

激光与光电子学进展

基于甲板漆的红外光谱预处理方法比较

古锬山, 王继芬*

中国人民公安大学侦查学院, 北京 100038

摘要 甲板漆是海事刑事案件中重要的物证之一, 对甲板漆进行快速无损准确的鉴定是法庭科学物证检验的重要分支。红外光谱是常用的快速无损鉴定方法, 红外光谱的预处理能够消除谱图的噪声和背景干扰, 提高谱图的识别率。实验收集了常见的 5 种不同品牌的甲板漆共计 100 个样本, 利用支持向量机和贝叶斯判别分析对多个红外预处理组合处理后的甲板漆谱图进行识别, 根据识别率来筛选甲板漆红外预处理的最优方法, 同时对支持向量机和贝叶斯判别分析的识别效果进行比较。结果表明, 贝叶斯判别分析的效果要优于支持向量机, 贝叶斯判别分析对经 Savitzky-Golay 二阶求导平滑和傅里叶自去卷积处理的甲板漆谱图训练集识别率为 100%, 测试集识别率为 96%。该方法分类效果好, 定性能力强, 对法庭科学中甲板漆的鉴定有一定的借鉴意义。

关键词 光谱学; 甲板漆; 光谱预处理; 化学计量学

中图分类号 O657.33

文献标志码 A

DOI: 10.3788/LOP202259.1030004

Comparison of Infrared Spectral Pretreatment Methods Based on Deck Paint

Gu Kunshan, Wang Jifen*

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

Abstract Deck paints are important evidence materials in maritime criminal cases. The rapid, accurate, and nondestructive identification of deck paints is crucial in forensic science. Infrared spectra are commonly used for such a fast and nondestructive identification analysis. Pretreatments can eliminate the noise and background interference of the spectra and improve the recognition rate of the dataset. In an experiment, 100 samples from five brands of deck paints were collected. Support vector machine (SVM) and Bayes discriminant analysis (BDA) were used to identify the deck paint spectra after several infrared pretreatment combinations, and the optimal infrared pretreatment method was screened according to the recognition rate. Results show that the effect of BDA is better than that of SVM. The recognition rate of the training dataset of the deck paint spectra processed using Savitzky-Golay second derivative smoothing and Fourier self-deconvolution is 100%, and the recognition rate of the test dataset is 96%. The selected method shows a good classification effect and strong qualitative ability and can be used as a reference for identifying deck paints in forensic science.

Key words spectroscopy; deck paint; spectral pretreatment; stoichiometry

1 引言

甲板漆是油漆的一种, 常被用于船舶甲板上起

到防水、吸附底漆的作用。船舶甲板油漆作为法庭科学中常见的微量物证, 通常会出现在涉及船只碰撞、非法运输以及劫持船只等海事案件中。该类物

收稿日期: 2021-04-13; 修回日期: 2021-05-11; 录用日期: 2021-05-20

通信作者: *wangjifen58@126.com

证通常可以在犯罪嫌疑人的衣物、鞋底以及集装箱货柜的表面等进行提取^[1]。检验人员对油漆碎片进行分析从而得到甲板油漆的厂家和类型等相关信息,为执法人员追踪嫌疑逃犯以及案件调查提供线索和方向,缩小侦察范围。

