

基于支持向量机建模的重质矿物油光谱模式识别

侯伟, 王继芬*, 何欣龙

中国人民公安大学侦查学院, 北京 100038

摘要 重质矿物油的检验分析在交通事故案件处理过程中具有重要作用。为了实现对重质矿物油种类的准确区分, 本文采集了汽机油、柴机油、润滑油、齿轮油和液压油 5 种重质矿物油共计 120 份样本的红外光谱和拉曼光谱数据, 结合光谱融合的相关方法, 建立了基于支持向量机(SVM)的重质矿物油分类判别模型。实验结果表明: 使用单一光谱数据进行建模分类的准确率较低; 对初级光谱融合数据进行建模分析时, 5 种重质矿物油的分类识别准确率稍高于前者, 最高可达 75%; 使用中级光谱融合数据结合主成分分析建模能够实现 5 种重质矿物油的完全区分, 在 26 维矩阵上特征提取得最好, 分类识别率为 100%。使用光谱数据融合结合 SVM 建模分析, 能够实现重质矿物油的完全区分, 该方法提高了检验鉴定效率, 能够满足公安机关提出的快速、准确的检验要求, 为基层民警处理相关案件提供了理论支撑和方法参考。

关键词 光谱学; 重质矿物油; 光谱数据融合; 多元分析

中图分类号 O433.4

文献标志码 A

doi: 10.3788/LOP202158.0630001

Spectral Pattern Recognition of Heavy Mineral Oil Using Support Vector Machine Modeling

Hou Wei, Wang Jifen*, He Xinlong

School of Investigation, People's Public Security University of China, Beijing 100038, China

Abstract The inspection and analysis of heavy mineral oil plays an important role in dealing with traffic-accident cases. In order to obtain accurate classifications of heavy mineral oils, we collected infrared and Raman spectral data for 120 samples of five kinds of heavy mineral oils, including gasoline engine oil, diesel engine oil, grease, gear oil, and hydraulic oil. We established a classification and discrimination model for heavy mineral oil by using a support vector machine (SVM) combined with a spectral-fusion method. The results showed the accuracy of modeling classification using single-spectrum data to be rather low. When we modeled and analyzed the data obtained from primary spectral fusion, the classification and recognition rates for the five heavy mineral oils were slightly better, with an accuracy up to 75%. However, modeling that used data from intermediate spectral fusion combined with principal component analysis achieved complete differentiation among the five heavy mineral oils, with feature extraction from the 26-dimensional matrix being the best, with an accuracy up to 100%. In summary, spectral-fusion data combined with SVM modeling analysis can achieve complete separation among heavy mineral oils. The method improves the efficiency of inspection and identification, which fulfills the goal of rapid and accurate inspection for frontline law-enforcement personnel. It also provides theoretical support and a reference method for relevant cases.

Key words spectroscopy; heavy mineral oil; spectral data fusion; multianalysis

OCIS codes 300.6340; 000.3870

收稿日期: 2020-07-14; 修回日期: 2020-08-05; 录用日期: 2020-08-18

基金项目: 中央高校基本科研业务费专项资金(2019JKF223)

* E-mail: wangjifen58@126.com

1 引言

重质矿物油是交通肇事案件现场遗留的一种常见微量物证,对其进行检验分析可在一定程度上为嫌疑车辆的认定提供侦察方向和摸排线索,在案件处理过程中往往可以起到关键作用。然而,目前针对重质矿物油的研究主要集中在环境和食品安全等领域^[1-2],在法庭科学领域的相关成果较少。张振宇等^[3]采用固相微萃取-气质联用(SPME-GC/MS)方法对火灾现场润滑油的燃烧残留物进行了检验,结果表明,该方法能够对润滑油的种类和成分进行认定;张永国等^[4]采用气相色谱-质谱联用(GC/MS)方法对 4 种航空发动机润滑油进行了检验分析,分析结果为控制润滑油的高温氧化提供了科学依据;Eschner 等^[5]采用 GC/MS 方法对矿物油样本进行了分析,实现了全面的 2D 表征。虽然采用 GC/MS 技术检验矿物油的方法具有灵敏度高、选择性好的优点,能够实现样本之间细微差别的准确区分,但由于公安工作具有一定的特殊性,此类方法有时难以应用于公安实战。首先,复杂的前处理过程和较长的分析时间无法满足快速办案的要求;其次,大型仪器设备无法应用于现场检材的快速筛查,在一定程度上影响了办案的进度和效率;最后,上述方法会对样本造成不可逆的损坏,不利于证据的保存和后续检验,破坏了证据链的完整性。

