

# 基于特征波段-Fisher-K 近邻的木器漆拉曼光谱的快速无损鉴别

何亚, 王继芬\*

中国人民公安大学刑事科学技术学院, 北京 100038

**摘要** 为了实现木器漆的快速无损检测以及精确识别与分类, 采集并获取了晨阳等 3 种品牌木器漆样本的拉曼光谱, 并考察了基线校正、Savitzky-Golay 九点平滑法、一阶导数和二阶导数等不同预处理方法的处理效果, 建立了特征波段比值、Fisher 判别、K 近邻(KNN)模型。结果表明, 特征波段比值法能以  $1358\text{ cm}^{-1}/1239\text{ cm}^{-1}$  表征 3 种木器漆的特征; 基于 Fisher 判别的基线校正、平滑和二阶导数处理的拉曼光谱模型的分类准确率最高, 能够实现 100% 区分; 在相同的预处理下, KNN 判别模型的准确率仅为 88.5%。基于二阶导数的拉曼光谱结合特征波段-Fisher-KNN 法能为不同品牌木器漆的准确检测提供一种新的快速无损分析手段, 具有普适性和一定的借鉴意义。

**关键词** 光谱学; 拉曼光谱; 木器漆; Fisher 判别; K 近邻

中图分类号 O433.4

文献标志码 A

doi: 10.3788/LOP57.013001

## Rapid Nondestructive Identification of Wood Lacquer Using Raman Spectroscopy Based on Characteristic-Band-Fisher-K Nearest Neighbor

He Ya, Wang Jifen\*

*Institute of Forensic Science and Technology, People's Public Security University of China, Beijing, 100038, China*

**Abstract** In this study, the Raman spectra of three brands of wood lacquer samples, such as Chenyang brand, are examined to realize the rapid nondestructive detection as well as the accurate identification and classification of wood lacquer. Subsequently, the processing effects of different preprocessing methods, such as the baseline correction, Savitzky-Golay nine-point smoothing, first derivative, and second derivative, are investigated. Three models of characteristic-band ratio, Fisher discriminant, and K nearest neighbor (KNN) are also established. The experimental results indicate that the characteristic-band ratio method is capable of characterizing the features of the three wood lacquer samples with  $1358\text{ cm}^{-1}/1239\text{ cm}^{-1}$ ; further, the Raman spectral model combined with Fisher discriminant based preprocessing methods of baseline correction, smoothing, and second derivative exhibits an optimal classification accuracy of 100%. However, the accuracy of the KNN discriminant model is limited to 88.5% under the same preprocessing. Therefore, the second-derivative based Raman spectroscopy combined with the characteristic-band-Fisher-KNN method can provide a new rapid nondestructive analysis method for the accurate detection of different brands of wood lacquer, exhibiting universality and certain reference significance.

**Key words** spectroscopy; Raman spectroscopy; wood lacquer; Fisher discriminant; K nearest neighbor

**OCIS codes** 300.6450; 120.4630

## 1 引言

木器漆作为微量物证常出现于室内现场, 如作案工具上残留的木器漆附着物、碎尸包裹物上粘附

的木器漆碎片等, 通过对现场采集的木器漆物证进行同一性鉴定, 能够认定或者排除作案工具, 为侦查提供线索, 并为揭露和证实犯罪提供科学依据。木器漆分为水性木器漆和油性木器漆两种, 油性木器

收稿日期: 2019-05-05; 修回日期: 2019-06-14; 录用日期: 2019-07-02

基金项目: 中国人民公安大学 2019 年度基本科研业务费项目库重点项目(2019JKF223)

\* E-mail: wangjifen58@126.com

漆因其溶剂中含有大量挥发性有害物质(VOCs)而逐渐被弃用<sup>[1]</sup>。目前,市售木器漆多为水性木器漆,水性漆是以水作为稀释剂,以丙烯酸树脂乳液(有纯丙、苯丙等各种类型)为主要成分,以成膜助剂、消泡剂和润湿剂为辅助成分<sup>[2]</sup>。目前,针对木器漆的研究多针对在其制作工艺和性能改良方面,而在微量物证检验领域,相关的研究报道很少,如何对其开展快速、准确、无损检验,已成为物证鉴定工作者关注的重点之一。