目前对于甲板油漆检验方法有很多,例如热重分析法^[2-3]、扫描电镜能谱法^[4-6]、电子探针法^[7]以及光谱分析法^[8-10]等。物证的无损鉴定是法庭科学的基础。然而对于上述方法来说除了红外光谱法以外,其他方法均存在实验仪器昂贵、操作复杂以及反应过程耗时费力等问题;并且最重要的是红外光谱法能够实现对物证的快速无损鉴别,而其他方法均会消耗一定量的检材,一旦检材受损将直接影响案件的进一步调查。然而当样本数量增多时,直接通过肉眼对谱图进行分析会导致主观性强、误差大等问题。因此近年来将机器学习和红外谱图分析相结合的研究有很多,例如 Ferreira 等^[11]利用高光谱成像可见光/近红外光谱技术(HSI-UV/VIS/NIR)对收集的 12 个不同品牌共计 38 个油漆样品进行分析;同时还评估了该技术与主成分分析(PCA)结合作为鉴定汽车油漆的潜力。PCA 在不同的颜色数据集中单独执行,可以 100% 区分来自不同品牌或来自同一制造商的不同样本。此外,利用从不同来源的相同品牌、颜色的车辆采集油漆样本碎片,成功地检验了 PCA 在识别相似漆片方面的可靠性;表明 HSI-UV/VIS/NIR 可以作为其他光谱仪的快速、简单和高效的替代技术,用于汽车油漆的准确识别。谢有超等^[12]采集了 10 种烟丝的红外光谱数据,利用连续小波变换(CWT)方法进行预处理以及概率主成分分析(PPCA)方法进行光谱数据降维;比较了烟丝红外光谱数据的支持向量机(SVM)和线性判别分析两种模型;结果表明,支持向量机结合 CWT 以及 PPCA 对烟丝的识别率高达 97.2%。充分证明机器学习结合红外光谱分析是可行的,结合以后能够很好地消除人为导致的误差。

光谱预处理算法主要是指利用求导、平滑以及滤波等方法减少由于仪器或者样本本身所导致的红外光谱谱峰重叠以及基线漂移等干扰现象,同时消除谱图的噪声和背景干扰,从而提高模型对红外光谱识别的稳定性以及准确率^[13]。常见的求导算法有一阶导数^[14]以及二阶导数^[14]等;常见的平滑算法有 Savitzky-Golay(S-G)平滑^[14]以及 Norris 平滑^[15]

等;常见的滤波算法有傅里叶自去卷积(FSD)^[16]以及 CWT^[17]等。

利用机器学习方法来选择光谱预处理效果的比较报道相对较少,例如王淑贤等^[18]利用偏最小二乘法比较了不同预处理模型下普洱茶中添加香精类型的预测准确度,同时还利用预测误差均方根对各预处理模型的稳定性进行检验。实验结果表明,该方法能够有效地筛选出普洱茶的光谱最佳预处理手段,提高模型的精密度;Liu 等^[19]利用肯纳德·斯通(KS)算法对土壤有机质的求导平滑等 6 种光谱预处理方法进行筛选,并利用均方根误差(RMSE)进行验证。实验结果表明,通过 KS 算法筛选出来的最佳预处理方法可以优化荷电状态(SOC)模型的性能,充分证明利用机器学习方法来筛选光谱预处理方法是切实可行的。与此同时,关于利用机器学习方法对甲板漆的单一或多个红外光谱预处理组合效果的比较报道也较少。本实验拟将收集的甲板漆光谱图进行单一或多个光谱预处理;然后建立贝叶斯判别(BDA)和 SVM 模型对预处理后的甲板漆谱图进行识别,最终以正确识别率来比较单一和多预处理方法的效果。

2 实验部分

2.1 实验样本

从市场上购买常见的 5 种不同品牌厂家的甲板漆样本共计 100 份,其中醇酸甲板漆 14 份;酚醛甲板漆 14 份;氧化橡胶甲板漆 18 份;环氧甲板漆 16 份;甲板防滑漆 38 份;按顺序以 P1、P2、P3、P4、P5 来对上述 5 种样本进行标号。

2.2 实验条件及前处理

在温度为 $(25 \pm 3)^\circ\text{C}$, 相对湿度为 30% 的实验条件下,用含乙醇的棉球擦拭甲板漆的样本表面以除去其表面的灰尘等杂质,之后将甲板漆样品放入去离子水中进行超声清洗 3 次,每次约 20 min,清洗完成后利用含乙醇的棉球擦干样品表面的水分。