光谱技术在法庭科学领域有着其他技术所不具备的特殊优势,其以无损检验的特性被广泛用来对案件中的检材进行分析。光谱数据融合技术可以整合多类光谱的有效信息,克服单一光谱的信息缺陷,实现不同光谱数据间的优势互补,从而提高检验效率^[6]。光谱数据融合技术可分为初级、中级和高级三个层次,其中:初级光谱数据融合是指将多种光谱的原始数据进行简单叠加,形成更加丰富的数据集;中级光谱融合是指将多种光谱数据的特征变量进行叠加和整合,挖掘关键信息,以精简数据集;高级光谱融合是指整合多种数据模型,以获取最佳结果。一般来说,通过初级和中级光谱数据融合得到的分析结果基本能够满足研究的需要,因此高级光谱数据融合技术在目前的相关研究中较少得到使用^[7-8]。

本文首先采集并分析了 5 种重质矿物油共计 120 份样本的红外光谱和拉曼光谱数据,然后借助光谱融合的相关方法,建立了基于支持向量机

(SVM)的重质矿物油分类判别模型,该模型可实现 5 种重质矿物油的完全区分。与 GC/MS 等方法相比,本文方法具有成本低、污染小、迅速准确、不破坏样本、分类结果可视化等优点,能够满足公安机关办理案件的实际需要。

2 实验

2.1 样本与仪器

样本:不同品牌不同种类的重质矿物油样品共 120 份,其中包含 60 份汽机油样品、33 份柴油油样品、12 份润滑脂样品、9 份齿轮油样品和 6 份液压油样品。

仪器:Nicolet 5700 型衰减全反射-傅里叶变换红外光谱仪和傅里叶变换拉曼光谱仪。

采用商业数学软件 MATLAB R2014a 进行数据的预处理、降维以及建模分析。

2.2 实验方法

分别采集 120 份重质矿物油样本的傅里叶变换红外光谱和拉曼光谱,两种光谱的采集范围分别为 $4000\sim 400\text{ cm}^{-1}$ 和 $3700\sim 100\text{ cm}^{-1}$ 。扫描次数为 64 次,分辨率为 4 cm^{-1} 。每个样本连续进行 3 次光谱曲线的采集,采用 3 次光谱数据的平均值作为后续分析的实验样本^[9-11]。

2.3 数据预处理方法

对采集后的光谱进行自动基线校正(automatic baseline correction)^[12]、纵坐标归一化(ordinate normalization)^[13]、标准正态变换(standard normal variate, SNV)^[14]、多元散射校正(multiplicative scatter correction)^[15]和 S-G 平滑滤波^[16-17]等预处理,将光谱数据转化为数值数据,然后采用 Z-score 方法^[18]进行标准化处理。

2.4 多元建模分析

SVM 是一种二值分类数学模型,该模型被定义为特征空间上的间隔最大的线性分类器。SVM 就是在多维数据空间中构建一个超平面对两类数据进行区分,最终将两类数据的区分转化成为一个凸二次规划问题的求解^[19-21]。

SVM 算法最初是为二值分类问题设计的,无法直接用于区分多类数据。当进行多分类任务时,需要构建合适的多类分类器。目前解决此类问题的办法就是对数据进行拆分,将多分类任务拆分成单独的二分类任务,对每个二分类器的预测结果进行集成,以获得最终的多分类结果。常见的拆分策略有一对一法(OVO SVMs)、一对多法(OVR SVMs)和

多对多法 (MVM SVMs)。下面以一对多法为例进行介绍。

对于 n 类数据中的某一类, 将其视作 +1 类, 而其余 $n-1$ 类的所有样本均视作 -1 类。构建一个 binary SVM 模型, 用该模型将第 i 类和其余 $n-1$ 类分开, 分开的过程就是求解二次规划问题。超平面可表示为如下集合 $\{x | w^T x + b = 0\}$, 该问题即可转化为求解最优的权重向量 w 和偏置量 b , 使该平面离样本点的距离最大, 数学表达式为

$$\begin{aligned} \min_{w^i, b^i, \zeta^i} & \frac{1}{2} (w^i)^T w^i + C \sum_{t=1}^N \zeta^i, \\ \text{s. t. } & (w^i)^T \phi(x_t) + b^i \geq 1 - \zeta^i, \text{ if } y_t = i, \\ & (w^i)^T \phi(x_t) + b^i \leq -1 + \zeta^i, \text{ if } y_t \neq i, \\ & \zeta^i \geq 0, \end{aligned} \quad (1)$$

