear magnetic resonance (NMR) spectroscopy, were set as inputs in partial least squares discrimination analysis (PLS-DA) and a least-squares support vector machine (LS-SVM) to establish the discrimination models. The discrimination results of the PLS-DA and LS-SVM models were subsequently compared. The results showed that LNIR achieved the highest discriminant accuracy, and the accuracies of modeling set and prediction set were up to 100%. The LS-SVM model using MIR and NMR spectroscopy also yielded a discriminant accuracy of 100%. On the other hand, the discriminant accuracies of those based on visible spectroscopy and SNIR were poor. In addition, LS-SVM was more suitable than PLS-DA to build identification models for fish oil brands using spectroscopic data. The results indicated that LNIR spectroscopy technique could effectively distinguished fish oil brands, providing the technical support and theoretical basis for developing portable instruments for the analysis of fish oil quality in the future.

**Key words:** fish oil; brand discrimination; visible and near infrared spectroscopy; nuclear magnetic resonance; partial least squares discriminant analysis; least-squares support vector machines

鱼油是含有丰富营养成分的鱼类加工副产品, 鱼油包括鱼体油、鱼肝油和鱼脑油, 并包含多种脂肪酸, 以及维生素 A 和 D、磷脂、甘油醚、类固醇和鱼蛋白降解物等营养成分<sup>[1]</sup>。研究表明, 鱼油中的 EPA 和 DHA 具有功能性保健效果, 对一些疾病具有治疗和缓

收稿日期: 2014-05-08

基金项目: 国家自然科学基金 (31072247)

作者简介: 张瑜 (1977-), 女, 副教授, 主要研究方向为农业工程

通讯作者: 何勇 (1963-), 男, 博士, 教授, 主要研究方向精细农业与农业物联网

解的作用<sup>[2]</sup>。

随着对鱼油的保健和药用功能研究与应用的深入, 鱼油作为商品已经受到越来越多的人的关注和欢迎。市场上出现了许多不同品牌的鱼油, 而不同品牌的鱼油在其营养成分含量上都有较大区别, 且价格也存在较大的差异。鱼油较高的经济价值和药用价值导致部分不法厂商以次充好, 以假乱真, 造成鱼油市场的混乱。而目前, 对鱼油的品质评定与鉴别仅限于实验室中进行, 不仅耗时费力, 也浪费实验试剂, 成本较高。市场上种类繁多的鱼油需要研究快速无损的检

测技术进行品质评定和鉴别。

可见-近红外光谱技术 (Visible/ Short Wave Near Infrared Spectroscopy, Vis/SWNIR), 长波近红外光谱技术 (Long Wave Near Infrared Spectroscopy, NIR) 和中红外光谱技术 (Mid-Infrared Spectroscopy, MIR) 和是目前常见的应用于食品分析的光谱分析技术<sup>[3-5]</sup>。可见光光谱到中红外光谱可以获得物质化学基团相关的信息, 从而得到物体的颜色信息和内部化学成分信息, 结合化学计量学方法实现对物质品质定性或定量的检测<sup>[6-7]</sup>。核磁共振技术 (Nuclear Magnetic Resonance, NMR) 是基于原子核磁性的一种波谱技术<sup>[8]</sup>, 能实现对物质定性定量分析, 在食品中的应用和发展越来越广泛<sup>[9]</sup>。光谱分析方法最主要的特点之一是不需或只需少量的样本处理, 操作简单方便, 且检测速度较快。本文的研究目的有以下 3 点: 1) 研究不同品牌鱼油在不同光谱分析技术下的光谱曲线特征; 2) 采用 PLS-DA 和 LS-SVM 算法建立不同光谱分析技术下鱼油品牌的鉴别模型; 3) 分析并比较算法效果, 挑选对鱼油识别效果较好的分析技术与模型。