拉曼光谱分析技术属于非弹性光散射技术,它可以实现快速、高效的样品测量,具有操作简便、测量时间短、无损分析等特点,已被广泛应用于食品安全、化工生物、医学、珠宝鉴定、环境保护等领域<sup>[3-5]</sup>。与红外光谱法相比,拉曼光谱法在油漆鉴定方面更具优势,在鉴别车辆涂漆、墙漆等方面的应用已成熟<sup>[6-7]</sup>,但在木器漆领域尚无报道。

本文以拉曼图谱数据为基础,借助数学算法,建立了基于特征波段比值法、Fisher 判别、K 近邻(KNN)的原始图谱、基线校正(BL)和平滑图谱、一阶导数图谱和二阶导数图谱的分类模型,实现了对木器漆的准确分类,期望能快速准确地对木器漆进行鉴别。

## 2 实验

### 2.1 实验样本

采集了 3 个不同品牌、不同生产厂家的木器漆作为实验样本,其中晨阳(CY)样本 54 个,华彩士(HC)样本 60 个,雀尚(QS)样本 50 个,总计 164 个。其中 6 个样本的部分信息如表 1 所示,采样时采取“井”字格取样法。

表 1 6 个样本的基本信息

Number	Brand	Category	Manufacturer
CY003	CYSO	Water-based multifunctional paint	CYSO
CY008	CYSO	Water-based multifunctional paint	CYSO
HC004	HUACAI	Water-based paint	Chaomeiyaqi
HC013	HUACAI	Water-based paint	Chaomeiyaqi
QS001	TRYKON	Waterborne furniture finish	Changfeng
QS010	TRYKON	Waterborne furniture finish	Changfeng

### 2.2 仪器参数与设置

实验仪器为美国 Thermo Fisher Scientific 公司的 Nicolet Omega XR 显微激光拉曼光谱仪。将涂有木器漆样本的载玻片置于载物台中央,调整位置使激光束通过待测样品中心,于 50 倍镜头下进行拉曼光谱的测定。光谱测定时的仪器参数如下:测试条件为室温,相对湿度为 34%,拉曼光谱波长设定为 780nm,采用 1800 line/mm 光栅分光,波长能量为 2.0 mW,扫描时间为 8 s,曝光次数 4,电荷耦合器件(CCD)探测器接收拉曼信号,针孔直径为 300  $\mu\text{m}$ ,测量范围为 300~3000  $\text{cm}^{-1}$ 。在显微镜下累计采集 3 次,通过比较分析排除来自其他杂质的信号。

### 2.3 光谱预处理

拉曼光谱仪的信号会受到如下因素的干扰: CCD 的发射噪声、暗电流噪声及读出噪声,激发激光的波动,样品的荧光背景、样品及其周围环境的黑体辐射<sup>[8]</sup>等。这些干扰会导致基线漂移,给后续分析造成了影响,进而影响到校正模型的建立和待测样品的预测效果。因此,对光谱采用 BL、Savitzky-Golay 九点平滑法(Smoothing)以及导数法(一阶导数、二阶导数)进行预处理,以达到消除噪声和其他背景干扰的目的。

### 2.4 判别分析模型

模式识别中的判别分析方法可以通过光谱数据对不同的样本按照某些共同的特征进行分类识别,从而发现被测样本之间的内在联系,获得决策性信息<sup>[9]</sup>。因此,判别方法是将光谱数据转化为解决实际问题所需信息的一种重要方法<sup>[10]</sup>。本文选用 Fisher 和 KNN 模型。

#### 2.4.1 Fisher 判别分析

Fisher 线性判别也称线性判别式分析,是基于样本类别进行整体特征提取的一种有效方法。Fisher 判别分析的思想是:将多维数据投影到某个方向上,投影的原则是将总体与总体之间尽可能地放开,然后再选择合适的判别规则,将新的样品进行分类判别<sup>[11]</sup>。

设共有  $c$  个类别,  $\mathbf{x}_j^{(i)}$  为第  $i$  类中的第  $j$  个样本特征向量,样本的类内离散度矩阵  $\mathbf{S}_w$  可以表示为

$$\mathbf{S}_w = \frac{1}{N} \sum_{i=1}^c \sum_{j=1}^{N_i} [\mathbf{x}_j^{(i)} - \mathbf{x}^{(i)}] [\mathbf{x}_j^{(i)} - \mathbf{x}^{(i)}]^T, \quad (1)$$