2.3 实验仪器及设备

傅里叶变换红外光谱仪(FTIR, Nicolet 5700, Thermo Fisher Scientific 公司, 美国);KBr 分束器(美国 Thermo Fisher Scientific 公司)。扫描次数为 32 次,光谱分辨率为 4 cm^{-1} ,波数采集范围为 $4000 \sim 400\text{ cm}^{-1}$ ($500 \sim 400\text{ cm}^{-1}$ 处光谱噪声大,故将其剔除),每个样本采集 3 次光谱曲线,取平均值作为最终的光谱数据^[20-21]。

2.4 实验预处理

将采集的红外光谱统一进行峰面积归一化^[22]、多元散射校正^[23]以及自动基线校准^[24]处理后得到初步预处理的光谱,然后将初步预处理的光谱进行单一或多个预处理组合(求导处理、S-G平滑处理、Norris平滑处理以及傅里叶自去卷积处理),表1为得到的预处理组合,将10种预处理方法用M1、M2、...、M10代替,即M1为N,M2为SGFD,M3为SGSD,M4为NFD,M5为NSD,M6为N+FSD,M7为SGFD+FSD,M8为SGSD+FSD,M9为NFD+FSD,M10为NSD+FSD。

表1 光谱的预处理组合

Table 1 Pretreatment combination of spectra

No.	Pretreatment method	Abbreviation
1	None	N
2	S-G 1 st derivative	SGFD
3	S-G 2 nd derivative	SGSD
4	Norris 1 st derivative	NFD
5	Norris 2 nd derivative	NSD
6	None+FSD	N+FSD
7	S-G 1 st derivative+FSD	SGFD+FSD
8	S-G 2 nd derivative+FSD	SGSD+FSD
9	Norris 1 st derivative+FSD	NFD+FSD
10	Norris 2 nd derivative+FSD	NSD+FSD

2.5 实验流程图

本实验的流程图如图1所示。

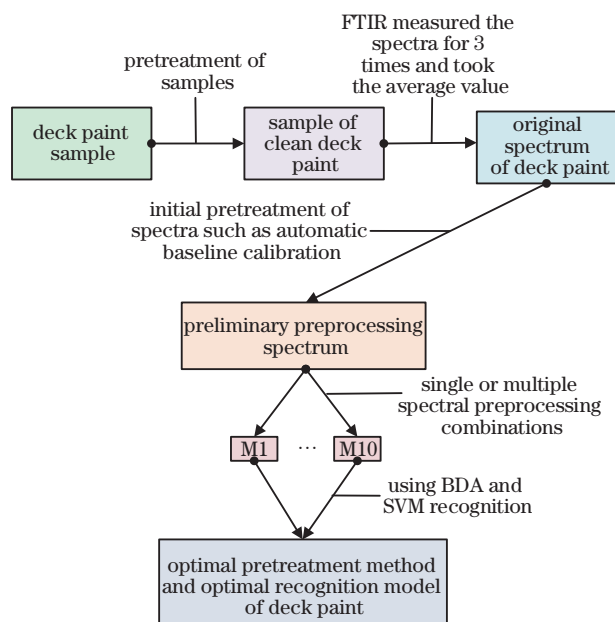


图1 实验流程图

Fig. 1 Experimental flow chart

3 实验建模

3.1 支持向量机

支持向量机是统计学中一种常见的二分类模型,该模型主要通过线性分类器来获得特征空间上的最大间隔,其目的是寻找一个线性函数即超平面来分割总体样本,分割的原则是使各样本的间隔最大化,分割以后最终转变为求样本的凸二次规划问题的最优解^[25-27]。该模型主要有三种类型:1)当样本线性可分时,通过硬间隔使得样本之间的间隔最大化,得到一个线性可分SVM;2)当样本接近线性可分时,通过软间隔使得样本之间的间隔最大化,得到一个线性SVM;3)当样本根本线性不可分时,通过软间隔以及核技巧使得样本之间的间隔最大化,得到一个非线性SVM。其原理如下。