## 1 材料与方法

### 1.1 实验样本

美国安利深海鱼油、纽崔莱深海鲑鱼油、汤臣倍健鱼油、汤臣倍健浓缩鱼油、安利心血管鱼油、GNC 双倍浓度深海鱼油和 GNC 三倍浓度深海鱼油等 7 种市场上常见的鱼油品牌被用来作为鱼油样本进行分析。鱼油采集的是保质期内的鱼油以避免储藏时间对实验结果的影响。根据厂商提供的信息, 各个品牌的鱼油的成分含量有一定的差异。且由肉眼观察, 不同品牌的鱼油的颜色存在较小的差异。

### 1.2 数据采集

测量鱼油的可见-短波近红外光谱透射光谱时, 鱼油置于比色皿中, 采用 USB4000 可见-近红外光谱仪 (200~1100 nm) 进行光谱扫描, 每个样本光谱测量 10 次, 取其平均值为鱼油样本的透射光谱, 每个品牌的鱼油采集五个样本的光谱。测量鱼油的长波近红外光谱时, 与可见-短波近红外光谱技术类似, 鱼油置于比色皿中, 采用 NIR256-2.5 光谱仪 (870~2537 nm) 进行光谱扫描, 每个样本光谱测量 10 次, 取其平均值为鱼油样本的透射光谱, 每个品牌的鱼油采集五个样本的光谱。测量鱼油中红外光谱时, 鱼油通过毛细管滴到 ATR 附件的样品槽中, 然后将 ATR 附件旋紧, 中红外光谱由 Nicolet iS10 傅里叶变换 ATR 中红外光

谱仪 (4000~650  $\text{cm}^{-1}$ ) 采集光谱, 每个样本扫描 64 次, 取其平均作为鱼油样本的中红外透射光谱, 与前两种方法一样, 每个品牌的鱼油采集五个样本的光谱。测量鱼油核磁共振谱时, 采用 AVANCE III 500MHz 核磁共振波谱仪测量鱼油  $^1\text{H}$  核磁共振谱, 在 298 K 条件下每个样品扫描 16 次谱线。基于核磁共振技术每个品牌采集 4 个样本。

### 1.3 模型建立和评价标准

偏最小二乘判别分析法 (partial least squares-discriminant analysis, PLS-DA)<sup>[10-11]</sup>和最小二乘支持向量机 (least-squares support vector machine, LS-SVM)<sup>[12-13]</sup>被用来建立判别分析模型时。并利用所建立的模型预测样本的品牌。

偏最小二乘判别分析 (PLS-DA) 是一种基于偏最小二乘回归 (partial least squares regression, PLSR) 算法的有监督判别分析方法。PLS-DA 算法以光谱数据为 X, 代表类别的整数值为 Y (例如在本研究中, 以代表鱼油品牌类别的不同整数作为 Y) 建立 PLSR 模型, 并进行预测, 回归分析模型得到的预测值为实际的数值而不是代表类别的整数, 因此, 在判别分析时需对预测结果进行判定以得到样本归属的类别。一般采用阈值法判断预测样本归属的类别。阈值的选取是任意的, 即对于已知类别的建模集样本与预测集样本, 若实际值与预测值的差值的绝对值小于阈值, 则表明判别正确, 反之则判别错误。对于未知样本, 则以预测值与对应最相近的整数的差值的绝对值判别, 若小于阈值, 则判别为该整数代表的类。一般研究中将阈值设定为 0.5, 以确保对样本类别的判别分析<sup>[10-11]</sup>。

最小二乘支持向量机 (LS-SVM) 是支持向量机 (support vector machine, SVM) 算法的扩展, 可用于求解线性与非线性的问题, 对小样本和高维数据具有较好的求解效果。LS-SVM 是在 SVM 的基础上, 将不等式约束转化为等式约束, 将 SVM 中的二次规划问题转化为求解线性方程组的解, 并能得到唯一最优解。LS-SVM 算法较 SVM 算法计算更为简单, 求解速度更快, 收敛精度高<sup>[12-13]</sup>。本研究中采用 LS-SVM 回归方法, 通过对预测值设定阈值判断样本归属的类别, 与 PLS-DA 相似, 判别阈值设定为 0.5。核函数是 LS-SVM 算法的关键, 本研究中采用径向基核函数 (radial basis function, RBF) 建立 LS-SVM 模型。