式中: ζ^i 为引入的松弛变量; ζ_t^i 为训练样本的松弛变量; C 为惩罚系数; N 为变量的总数; ϕ 为输入空间到特征空间的非线性映射; x_t 为训练样本的空间坐标; $y_t \in \{+1, -1\}$; 下标 t 表示样本的索引; 上标 $i \in \{1, \dots, n\}$ 。共需 n 个 binary SVM 来求解 n 个包含有 N 个变量的二次规划问题, 其中第 i 类的决策函数为

$$\begin{aligned} d_{\text{new}}^i &= (w^i)^T \phi(x_{\text{new}}) + b^i = \\ & \left(\sum_{t=1}^N y_t \alpha_t^i x_t \right)^T x_{\text{new}} + b^i, \end{aligned} \quad (2)$$

式中: α_t^i 为拉格朗日乘子; x_{new} 为测试集中未知样本的空间坐标。

在测试阶段, 对于 x_{new} , n 个决策函数共有 n 个输出, 选择使 (2) 式最大的类 i 作为 x_{new} 的预测, 亦即采用 (3) 式所示的决策函数对其进行分类。

$$\begin{aligned} f(x) &= \arg \max_{i \in \{1, \dots, n\}} [(w^i)^T \phi(x_{\text{new}}) + b^i] = \\ & \arg \max_{i \in \{1, \dots, n\}} \left[\left(\sum_{t=1}^N y_t \alpha_t^i x_t \right)^T x_{\text{new}} + b^i \right], \end{aligned} \quad (3)$$

若 x_{new} 位于 i 类区域, 则 d_{new}^i 最大, x_{new} 就被预测为 i 类。

本实验采用 SVM 进行数据建模。在预实验过程中比较了 4 种核函数 (线性核函数、多项式核函数、Sigmoid 核函数以及径向基核函数) 下重质矿物油样本的分类识别准确率, 结果发现选择径向基函数 (RBF) 作为核函数时的效果较好, 分类识别准确率较高, 故本文选择径向基函数作为 SVM 模型的核函数, 其余参数采用 MATLAB 系统的默认值 (规范化参数为 10, 回归精确度为 0.1, 惩罚系数 C 和 RBF 伽马值均为 0.1)。径向基核函数的表达式为

$$\varphi(\|X - X_c\|) = \exp\left(\frac{-\|X - X_c\|^2}{2\sigma^2}\right), \quad (4)$$

式中: X_c 为核函数中心; σ 为函数的宽度参数, 控制着函数的径向作用范围。若 X 和 X_c 很相近, 则核函数值为 1; 若 X 和 X_c 相差较大, 则核函数值约为 0。该函数也被称为高斯函数。

3 结果及分析

3.1 各组分样本的红外光谱和拉曼光谱分析

借助衰减全反射-傅里叶变换红外光谱仪和拉曼光谱仪分别采集包括汽机油、柴机油、润滑脂、齿轮油和液压油在内的 120 份重质矿物油样本的红外光谱和拉曼光谱。不同种类重质矿物油部分样本的红外光谱和拉曼光谱分别如图 1 和图 2 所示。

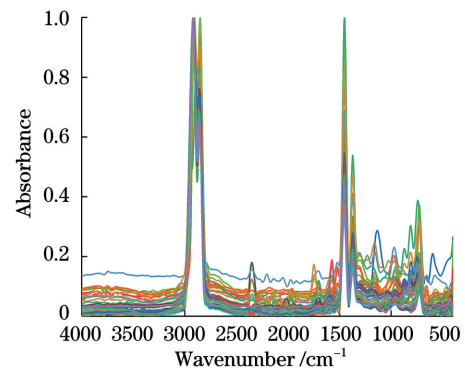


图 1 部分样本的红外光谱

Fig. 1 Infrared spectra of some samples

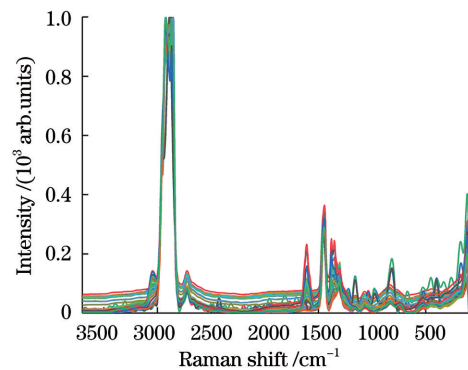


图 2 部分样本的拉曼光谱

Fig. 2 Raman spectra of some samples

由图 1 可知: 各样本红外光谱中的峰形、峰的走向及出峰位置基本一致, 在波数为 2900 cm^{-1} 处均有一双峰, 在波数为 1450 cm^{-1} 和 1380 cm^{-1} 处均有一峰; 个别峰的个数和相对峰高有所区别, 如部分样本在波数为 1750 cm^{-1} 和 1590 cm^{-1} 处有一窄峰, 部分样本则没有; 在波数为 $400 \sim 1300 \text{ cm}^{-1}$ 的光谱“指纹区”, 各样本的峰形变化较大, 有助于区分不同重质矿物油样本结构上的微小差异。