式中:  $\mathbf{x}^{(i)}$  为第  $i$  类样本特征向量的均值;  $N_i$  为第  $i$  类样本的样本数;  $N$  为样本总数,即  $N = \sum_{i=1}^c N_i$ 。

该样本的离散程度矩阵  $\mathbf{S}_b$  为

$$S_b = \frac{1}{N} \sum_{i=1}^c N_i [x^{(i)} - \bar{x}] [x^{(i)} - \bar{x}]^T, \quad (2)$$

式中:  $\bar{x}$  为所有样本特征向量均值。

Fisher 线性判别的准则函数定义为

$$J(W) = \frac{|W^T S_b W|}{|W^T S_w W|}. \quad (3)$$

式中:  $W$  为所求投影变换后的解向量。使函数  $J(W)$  达到最大值的  $W$  即为最佳向量  $W^*$ , 由此可以确定线性判别函数。

### 2.4.2 K 近邻算法

K 近邻(KNN)算法采用测量不同特征值之间距离的方法进行分类, 即假设已知样本集中各特征值及其类别, 在输入没有类别标签的新数据后, 将新数据的每个特征与样本集中数据对应的特征进行比较, 计算两数据点之间的欧几里得距离, 选择前  $K$  个距离最小(特征最相似)的数据所对应的分类标签, 即出现次数最多的分类, 作为新数据的分类<sup>[12]</sup>。具体描述如下: 以训练集  $X$  为样本的特征向量集,  $Y$  为样本的类别,  $T$  为一个特征空间, 根据给定的距离度量, 在训练集  $X$  的特征空间  $T$  中找出与  $x$  最邻近的  $K$  个点, 将包含这  $K$  个点的  $x$  的邻域记作  $N_K(x)$ , 在  $N_K(x)$  中根据分类决策规则(如多数表决)决定  $x$  的类别  $y$ , 其表达式为

$$y = \arg \max_{x_i \in N_K(x)} I(y_i = c_j),$$

$$i = 1, 2, 3, \dots, N, \quad j = 1, 2, 3, \dots, K, \quad (4)$$

式中:  $x_i$  为实例的特征向量;  $y_i \in Y \{c_1, c_2, \dots, c_k\}$

为实例的类别;  $I(\cdot)$  为指示函数, 即当  $y_i = c_j$  时,  $I$  为 1, 否则  $I$  为 0。

### 2.5 特征波段比值模型

特征波段比值法就是在已预处理过的光谱中提取反映各样本特征信息的波段, 通过其相互比值来反映各样本的差异, 并以特征波段的比值为类间距离作图分类<sup>[13]</sup>, 具有简单、易操作等优点。

## 3 结果与讨论

### 3.1 数据预处理

虽然拉曼光谱的扫描范围比较广, 包含的信息量多, 但光谱中存在较严重的噪声, 因此, 全谱范围建模方法在避免信息丢失的同时, 也将不可避免地受到其他因素和噪声的干扰<sup>[14]</sup>。选择合适的光谱区域进行后续分析, 可以提高模型的效果和稳定性。对木器漆拉曼光谱进行综合分析后可知, 木器漆的信息主要集中在  $500 \sim 1700 \text{ cm}^{-1}$ , 因此本研究选择  $500 \sim 1700 \text{ cm}^{-1}$  波段谱图进行木器漆信息的考察。

为了消除噪声和基线干扰因素, 对样品数据分别进行 BL、九点平滑法、一阶导数以及二阶倒数处理后再进行建模分析, 考察预处理方法对模型预测准确率的影响。CY003 样品的拉曼谱图如图 1 所示, 其中, 图 1(a) 为自动 BL 后的图谱, 图 1(b) 为九点平滑图谱, 图 1(c) 为一阶导数图谱, 图 1(d) 为二阶导数图谱。