建模时, 不同鱼油品牌分别赋予虚拟整数值作为鱼油品牌的标准值, 品牌赋值如表 1 所示。本研究中以建模集和预测集识别正确率作为模型效果和分析方法

的评价指标。识别正确率越高，说明模型和分析方法越好。

表 1 不同品牌鱼油的赋值

**Table 1 Assigned integers assigned to different brands of fish oil**

品牌名称	美国安利 深海鱼油	纽崔莱深 海鲑鱼油	汤臣倍 健鱼油	汤臣倍健 浓缩鱼油	安利心血 管鱼油	GNC 双倍浓 度深海鱼油	GNC 三倍浓 度深海鱼油
赋值	1	2	3	4	5	6	7

## 2 结果与讨论

### 2.1 鱼油的光谱曲线分析

不同光谱分析技术得到的鱼油谱线如图 1~4 所示。为去除光谱明显噪声，可见-短波近红外光谱曲线分析 429~1048 nm 的光谱；长波近红外光谱曲线分析用 1031~2336 nm 范围的光谱数据；对于中红外光谱曲线，同时去除包含明显噪声与无特征吸收峰的波段，仅对 650~1493、1734~1782、2819~3036  $\text{cm}^{-1}$  波段的光谱数据进行分析。分析 NMR 谱线时，去除响应值小于 100 的变量，对剩下的 3636 个 NMR 波谱变量进行分析处理。通过分析鱼油不同谱线可知，不同鱼油的品牌在四种谱线中相互之间没有明显差异且相互重合，表明无法从光谱曲线上实现对鱼油品牌的识别，需要根据鱼油的谱线与品牌赋值建立判别分析模型进行分析。