由图 2 可以看出,各样本在波数为 2900 cm^{-1} 处均有一宽峰,在波数为 1380 cm^{-1} 处有一左高右低的双峰,在波数为 1450 cm^{-1} 处均有一明显的窄峰,在波数为 2600 cm^{-1} 处的相对峰高有所区别。通过观察谱图结构可知,各样本的谱图结构具有高度相似性,仅仅依靠谱图特征对重质矿物油种类进行区分的难度较大。因此,本研究借助化学计量学的相关知识对 120 份重质矿物油样本开展分类工

作,以达到快速、准确区分重质矿物油的目的。

3.2 基于红外光谱及拉曼光谱的 SVM 模型分类识别结果

采用 SVM 建模方法分别对全部重质矿物油样本的红外光谱数据和拉曼光谱数据建立分类模型,所用时间分别为 1 min 40 s 和 1 min 35 s。两类光谱数据模型下 5 种重质矿物油的分类识别准确率如图 3 所示。

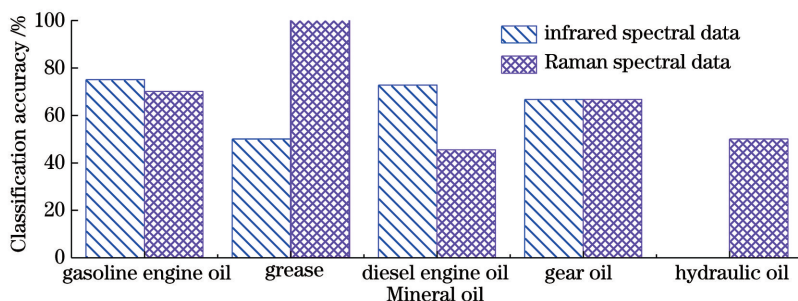


图 3 5 种重质矿物油在两类光谱数据模型下的分类识别准确率

Fig. 3 Classification accuracy of five kinds of heavy mineral oils under two types of spectral data models

由图 3 相关数据可知:利用红外光谱数据和拉曼光谱数据进行 SVM 建模,两者的总体分类识别准确率较低且数值相近,分别为 67.50% 和 65.50%;红外光谱数据模型对汽机油和柴机油的分类识别准确率高,但对润滑脂和液压油的分类识别准确率相对较差,两种模型对齿轮油的分类识别准确率相同(均为 66.67%)。由此可见,两种不同类型的光谱数据中包含的可用于分类的信息不均衡,仅利用单一光谱数据进行建模的分类识别准确率较低,无法完全将 5 种重质矿物油区分开来,需要对数据进行进一步处理,以便获取更加全面、准确的信息。

3.3 基于初级光谱融合数据的模式识别分类结果

红外光谱的“指纹区”通常指波数在 $1300\sim 400\text{ cm}^{-1}$ 范围内的一段区域,该区域光谱的峰形变化较为明显,具有较强的特征性,通常包含能够反映

不同样本之间细微差别的重要信息。拉曼光谱的部分区域不产生吸收峰,过多的无关变量会造成数据的冗余,将此类噪声区域剔除后便可得到拉曼光谱出峰区域。在重质矿物油拉曼光谱中,“指纹区”通常是波数在 $3100\sim 2400\text{ cm}^{-1}$ 和 $1700\sim 100\text{ cm}^{-1}$ 这两部分区域。分别对 120 份重质矿物油样本的红外光谱指纹区和红外光谱全波段、红外光谱指纹区和拉曼光谱全波段、拉曼光谱出峰区和红外光谱全波段、拉曼光谱出峰区和拉曼光谱全波段、红外光谱指纹区和拉曼光谱出峰区、红外光谱全波段和拉曼光谱全波段数据进行初级融合,将各部分数据变量进行叠加,便可得到 6 种初级光谱融合数据集。使用 SVM 建模方法分别对以上 6 种初级融合光谱数据构建分类判别模型,模型分别用 Q1、Q2、Q3、Q4、Q5 和 Q6 进行表示。6 种模型下各类别和总体的分类识别准确率分别见表 1 和图 4。

表 1 5 种重质矿物油在 6 种初级光谱融合数据模型下的分类识别准确率

Table 1 Classification accuracy of five kinds of heavy mineral oils under the six types of data models of primary spectral fusion

Heavy mineral oil	Classification accuracy / %					
	Q1	Q2	Q3	Q4	Q5	Q6
Gasoline engine oil	75.00	85.00	80.00	80.00	75.00	75.00
Diesel engine oil	81.82	18.18	63.64	45.00	36.36	81.82
Grease	25.00	75.00	75.00	100	100	100
Gear oil	66.67	0	33.33	0	66.67	66.67
Hydraulic oil	0	50.00	0	50.00	0	0