任意一个超平面都可用下面的线性方程来表述,

$$\mathbf{w}^T \mathbf{x} + b = 0, \quad (1)$$

式中: \mathbf{w} 为超平面的范数; b 为截距。

n 维空间,点 $\mathbf{x}=(x_1, x_2, \dots, x_n)$ 到直线 $\mathbf{w}^T \mathbf{x} + b = 0$ 的距离为

$$\frac{|\mathbf{w}^T \mathbf{x} + b|}{\|\mathbf{w}\|}, \quad (2)$$

式中: $\|\mathbf{w}\| = \sqrt{w_1^2 + w_2^2 + \dots + w_n^2}$

根据SVM定义可知,支持向量到超平面的距离为 D ,其余点到超平面的距离大于 D ,于是可以得到:

$$\begin{cases} \frac{|\mathbf{w}^T \mathbf{x} + b|}{\|\mathbf{w}\|D} \geq 1, y = 1 \\ \frac{|\mathbf{w}^T \mathbf{x} + b|}{\|\mathbf{w}\|D} \leq -1, y = -1 \end{cases}. \quad (3)$$

令 $\|\mathbf{w}\|D = 1$ 于是有:

$$\begin{cases} \mathbf{w}^T \mathbf{x} + b \geq 1, y = 1 \\ \mathbf{w}^T \mathbf{x} + b \leq -1, y = -1 \end{cases}. \quad (4)$$

将上述两方程合并,得到最优化约束式:

$$\min f(\mathbf{w}) = \min \frac{1}{2} \|\mathbf{w}\|^2, \text{ s.t. } y_\xi(\mathbf{w}^T \mathbf{x}_\xi + b) \geq 1, \quad \xi = 1, 2, \dots, n. \quad (5)$$

为解决上述约束式引入拉格朗日乘子 α_ξ ,且 $\alpha_\xi \geq 0$,代入式(5),于是无约束的拉格朗日目标函数为

$$L(\mathbf{w}, b, \alpha) = \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{\xi=1}^n \alpha_\xi [y_\xi(\mathbf{w}^T \mathbf{x}_\xi + b) - 1]. \quad (6)$$

为了得到求解的具体形式,令 $L(\mathbf{w}, b, \alpha)$ 对 \mathbf{w} 和 b 求偏导数,并令其为 0 可以得到:

$$\begin{cases} \mathbf{w} = \sum_{\xi=1}^n \alpha_{\xi} y_{\xi} x_{\xi} \\ \sum_{\xi=1}^n \alpha_{\xi} y_{\xi} = 0 \end{cases} \quad (7)$$

将上述两个等式代入目标函数消去 \mathbf{w} 和 b 得到 SVM 分类表达式:

$$L(\mathbf{w}, b, \alpha) = -\frac{1}{2} \sum_{\xi=1}^n \sum_{j=1}^n \alpha_{\xi} \alpha_j y_{\xi} y_j (x_{\xi} \cdot x_j) + \sum_{\xi=1}^n \alpha_{\xi} \quad (8)$$

3.2 贝叶斯判别

距离判别分析是一种简单实用的统计学方法,但是该方法没有把每个样本出现的概率考虑在内,也没有考虑到错误判别带来的影响。贝叶斯判别分析能够很好地解决距离判别分析带来的两个问题。

贝叶斯判别是统计学中的经典判别分析方法,其过程主要是先根据总体样本的先验概率得到样本的概率密度函数,然后利用贝叶斯判别公式计算各样本落入各子域的概率即后验概率,最后根据后验概率的大小进行判别,后验概率大的即为该样本的类型。其主要原理如下。