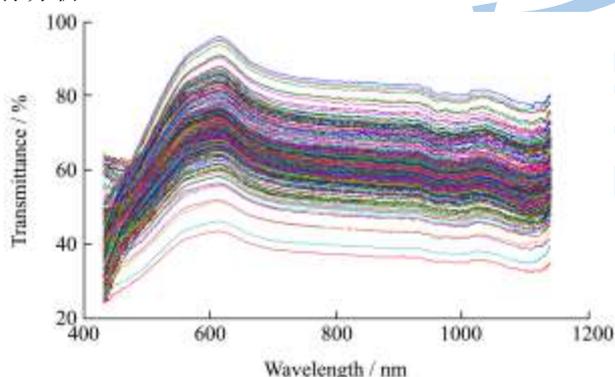


图 1 鱼油的可见-短波近红外光谱  
Fig.1 Vis/SW NIR spectra of fish oil

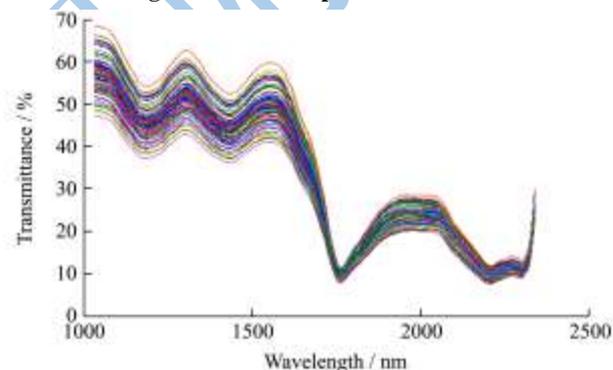


图 2 鱼油的长波近红外光谱  
Fig.2 LW NIR spectra of fish oil

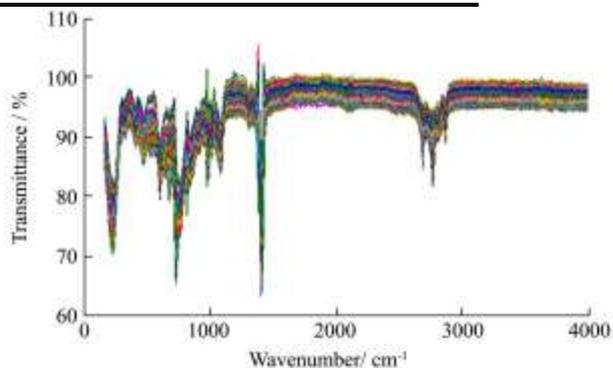


图 3 鱼油的中红外光谱

Fig.3 MIR spectra of fish oil

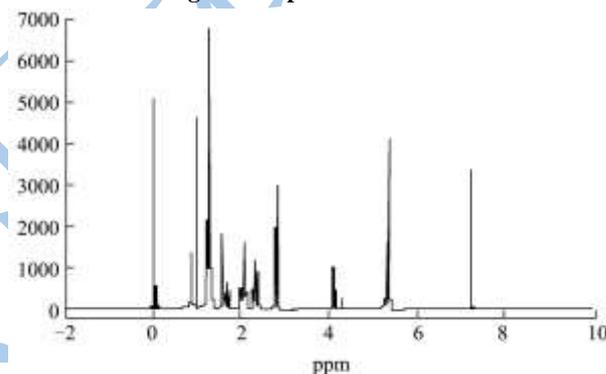


图 4 鱼油的核磁共振谱

Fig.4 NMR spectra of fish oil

### 2.2 偏最小二乘判别分析模型

采用 430~700 nm 范围的可见光谱建立 PLS-DA 判别模型，其建模集样本的判定正确率为 82.14%，而预测集样本为 71.43%，判别正确率不高。其中，GNC 两个不同品牌的样本错判较多，可能是因为可见光区域没有体现出这两个品牌鱼油的差异，不能够通过可见光谱很好地区分。采用 701~1049 nm 的短波近红外光谱建立 PLS-DA 模型时，建模集和预测集样本的判定正确率和可见光谱的 PLS-DA 模型相似，分别为 85.71% 和 71.43%。不同之处在于短波近红外光谱 PLS-DA 模型的错判样本分布在纽崔莱深海鲑鱼油、汤臣倍健鱼油和 GNC 双倍浓度深海鱼油三个品牌。当长波近红外和中红外光谱分别用于建立鱼油品牌判别模型时，两者的建模集判定正确率均为 100%。但是中红外光谱的预测集样本判定正确率只有 85.71%（纽崔莱深海鲑鱼油品牌判别错误），而长波近红外则

为 100%。在所有多源光谱中,核磁共振光谱的 PLS-DA 模型判别结果最差,其中建模集样本为 80.95%,而预测集仅为 57.14%。其中建模集样本有分属美国安利深海鱼油、安利心血管鱼油和 GNC 三倍浓度深海鱼油共四个样本判别错误,而预测集样本则有分属汤臣倍健鱼油、汤臣倍健浓缩鱼油和安利心血管鱼油供三个样本判别错误。综上所述,长波近红外光谱所建立的 PLS-DA 模型对于鱼油品牌判别能力最佳,其次是中红外光谱。

表 2 基于偏最小二判别分析法(PLS-DA)和最小二乘支持向量机(LS-SVM)所建立的可见光谱(Vis)、短波近红外光谱(SWNIR)、长波近红外光谱(LWNIR)、中红外光谱(MIR)和核磁共振光谱(NMR)模型的鱼油品牌识别结果

**Table 2 Identification results of fish oil brands using PLS-DA and LS-SVM models based on the Vis, SWNIR, LWNIR, MIR and NMR spectra**

算法	建模	Vis	SWNIR	LWNIR	MIR	NMR
PLS-DA	Calibration	82.14%	85.71%	100.00%	100.00%	80.95%
	Prediction	71.43%	71.43%	100.00%	85.71%	57.14%
LS-SVM	Calibration	100.00%	100.00%	100.00%	100.00%	100.00%
	Prediction	85.71%	71.43%	100.00%	100.00%	100.00%