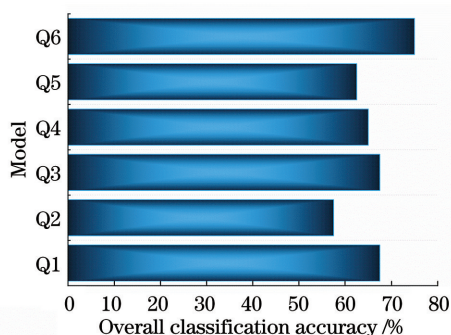


图 4 5 种重质矿物油在 6 种初级光谱融合数据模型下的总体分类识别准确率

Fig. 4 Overall classification accuracy of five kinds of heavy mineral oils under the six types of data models of primary spectral fusion

由图 4 可知,基于红外-拉曼初级融合光谱构建的 SVM 分类模型的总体分类识别准确率(最高可达 75%)优于其他 5 种初级融合光谱数据模型,这可能是由于全波段光谱能够包含完整的样本信息,有利于 SVM 模型基于全部数据对 5 种重质矿物油进行识别和分类。

由图 3 和图 4 中的相关数据可知,对红外-拉曼全波段光谱进行初级融合后的数据进行建模分析时,5 种重质矿物油的总体分类识别准确率较单一红外光谱和拉曼光谱分别提高了 7.5 个百分点和 9.5 个百分点。对比图 3 和表 1 可以看出:汽机油和润滑脂初级光谱融合数据模型的分类识别准确率分别与红外光谱、拉曼光谱数据模型的结果相同;齿轮油的分类识别准确率未发生明显变化,仍为 66.67%;柴油油的分类识别率高于两种单一光谱数据模型的结果,达到了 81.82%,能够体现光谱融合的优势;而液压油的分类识别准确率低于拉曼光谱数据模型的结果,这可能是由于光谱融合后维度过高,冗杂的无关变量阻碍了模型对关键信息的挖掘,进而导致模型精度下降,这也是造成初级融合光谱数据模型无法完全区分 5 种重质矿物油的原因所在,因此必须合理控制数据维度,注重有效信息的采集。

3.4 基于中级光谱融合数据的模式识别分类结果

采用主成分分析(PCA)法分别对 935 维的红外光谱和拉曼光谱数据进行降维处理,提取特征向量,用时分别为 35 s 和 39 s。两种光谱的前 39 个成分均累计包含各自全部变量 100% 的信息,因此将红外光谱和拉曼光谱数据经降维处理后的前 39 个特征变量进行叠加,便可得到基于特征变量的中级融合光谱数据;对该数据再次进行主成分分析(PCA)

处理,对特征变量进行整合以去除干扰信息,共提取到 39 个特征变量,这些特征变量可以包含该中级融合光谱数据全部变量 100% 的信息;最后采用 SVM 方法对中级融合光谱特征变量进行建模分析。5 种重质矿物油在不同维度下的总体分类识别准确率如图 5 所示。

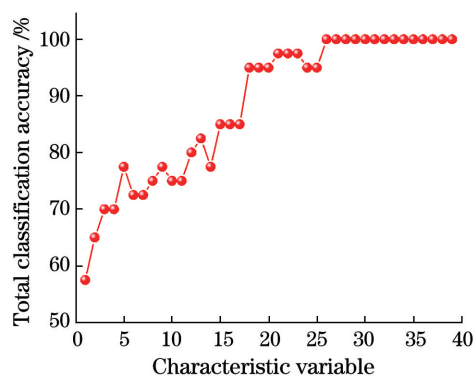


图 5 中级融合光谱在不同维度特征变量下的总体分类识别准确率

Fig. 5 Total classification accuracy of intermediate fusing spectrum under different dimensional characteristic variables

由图 5 可知,对中级融合光谱特征变量进行 SVM 建模分析,在 26 维矩阵上 5 种重质矿物油的分类识别准确率能够达到 100%,实现完全区分,且分类速度远快于基于初级融合光谱的 SVM 模型。汽机油、柴油油、润滑脂、齿轮油和液压油这 5 种重质矿物油的空间分布结果如图 6 所示,其中 X、Y 和 Z 表示 SVM 模型三维分类空间的特征轴。