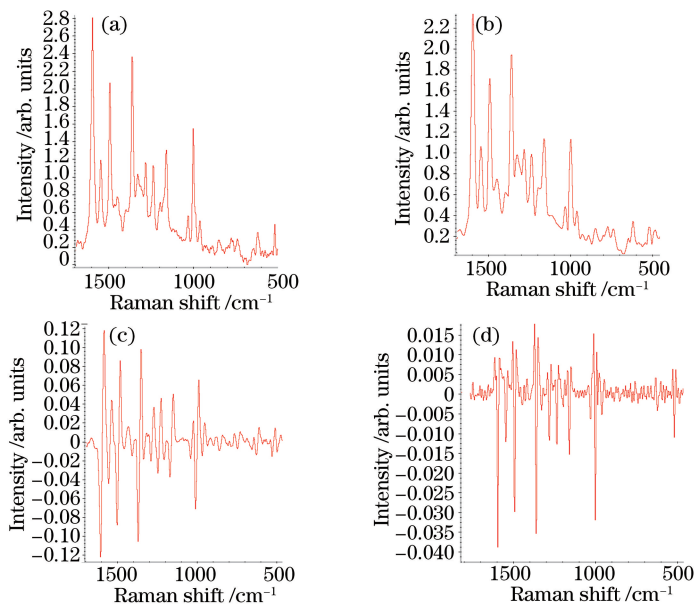


图 1 CY003 样品拉曼光谱图的预处理结果。(a) BL; (b) 九点平滑法; (c) 一阶导数; (d) 二阶导数

Fig. 1 Preprocessing results of Raman spectra of CY003 samples. (a) BL; (b) nine-point smoothing; (c) first derivative; (d) second derivative

### 3.2 特征值分析

由图 1(a)可知:该样品在  $997, 1160, 1239 \text{ cm}^{-1}$  等 7 个峰处出现了较明显的特征吸收,其中  $1160, 1283, 1358 \text{ cm}^{-1}$  这 3 个特征峰均与聚酯  $\text{C}=\text{O}$  双键相关。取 BL 图谱中特征拉曼位移频率为  $1160, 1283, 1358 \text{ cm}^{-1}$  的 3 个特征峰值的强度作为鉴别不同种类木器漆的依据,以  $1239 \text{ cm}^{-1}$  处的特征峰值强度为基准,分别用另外 3 组特征峰值强度与其相除得到 3 列新数据:  $1160 \text{ cm}^{-1}/1239 \text{ cm}^{-1}$ 、 $1283 \text{ cm}^{-1}/1239 \text{ cm}^{-1}$ 、 $1358 \text{ cm}^{-1}/1239 \text{ cm}^{-1}$ 。然后,以  $1358 \text{ cm}^{-1}/1239 \text{ cm}^{-1}$  为纵轴,以  $1283 \text{ cm}^{-1}/1239 \text{ cm}^{-1}$  为横轴作二维散点图,得到木器漆品牌的分类图,如图 2 所示。

由图 2 可知:晨阳品牌木器漆样品的分布均高于  $Y=1.44$ ;而华彩士、雀尚两品牌木器漆均分布在  $B < 1.45$  的范围内;晨阳品牌与华彩士、雀尚两品牌区分显著,但华彩士、雀尚两品牌木器漆有样本混杂,区分效果一般。

此外,分别以  $A = 1283 \text{ cm}^{-1}/1239 \text{ cm}^{-1}$ 、 $B = 1358 \text{ cm}^{-1}/1239 \text{ cm}^{-1}$ 、 $C = 1160 \text{ cm}^{-1}/1239 \text{ cm}^{-1}$  作为坐标轴作出三维散点图,但区分度与二维图类似,木器漆样品的聚集分类情况主要取决于  $1358 \text{ cm}^{-1}/1239 \text{ cm}^{-1}$  的大小。木器漆为混合物,

表 2 不同预处理方法的判别结果

Table 2 Discriminant results of different preprocessing methods

Index	Original spectrum	BL+Smoothing	BL+Smoothing+ 1 <sup>st</sup> derivative	BL+Smoothing+ 2 <sup>nd</sup> derivative
Overall accuracy /%	68.0	79.8	96.0	100
Number of errors	55	35	7	0
Cross validation rate /%	65.1	69.4	94.1	100