### 2.3 最小二乘支持向量机( LS-SVM) 模型

LS-SVM 算法被同样用于分别建立了可见光谱、短波近红外光谱、长波近红外光谱、中红外光谱和核磁共振光谱模型,结果如表 1 所示。其中长波近红外光谱、中红外光谱和核磁共振光谱均获得了满意的判别结果,无论是建模集样本还是预测集样本的判别正确率均为 100%。另一方面,可见光谱和短波近红外光谱的判别结果则较差。虽然两者建模集的判别正确率也达到了 100%,但预测集却分别只有 85.71% 和 71.43%。

综合分析表 2,比较 PLS-DA 算法和 LS-SVM 算法在处理不同光谱分析技术判别鱼油品牌的能力。其中 PLS-DA 是一种线性建模算法,能够用线性关系式建立鱼油光谱数据和品牌编号之间的判别模型;而 LS-SVM 是一种非线性建模算法,其建立的鱼油光谱和品牌之间的判别模型是以非线性关系形式表达。如表 2 所示,LS-SVM 算法较 PLS-DA 算法而言可以获得更好的判别结果。其中中红外光谱模型的预测集判别正确率从 PLS-DA 模型的 85.71% 提高到了 LS-SVM 模型的 100%,而短波近红外光谱模型的建模集判别正确率从 PLS-DA 模型的 85.71% 提高到了 LS-SVM 模型的 100%。另外,建模集和预测集的判别正确率同

时提高的光谱包括可见光谱和核磁共振光谱。其中可将光谱的建模集判别正确率从 PLS-DA 模型的 82.14% 提高到了 LS-SVM 模型的 100%,而其预测集判别正确率从 PLS-DA 模型的 71.43% 提高到了 LS-SVM 模型的 85.71%。核磁共振光谱的判别正确率则通过采用 LS-SVM 算法获得了较 PLS-DA 算法而言大幅度的提高,建模集和预测集判别正确率分别从 80.95% 和 57.14% 提高到了 100% 和 100%。由此可以推断出核磁共振光谱和鱼油品牌之间主要存在非线性关系,因此当采用 PLS-DA 算法时无法获得准确的判别结果,而当采用 LS-SVM 时则获得了令人满意的判别正确率。比较所有本研究采用的光谱波段和建模算法的结果最后发现,有四组模型的建模集样本和预测集样本均能够同时获得 100% 的判别正确率,它们分别是基于 LS-SVM 算法的中红外光谱模型,基于 LS-SVM 算法的核磁共振光谱模型,以及基于 PLS-DA 和 LS-SVM 的长波近红外光谱模型。特别需要注意的是,无论是采用 PLS-DA 还是 LS-SVM 算法,长波近红外光谱均能够获得 100% 的判别结果,说明该波段光谱非常适合于鱼油品牌的鉴别。另一方面,长波近红外光谱的采集也较中红外光谱和核磁共振光谱而言快速、易于操作,并且成本低。

### 3 结论

本研究分别基于 5 种不同的光谱分析技术建立鱼油品牌的 PLS-DA 和 LS-SVM 鉴别模型。长波近红外光谱在鱼油品牌的鉴别中取得了最佳的效果,无论是采用 PLS-DA 算法还是 LS-SVM 算法,其建模集样本和预测集样本均能够获得 100% 的判别正确率。中红外光谱和核磁共振光谱的判别能力较长波近红外光谱次之,当采用 LS-SVM 算法时同样可以获得 100% 的正确率,而当采用 PLS-DA 算法时,判别结果则不甚理想。而可见光谱和短波近红外光谱则无法获得令人满意的判别结果。此外,LS-SVM 算法由于可以建立光谱值和品牌之间的非线性关系,因此获得了较 PLS-DA 更好的判别正确率。上述研究结果表明,利用长波近红外光谱分析技术和数据挖掘算法可以实现鱼油品牌信息的快速识别。