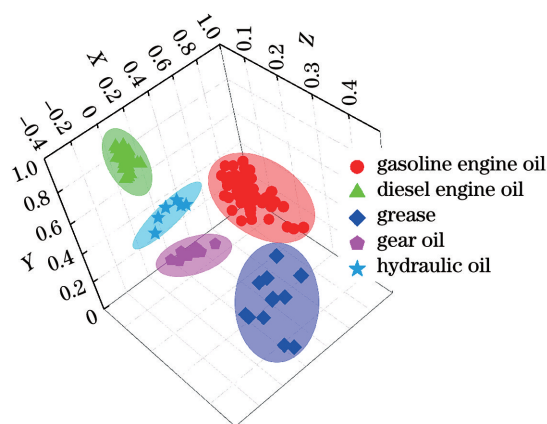


图 6 5 种重质矿物油的空间分布图

Fig. 6 Spatial classification details of five kinds of heavy mineral oils

如图 6 所示,5 种重质矿物油共计 120 份样本实现了 100% 的准确区分,其中:柴油油和汽机油样本的聚敛程度较高,分布较为集中,表明其中所

含的基础油和添加剂的种类和含量差异较小; 润滑脂的 12 份样本分布得较为分散, 表明其所含成分的种类和含量差异较大; 齿轮油的 9 份样本和液压油的 6 份样本分布间隔较近, 表明两种重质矿物油的成分较为相似, 这是使用单一光谱数据建模时难以对液压油进行准确识别和分类的原因。

4 案例分析

2020 年 3 月 18 日, 北京市西城区某地发生一

起交通肇事案件。公安机关在现场提取到疑似重质矿物油检材两份, 分别命名为 JC-1 和 JC-2, 经技术部门检验后获得了两份检材的傅里叶变换红外光谱和拉曼光谱, 如图 7 所示。

由图 7 可知, JC-2 的红外光谱在波数为 2320 cm^{-1} 处有一左高右低的双峰, 而在 JC-1 的相同位置未发现吸收峰。通过比较 JC-1 和 JC-2 的拉曼光谱可知, 两者在波数为 1600 cm^{-1} 和 1350 cm^{-1} 处的峰形不同, 同时在波数为 1450 cm^{-1} 处的相对峰高有所差异。

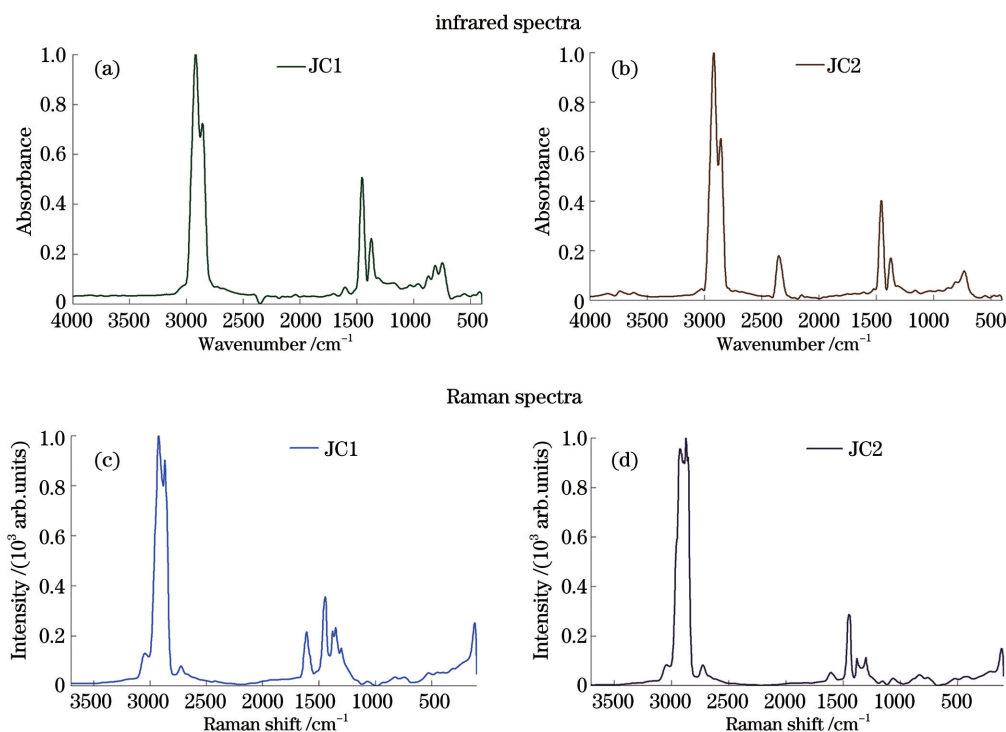


图 7 JC-1 和 JC-2 的红外光谱和拉曼光谱

Fig. 7 Infrared and Raman spectra of JC-1 and JC-2

以上述 120 份已知样本作为训练集, 两份未知检材作为验证集, 采用中级光谱融合结合 SVM 建模方法构建分类判别模型, 两种待判定检材的分类

结果如表 2 所示。结果显示, 训练集的分类识别准确率为 100%, 预测 JC-1 为齿轮油, JC-2 为汽机油。最终破案后, 经比对检验证实了预测结果的正确性。