设有一 p 维样本,分别记为 T_1, T_2, \dots, T_k ,各样本的总体概率密度函数分别记为 $f_1(x), f_2(x), \dots, f_k(x)$,各样本先验概率为 $p_j = P(T_j)$, $\sum_{j=1}^k P_j = 1$,可用贝叶斯判别公式来判断第 i 个样本的后验概率:

$$P(T_i|x) = \frac{p_i P(x|T_i)}{\sum_{j=1}^k p_j P(x|T_j)} = \max P(x|T_j), x \in T_i \quad (9)$$

贝叶斯判别准则:将待判样本输入贝叶斯判别式,如果在所有样本的后验概率中 $P(T_i|x)$ 是最大的,则该样本判定为属于 i 样本^[28]。

4 结果与讨论

4.1 光谱预处理谱图比较

部分甲板漆样本的原始谱图、SGFD 谱图、SGSD 谱图、NFD 谱图、NSD 谱图、N+FSD 谱图、SGFD+FSD 谱图、SGSD+FSD 谱图、NFD+FSD 谱图以及 NSD+FSD 谱图如图 2 所示。

观察图 2 可以得知,未使用傅里叶自去卷积处理的原始谱图和导数谱图的出峰位置大致相同。原始谱图在 1700 cm^{-1} 左右处有一个很强的尖峰,是 C=O 伸缩振动吸收峰;在 1300 cm^{-1} 左右处有一个较强的尖峰,是 C—H 弯曲振动吸收峰;在约 1100 cm^{-1} 处有一个左低右高的强峰;在 700 cm^{-1} 左右附近有一个左高右低的中强峰。经过求导处理后的谱图能将原始谱图种的重叠吸收峰分离出来,使得谱图的各吸收峰更明显。使用傅里叶自去卷积处理后的谱图会与原谱图产生很大的差异,可以看出,经过傅里叶自去卷积处理的谱图其信号峰的峰形会变窄和变尖,但是其出峰的位置以及峰面积没有发生变化^[29]。谱图通过傅里叶自去卷积处理能够将原本谱图中信号不明显的地方进行一定的放大,使得谱图更具辨识度。但是仅仅通过人眼来对谱图进行区分会产生很强的主观性和误差,因此需要引进机器学习来帮助识别谱图,并对预处理方法进行准确的评估。

4.2 支持向量机建模分析

SVM 处理线性的数据可以通过式(8)进行求解,而大部分的数据都是非线性的。此时则需要引入核函数,通过核函数将非线性的原始数据映射到高维空间,从而解决 SVM 中的数据非线性问题。

SVM 中常见的核函数有线性核、多项式核、径向基函数(RBF)核以及拉普拉斯核等。本实验通过预处理实验对比了上述四种核函数的效果。通过不断增加迭代的次数,得到不同迭代数下核函数的最小分类误差,从而筛选出最优核函数。图 3 为 4 种核函数的最小分类误差图。

如图 3 所示,4 种核函数的最小分类误差随着迭代数的增加呈现出不同的变化。其中线性核函数、拉普拉斯核函数以及多项式核函数的最小分类误差值随着迭代数的变化会产生较大的上下波动。而径向基核函数的最小分类误差会随着迭代数的变化波动比较稳定,整体呈现先下降后稳定的趋势,且在迭代数 18 以后出现最小的分类误差值 0.21398。该值不仅是径向基核函数的最小分类误差值,还是 4 种核函数的最小分类误差值,因此本实验拟采用径向基核函数作为 SVM 的核函数,其余参数为 Matlab 2019b 系统默认,即惩罚系数 C 和 Gamma 值为 0.1。将不同预处理得到的数据输入 SVM,得到的实验结果如图 3 和图 4