由表 2 可知:直接对原始谱图进行判别,正确率仅为  $68.0\%$ ,交叉验证率仅为  $65.1\%$ ,这是因为原始光谱受到了噪声的干扰,噪声掩盖了原始信息;与原始光谱相比,经过预处理的光谱判别率均有显著提升,采用 BL 结合九点平滑法处理后,正确率达到  $79.8\%$ ,交叉验证率为  $69.4\%$ ,表明这种预处理方法能较好地降低噪声,提高信噪比。对光谱求导能消除基线漂移和其他背景的干扰,强化特征谱带,克服谱带重叠<sup>[15]</sup>。采用 Fisher 判别模型结合 BL、九点

表 3 Fisher 函数判别摘要表

Table 3 Summary of Fisher-function discriminant

Function	Eigen value	Cumulative variance	Regular correlation	Test function	Wilks' Lambda	Significance
1	113.871	90.1	0.996	1 and 2	0.001	0
2	12.455	100	0.962	1	0.074	0.041

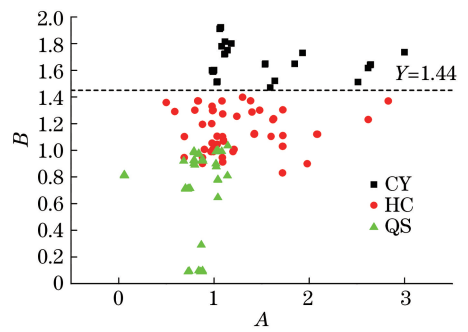


图 2 根据特征波段比值法获得的 3 个木器漆品牌样本分类图

Fig. 2 Classification of three brands of woodlacquer samples obtained by characteristic-band ratio method

其拉曼光谱的主要差异是由丙烯酸树脂聚合物的不同导致的, $B = 1358 \text{ cm}^{-1}/1239 \text{ cm}^{-1}$  表示的是  $\text{C}=\text{O}$  相对不饱和度的大小,因此,利用拉曼光谱提取木器漆的信息来鉴别不同品牌木器漆是可行的。

### 3.3 Fisher 判别分析

采用主成分分析法对拉曼数据降维后,再借助 Fisher 判别构建模型,比较了基础处理(九点平滑法和 BL)、导数处理以及不同处理方法组合对模型判别率的影响,得到了各模型的分类结果,如表 2 所示。

平滑法以及一阶导数方法的预处理可以实现  $96.0\%$  的正确率,交叉验证率也达到了  $94.1\%$ ;而将 BL、平滑以及二阶导数方法相结合的预处理后的正确率能达到  $100\%$ ,误判数为 0。相比于一阶光谱,二阶导数光谱能呈现出许多原始谱图中被掩盖的谱峰的斜率变化特征,将样本谱图间的差异更为明显地表示出来。综上所述,BL、平滑以及二阶导数方法预处理分类效果最为理想,对其进行进一步分析,可以得到判别函数摘要表,如表 3 所示。



相关性表明了不同分组与各个函数之间的相关性。表3中函数1与函数2的相关性均大于0.9,这表明不同的分组在两函数上的相关性极强。两函数的累计方差贡献率达到了100%,很好地解释了样本的基本信息。Wilks' Lambda是组内平方和与总平方和之比。显著性值越小,说明某个量对模型的影响越显著。分别选择函数1和函数2作为判别函数,构建判别分类模型,得到了3个品牌木器漆样本的判别分布图,如图3所示。

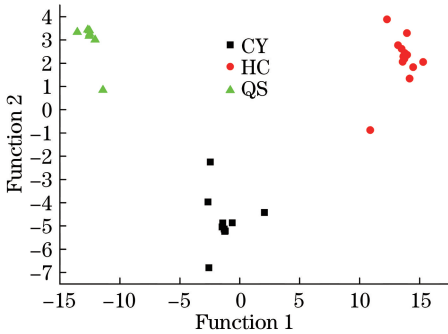


图3 3个品牌木器漆样本在函数1和函数2上建立的判别分布图

Fig. 3 Discriminant distribution of three brands of woodlacquer samples obtained by functions 1 and 2

由图3可知,3种类型的样本均实现了较为明显的区分,其中,晨阳木器漆样本与雀尚木器漆样本在函数1判别轴上区明显,华彩士木器漆样本与雀尚木器漆样本在函数2判别轴上实现了理想区分。