表 2 两种待判定检材的分类结果

Table 2 Classification of two samples to be determined

Sample to be determined	Gasoline engine oil	Diesel engine oil	Grease	Gear oil	Hydraulic oil
JC-1				✓	
JC-2	✓				

5 结 论

本研究以市面上常见品牌的 5 种重质矿物油共计 120 份样本为对象, 采用红外-拉曼光谱融合的相关方法, 借助支持向量机这一数学模型开展模式识

别。结果表明, 使用单一红外光谱或拉曼光谱数据构建分类模型时的分类识别准确率较低, 而采用红外-拉曼光谱融合方法结合 PCA 降维能够实现 5 种重质矿物油的快速准确区分和认定, 分类识别准确率可达 100%。

与其他针对矿物油研究中使用的 GC/MS 方法相比,使用光谱分析结合化学计量学方法能够快速、准确且无损地识别和区分案发现场遗留的重质矿物油检材,在短时间内实现分类结果的可视化,能够满足公安机关对检测时间和证据保存的相关要求。另外,使用光谱数据融合技术能够克服单一光谱数据信息不完整的问题,实现两类光谱的优势互补,提高样本的分类识别准确率。在后续研究中拟增加重质矿物油的种类和数量,继续完善重质矿物油分类判别模型,为公安机关检验相关检材提供理论支撑和方法参考。

参 考 文 献

- [1] Zheng J G, Chen Y F, Zhong H N, et al. Advances in analysis techniques for mineral oil hydrocarbons in foods and food contact materials [J]. *Journal of Instrumental Analysis*, 2019, 38(11): 1393-1399.
郑建国, 陈燕芬, 钟怀宁, 等. 食品及食品接触材料中烃类矿物油分析技术进展 [J]. *分析测试学报*, 2019, 38(11): 1393-1399.
- [2] Zhong H N, Zhu L, Lu Q, et al. Toxicology and risk management of the mineral oil hydrocarbons of food contact material [J]. *Chinese Journal of Food Hygiene*, 2019, 31(3): 284-290.
钟怀宁, 朱蕾, 卢倩, 等. 食品接触材料中烃类矿物油毒性和风险管理 [J]. *中国食品卫生杂志*, 2019, 31(3): 284-290.
- [3] Zhang Z Y, Wang G, Wang X X, et al. Using SPME-GC/MS to identify the lubricating oil from combustion residues likely found at the arson spot [J]. *Forensic Science and Technology*, 2016, 41(5): 395-397.
张振宇, 王冠, 王欣欣, 等. 火场润滑油燃烧残留物的 SPME-GC/MS 检验方法研究 [J]. *刑事技术*, 2016, 41(5): 395-397.
- [4] Zhang Y G, Xiao Y P, Zhao Y L, et al. Experimental research on high temperature oxidation of aviation engine oil by GC/MS [J]. *Lubrication Engineering*, 2014, 39(7): 77-82.
张永国, 校云鹏, 赵媛莉, 等. 航空发动机润滑油高温氧化 GC/MS 试验研究 [J]. *润滑与密封*, 2014, 39(7): 77-82.
- [5] Eschner M S, Gröger T M, Horvath T, et al. Quasi-simultaneous acquisition of hard electron ionization and soft single-photon ionization mass spectra during GC/MS analysis by rapid switching between both ionization methods: analytical concept, setup, and application on diesel fuel [J]. *Analytical Chemistry*, 2011, 83(10): 3865-3872.
- [6] Yang Q L, Deng X J, Sun X D, et al. Application and research progress of spectral data fusion technology in food testing [J]. *Science and Technology of Food Industry*, 2020, 41(18): 324-329.
杨巧玲, 邓晓军, 孙晓东, 等. 光谱数据融合技术在食品检测中的应用研究进展 [J]. *食品工业科技*, 2020, 41(18): 324-329.
- [7] Sun F, Chen Y, Wang K Y, et al. Identification method of ginger-processed *Pinelliaternata* based on infrared spectroscopy data fusion [J]. *Journal of Beijing University of Traditional Chinese Medicine*, 2019, 42(10): 862-868.
孙飞, 陈雨, 王凯洋, 等. 基于红外光谱数据融合的姜半夏鉴别方法研究 [J]. *北京中医药大学学报*, 2019, 42(10): 862-868.
- [8] Yao S, Li T, Liu H G, et al. Identification of geographical origin of *Boletus tomentipes* by multi-spectral data fusion [J]. *Food Science*, 2018, 39(8): 212-217.
姚森, 李涛, 刘鸿高, 等. 多光谱数据融合技术对绒柄牛肝菌产地的鉴别 [J]. *食品科学*, 2018, 39(8): 212-217.
- [9] Hou W, Wang J F. Rapid identification of the black marker ink based on infrared fingerprint spectroscopy [J]. *Laser Technology*, 2020, 44(4): 436-440.
侯伟, 王继芬. 基于红外指纹光谱的快速鉴别黑色记号笔墨水 [J]. *激光技术*, 2020, 44(4): 436-440.
- [10] He X L, Wang J F, Wang F, et al. Rapid identification of rubber particles based on second derivative infrared spectra [J]. *China Measurement & Test*, 2019, 45(9): 60-64, 83.
何欣龙, 王继芬, 王飞, 等. 二阶导数红外光谱快速鉴别轮胎橡胶颗粒 [J]. *中国测试*, 2019, 45(9): 60-64, 83.
- [11] He X L, Wang J F, Zhang Q, et al. Infrared spectroscopy analysis of marker ink based on multi-classification model [J]. *Chemistry*, 2019, 82(2): 169-174.
何欣龙, 王继芬, 张倩, 等. 基于多分类模型的记号笔墨水红外光谱分析 [J]. *化学通报*, 2019, 82(2): 169-174.
- [12] Shen X C, Xu L, Ye S B, et al. Automatic baseline correction method for the open-path Fourier transform infrared spectra by using simple iterative averaging [J]. *Optics Express*, 2018, 26(10): A609-A614.
- [13] He X L, Wang J F, Li Q S, et al. Identification of vehicle bumper debris based on multi-layer perception-Fisher discriminant and infrared spectroscopy [J]. *China Measurement & Test*, 2019,