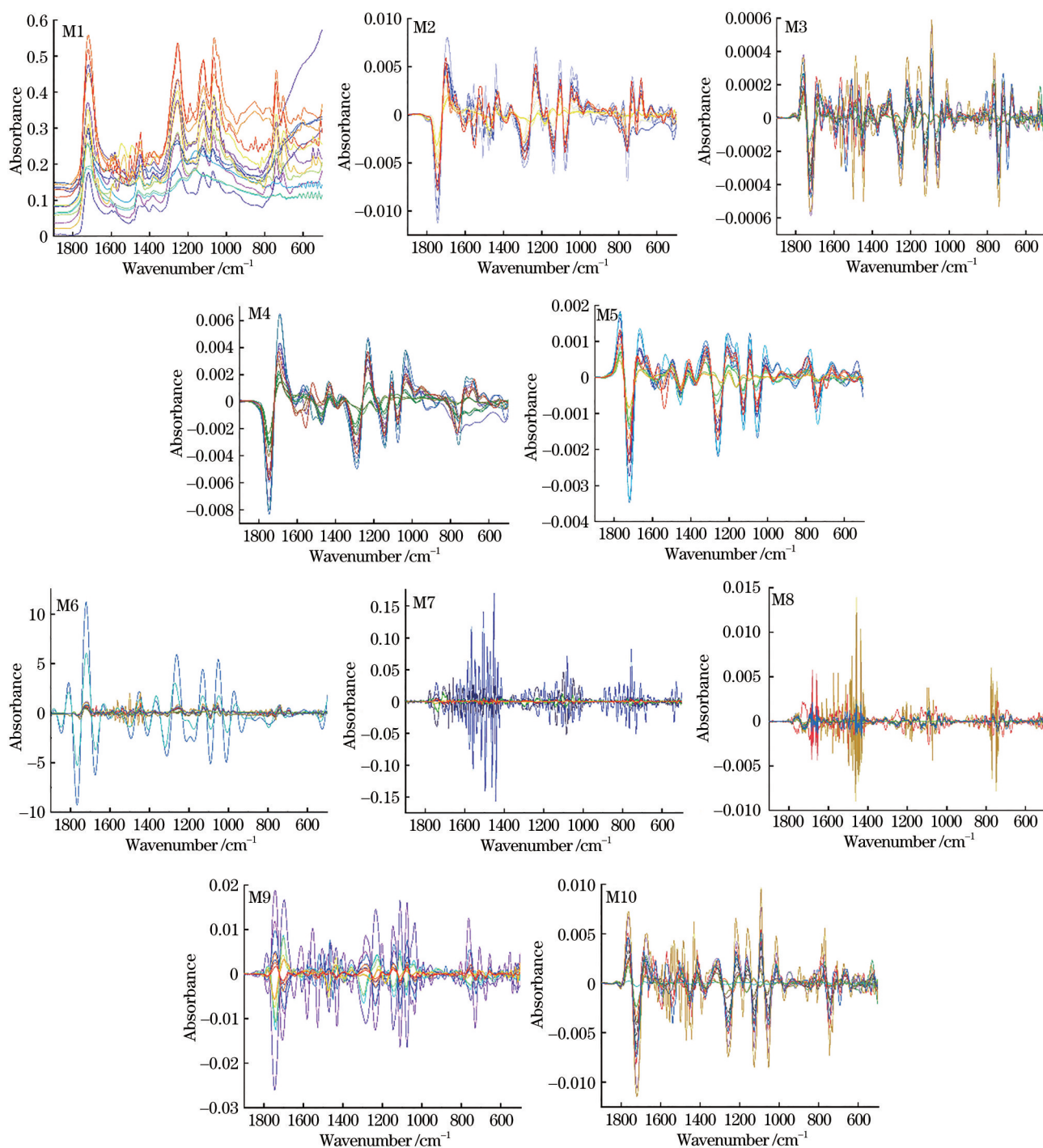


图 2 部分甲板漆的原始和预处理红外光谱图

Fig. 2 Original and pretreated infrared spectra of some deck paints

所示。

观察图 4 可以得知, 5 种甲板漆样本在不同预处理的条件下识别率差异比较大。P4 和 P5 样本在各预处理条件下均有很高的识别率, P5 样本只有在 NFD+FSD 和 SGSD+FSD 处理下没有达到 100% 的识别率, P4 样本只有在 NSD+FSD、NFD+FSD 以及 SGFD+FSD 处理的条件下没有