### 3.4 KNN 判别分析

根据欧几里得距离,选取 $K$ 值为1,建立分类模型(训练集与保持集中样本数量之比为3:1),以Fisher判别效果较好的基础预处理结合一阶导数或二阶导数预处理后的光谱数据为基础,建立KNN模型对各样本进行分类,得到了KNN模型下各样本的分类结果与Fisher判别对比表,如表4所示。

表4 两组模型下各样本的分类结果

Table 4 Classification results of samples for two models

Spectral type	Overall accuracy / %		Cross validation rate / %
	Fisher	KNN	
BL+Smoothing+ 1 <sup>st</sup> derivative	96	79.2	100
BL+Smoothing+ 2 <sup>nd</sup> derivative	100	88.5	100

由表4可知:采用基础加二阶导数预处理光谱构建的判别模型的准确率最高,这是因为二阶导数光谱能呈现出许多原始谱图中被掩盖的谱峰的斜率

变化特征<sup>[16]</sup>,将样本谱图间的差异更明显地表示了出来;相比较Fisher模型,KNN模型分类准确率均较低,这可能是由于样本不均,当样本不平衡时大容量样本会影响系统归类,从而导致误判。实验中,华彩士木器漆样本最多,有60个,而雀尚木器漆样本只有50个,这使得其误判概率增大。

综上所述,在木器漆拉曼光谱的分析中,特征波段比值法可以有效地将晨阳与其他品牌木器漆区分开来,但华彩士与雀尚品牌木器漆的区分效果不理想。为了更有效地对木器漆进行分类,需建立判别模型,分别采用BL、九点平滑法以及导数预处理光谱,结果表明,采用BL、平滑加二阶导数的预处理方法,可在最大程度减小噪声及基线信息等的干扰;对该预处理组合同样建立KNN模型,模型精度稍低于Fisher模型,能实现初步区分。可见,基础预处理加二阶导数对拉曼图谱进行预处理,然后结合Fisher判别,可以实现对木器漆油漆的快速、准确鉴别。

## 4 结 论

利用拉曼原始图谱及其BL和九点平滑法、一阶导数、二阶导数光谱预处理与特征波段比值法、Fisher判别和KNN分析对木器漆进行了识别与分类。拉曼特征波段比值法能在一定程度上区分木器漆,但若要实现精准分类就需建立判别模型。实测光谱中存在大量不相关的信息和噪声,采用BL、九点平滑法加二阶导数的预处理方法可以最大限度地弱化噪声区间和干扰区间的权重,构建的模型分类效果更好,结合Fisher判别分析的准确率可达100%,可有效开展对木器漆的区分鉴别。另外,KNN的分类模型具有无法映射出个别样本在某些特征空间中差异性的不足之处,数据样本不均衡会直接影响分类效果。在本实验中,不同种类的样本在数量上未达到均衡,导致KNN分类效果稍逊于Fisher判别。本实验方法可以为其他微量物证的分类鉴定提供一种新思路,借助判别分析法可以避免传统鉴定方法中主观判断的干扰臆断,从而实现准确无误的鉴别。

## 参 考 文 献

- [1] Hu L, Su J Q, Lin X, et al. Synthesis and properties of castor oil-based waterborne polyurethane wood coating cured through UV/air dual process [J]. Surface Technology, 2019, 48(3): 219-227.