- 45(5): 74-78,92.
何欣龙, 王继芬, 李青山, 等. 基于多层感知器-Fisher 判别分析的车用保险杠红外光谱鉴别[J]. 中国测试, 2019, 45(5): 74-78,92.
- [14] He X L, Wang J F, Wu F L, et al. Identification of the infrared spectra of tire rubber based on chemometrics [J]. *Journal of Analytical Science*, 2019, 35(3): 357-361.
何欣龙, 王继芬, 吴福璐, 等. 基于化学计量学的橡胶颗粒红外光谱分析鉴别[J]. 分析科学学报, 2019, 35(3): 357-361.
- [15] Johnson J, Mani J, Ashwath N, et al. Potential for Fourier transform infrared (FTIR) spectroscopy toward predicting antioxidant and phenolic contents in powdered plant matrices [J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2020, 233: 118228.
- [16] Shimbo G, Tagawa M, Matsumoto K, et al. Effects of scatter correction processing on image quality of portable thoracic radiography in calves [J]. *The Japanese Journal of Veterinary Research*, 2018, 66(2): 105-112.
- [17] Sivagami, Vaishali A, Ramakrishnan R, et al. Weather prediction model using Savitzky-Golay and Kalman filters [J]. *Procedia Computer Science*, 2019, 165: 449-455.
- [18] Qin J Q, Kong X Y, Sun X R. Influence of different data normalizing methods on the performance of Sevcik fractal dimension algorithm [J]. *Chinese Journal of Scientific Instrument*, 2016, 37(7): 1485-1491.
秦建强, 孔祥玉, 孙喜荣. 数据标准化对 Sevcik 分形维数算法的性能影响[J]. 仪器仪表学报, 2016, 37(7): 1485-1491.
- [19] Li D, Jin Y Y, Tong Y, et al. Intelligent detection and defect classification of infusion bags based on support vector machine [J]. *Laser & Optoelectronics Progress*, 2019, 56(13): 131502.
李丹, 金媛媛, 童艳, 等. 基于支持向量机的输液袋智能检测与缺陷分类[J]. 激光与光电子学进展, 2019, 56(13): 131502.
- [20] Chen Y, Yan X, Zhang X, et al. Surface-enhanced Raman spectroscopy quantitative analysis of polycyclic aromatic hydrocarbons based on support vector machine algorithm [J]. *Chinese Journal of Lasers*, 2019, 46(3): 0311005.
陈阳, 严霞, 张旭, 等. 基于支持向量机算法的多环芳烃表面增强拉曼光谱的定量分析[J]. 中国激光, 2019, 46(3): 0311005.
- [21] Wang X F, Zhang X Y, Xu X H. Comparison of multi-factor-considered blood glucose prediction models by near-infrared spectroscopy [J]. *Laser & Optoelectronics Progress*, 2019, 56(4): 041701.
王晓飞, 张欣怡, 徐馨荷. 考虑多种因素的近红外光谱血糖预测模型对比[J]. 激光与光电子学进展, 2019, 56(4): 041701.