达到 100% 识别率, 其余预处理的条件下均能达到 100% 的识别率; P3 样本的识别率比较平均但均未达到 100%, P1 和 P2 样本的识别效果不理想。P1 和 P2 识别效果不理想的原因可能与样本的类型和数目有关。样本的不均匀很可能导致 SVM 在识别的过程中把样品数目少的类型识别为样本数目多的类型; 由于在前期采集甲板漆样本时, P1 和

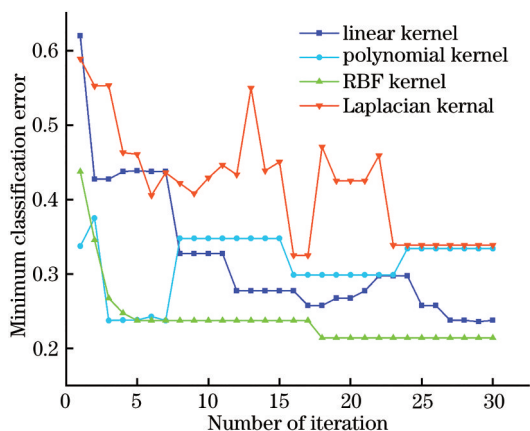


图 3 4种核函数的最小分类误差图

Fig. 3 Minimum classification error diagram of four kinds of kernel functions

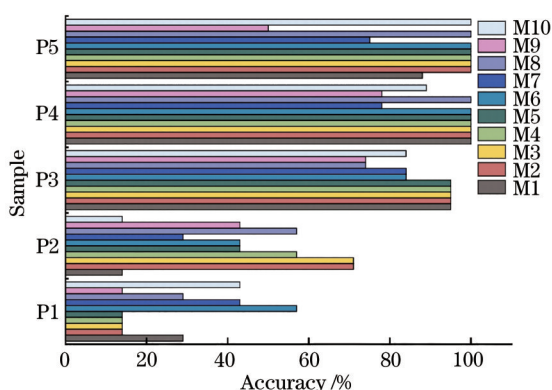


图 4 5种甲板漆在不同预处理方法下的识别率

Fig. 4 Recognition rate of five kinds of deck paints under different pretreatment methods

P2样本相对于P4样本来说采集偏少,导致在本实验中观察SVM识别的混淆矩阵时发现有部分P1、P2样本在分类时被误判为P4样本,这是导致P1和P2样本识别率较低的原因之一。观察SVM判别的混淆矩阵时还发现P1样本和P2样本之间也存在着大量的相互误判,即P1样本被识别为P2样本,P2样本被识别为P1样本;可能是P1和P2样本之间的特征差异太小导致SVM难以对其进行类型的定性,这也是导致P1和P2样本识别效果不理想的另一原因。

由图5可以发现,原始谱图经过求导平滑或只进行傅里叶自去卷积处理后的总体分类识别率均有一定程度的增加,表明这两种预处理手段能够优化甲板漆的光谱识别率。同时还发现导数光谱经过Savitzky-Golay平滑的效果要优于Norris平滑的效果,表明Savitzky-Golay平滑更适用于甲板漆的光谱预处理。然而当求导平滑以及傅里叶自去卷