### 参考文献

- [1] 陶宁萍,鲍丹.鱼油的营养和药用价值及其提取工艺的研究进展[J].上海水产大学学报,2005,14(2):197-201  
TAO Ning-ping, BAO Dan. Nutritive value and development of extraction technique on fish oil [J]. Journal of Shanghai Fisheries University, 2005, 14(2): 197-201

- [2] 肖玫,欧志强.深海鱼油中两种脂肪酸(EPA和DHA)的生理功效及机理的研究进展[J].食品科学,2005,26(8):522-526  
XIAO Mei, OU Zhi-qiang. Research progress of the physiological function and mechanism of two kinds of fatty acid (EPA and DHA) in the fish oil of deep sea [J]. Food Science, 2005, 26(8): 522-526
- [3] KIMIYA T, SIVERTSEN, A H, HEIA K, et al. VIS/NIR spectroscopy for non-destructive freshness assessment of atlantic salmon (*Salmosalar L.*) fillets [J]. Journal of Food Engineering, 2013, 116(3): 758-764
- [4] LUNA A S, da SILVA A P, PINHO J S, et al. Rapid characterization of transgenic and non-transgenic soybean oils by chemometric methods using NIR spectroscopy [J]. Spectro chimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2013, 100: 115-119
- [5] AOUIDI F, DUPUY N, ARTAUD J, et al. Rapid quantitative determination of oleuropein in olive leaves (*oleaeuropaea*) using mid-infrared spectroscopy combined with chemometric analyses [J]. Industrial Crops and Products, 2012, 37(1): 292-297
- [6] 汤丽华,刘敦华.基于近红外光谱的枸杞化学成分定量分析[J].现代食品科技,2013,29(9):2306-2310  
TANG Li-hua, LIU Dun-hua. Quantitative analysis of *lyciumbarbarum L.* composition based on near infrared spectrum [J]. Modern Food Science & Technology, 2013, 29(9): 2306-2310
- [7] 许洪勇,成莲,王东峰,等.傅立叶变换红外光谱法鉴别地沟油的研究[J].现代食品科技,2012,28(6):707-708,719  
XU Hong-yong, CHENG Lian, WANG Dong-feng, et al. Identification of recycled cooking oil by FT-IR [J]. Modern Food Science & Technology, 2012, 28(6): 707-708, 719
- [8] 陈卫江,林向阳,阮榕生等.核磁共振技术无损快速评价食品水分的研究[J].食品研究与开发,2006,27(4):125-127  
CHENG Wei-jiang, LIN Xiang-yang, RUAN Rong-sheng, et al. Study on quickly and non-destructive estimate the moisture content of food using NMR [J]. Food Research and Development, 2006, 27(4): 125-127
- [9] 李银,李侠,张春晖,等.利用低场核磁共振技术测定肌原纤维蛋白凝胶的保水性及其水分含量[J].现代食品科技,2013,29(11):2777-2781  
LI Yin, LI Xia, ZHANG Chun-hui, et al. Water-holding capacity and water content of myofibrillar protein gel by low-field nuclear magnetic resonance [J]. Modern Food Science & Technology, 2013, 29(11): 2777-2781
- [10] LIU F, YUSUF B L, ZHONG J L, et al. Variety identification of rice vinegars using visible and near infrared spectroscopy and multivariate calibrations [J]. International Journal of Food Properties, 2011, 14(6): 1264-1276
- [11] BEVILACQUA M, BUCCI R, MAGRI A D, et al. Tracing the origin of extra virgin olive oils by infrared spectroscopy and chemometrics: a case study [J]. Analytica Chimica Acta, 2012, 717: 39-51
- [12] JIANG L L, LIU F, HE Y, et al. A non-destructive distinctive method for discrimination of automobile lubricant variety by visible and short-wave infrared spectroscopy [J]. Sensors (Basel), 2012, 12(3): 3498-2511
- [13] LIU F, HE Y. Discrimination of producing areas of *auricularia auricula* using visible/near infrared spectroscopy [J]. Food and Bioprocess Technology, 2009, 4(3): 387-394