- 胡隆, 苏佳琦, 林兴, 等. 蓖麻油基 UV/空气双重固化水性聚氨酯木器漆的制备及性能研究[J]. 表面技术, 2019, 48(3): 219-227.
- [2] Huang L, Liang W H, Qin M H. Effect of temperature on formaldehyde and volatile organic compound emissions from a wood lacquer [J]. Building Science, 2018, 34(8): 71-76, 88.  
黄丽, 梁卫辉, 秦孟昊. 温度对木器漆甲醛与可挥发性有机化合物(VOCs)散发的影响[J]. 建筑科学, 2018, 34(8): 71-76, 88.
- [3] Zhu L L, Feng A M, Jin S Z, et al. Fluorescence suppression methods in Raman spectroscopy detection and their application analysis [J]. Laser & Optoelectronics Progress, 2018, 55(9): 090005.  
朱磊磊, 冯爱明, 金尚忠, 等. 拉曼光谱检测中荧光抑制方法及其应用分析[J]. 激光与光电子学进展, 2018, 55(9): 090005.
- [4] Zheng J W, Yang T W. Classification method of biological tissues based on Raman spectrum features [J]. Laser & Optoelectronics Progress, 2017, 54(5): 053001.  
郑家文, 杨唐文. 基于拉曼光谱特征的生物组织识别方法[J]. 激光与光电子学进展, 2017, 54(5): 053001.
- [5] Gao F, Xu L Z, Zhang Y J, et al. Analytical Raman spectroscopic study for discriminant analysis of different animal-derived feedstuff: understanding the high correlation between Raman spectroscopy and lipid characteristics[J]. Food Chemistry, 2018, 240: 989-996.
- [6] Yu J, Wang L Q, Wang J F, et al. The research and applications of Raman spectroscopy in the identification of paints [J]. Modern Instruments, 2012, 18(3): 21-23.  
余静, 王琳琦, 王继芬, 等. 拉曼光谱技术在涂料鉴定中的应用进展[J]. 现代仪器, 2012, 18(3): 21-23.
- [7] Xu Z S, Chen D J, Lü H R, et al. Raman spectra: recent developments to revolutionize trace evidence [J]. Shandong Chemical Industry, 2016, 45(9): 70-72.  
许中石, 陈东杰, 吕弘儒, 等. 拉曼光谱在微量物证鉴定中的应用[J]. 山东化工, 2016, 45(9): 70-72.
- [8] Saito R, Tatsumi Y, Huang S, et al. Raman spectroscopy of transition metal dichalcogenides [J]. Journal of Physics: Condensed Matter, 2016, 28(35): 353002.
- [9] Li H D, Xu Q S, Liang Y Z. LibPLS: an integrated library for partial least squares regression and linear discriminant analysis [J]. Chemometrics and Intelligent Laboratory Systems, 2018, 176: 34-43.
- [10] Li X, Huang M Z, Song B, et al. Nondestructive discrimination of Shoushan Stone and Laos Stone based on Raman spectroscopy and multivariate methods [J]. Acta Optica Sinica, 2016, 36(12): 1230001.  
李夏, 黄梅珍, 宋彪, 等. 拉曼光谱结合多变量统计方法在无损区分寿山石和老挝石中的应用[J]. 光学学报, 2016, 36(12): 1230001.
- [11] Li F, Wang J X, Chyu M K, et al. Weak fault diagnosis of rotating machinery based on feature reduction with supervised orthogonal local fisher discriminant analysis [J]. Neurocomputing, 2015, 168: 505-519.
- [12] Deng Z Y, Zhu X S, Cheng D B, et al. Efficient  $k$ NN classification algorithm for big data [J]. Neurocomputing, 2016, 195: 143-148.
- [13] Li Y P. Identification of edible oil and detection of its adulteration via Raman spectroscopy [D]. Guangzhou: Jinan University, 2016.  
黎远鹏. 基于拉曼光谱法的食用油定性鉴别与掺伪含量检测研究[D]. 广州: 暨南大学, 2016.
- [14] Li Z G, Peng S L, Yang N, et al. Quantitative analysis method of infrared spectra based on derivative spectra fusion modeling [J]. Chinese Journal of Analytical Chemistry, 2016, 44(3): 437-443.  
李志刚, 彭思龙, 杨妮, 等. 基于导数光谱融合建模的红外光谱定量分析方法[J]. 分析化学, 2016, 44(3): 437-443.
- [15] Liu X B, Zhang Z M, Liang Y Z, et al. Baseline correction of high resolution spectral profile data based on exponential smoothing [J]. Chemometrics and Intelligent Laboratory Systems, 2014, 139: 97-108.
- [16] Wang J M, Zhang J C, Zhang Z J. Rapid determination of nitrate nitrogen and nitrite nitrogen by second derivative spectrophotometry [J]. Spectroscopy and Spectral Analysis, 2019, 39(1): 161-165.  
王静敏, 张景超, 张尊举. 二阶导数光谱法快速测定硝酸盐氮和亚硝酸盐氮[J]. 光谱学与光谱分析, 2019, 39(1): 161-165.