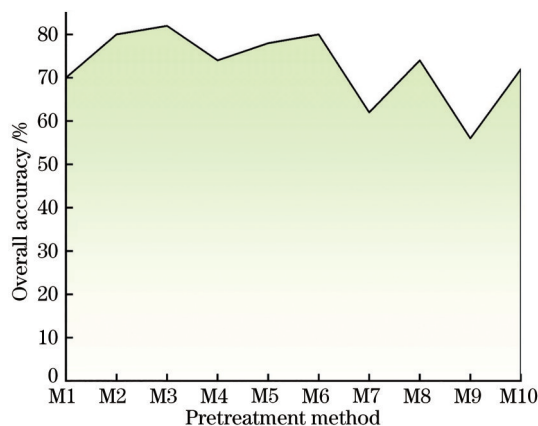


图 5 SVM对于不同预处理方法的总体分类识别率

Fig. 5 Overall classification recognition rate of SVM for different preprocessing methods

积处理相结合时其效果却远不如两种方法的单独使用,原因可能是通过求导平滑得到的光谱能够将原谱图中重叠的吸收峰分离出来,而傅里叶自去卷积处理则将各吸收峰的峰形变窄;使得求导平滑后分离的吸收峰经过傅里叶自去卷积处理后,各吸收峰之间的差异不明显,从而导致SVM对于这两种预处理方法叠加后的识别率大大下降。

4.3 贝叶斯判别分析

将100个甲板漆样本分成5份,其中4份样本作为训练样本(训练集),1份样本作为测试样本(测试集),测试集采用交叉验证的方法进行检验。图6为训练集和测试集在不同预处理条件下的识别率,图7为训练集和测试集在不同预处理条件下的误差棒。

观察图6可知,经过求导平滑处理或傅里叶自去卷积处理后的谱图识别率要高于原始谱图,且二

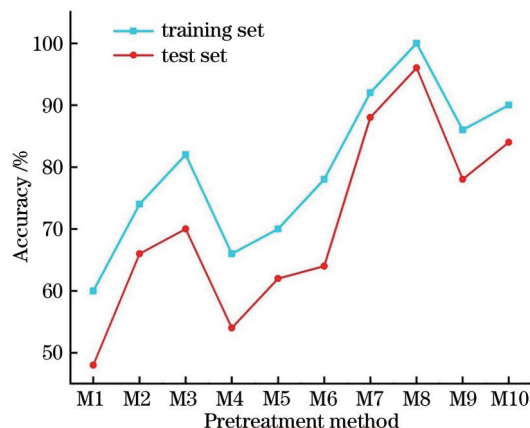


图 6 训练集和测试集在不同预处理方法下的识别率

Fig. 6 Recognition rate of training set and test set under different preprocessing methods

阶导数的识别率要高于一阶导数;在谱图识别率方面, Savitzky-Golay 平滑的效果要优于 Norris 平滑。贝叶斯判别对上述两种预处理方法结合后的识别效果更佳, 在 M8(SGSD+FSD) 预处理方法下达到最高的 100% 的训练集识别率和 96% 的测试集识别率。误差棒表示的是偏差情况, 误差棒越长, 说明两者之间的波动越大。观察图 7 可知, M8 预处理方法的误差棒最小, 表明 M8 预处理方法的训练集和测试集的波动很小, 进一步说明了 M8 预处理方法的稳定性要优于其他预处理方法。下面将进一步分析 M8 预处理条件下的贝叶斯判别分析模型。

M8 预处理条件下, 5 种甲板漆的贝叶斯判别函数分别为

$$\begin{cases} Z_1 = -202313.797x_1 - 433587.345x_2 + 193072.659x_3 - 295155.805x_4 \cdots - 221711.254x_{39} - \\ \quad 314981.164x_{40} + 409187.756x_{41} + 133621.623x_{42} - 52.401 \\ Z_2 = -591293.934x_1 - 1375638.832x_2 + 1261753.686x_3 + 521045.764x_4 \cdots - 725806.122x_{39} - \\ \quad 553086.479x_{40} + 1809836.231x_{41} + 689907.924x_{42} - 623.588 \\ Z_3 = 276477.901x_1 + 346793.925x_2 - 139975.671x_3 - 542649.165x_4 \cdots + 341224.024x_{39} + \\ \quad 319724.791x_{40} - 543573.729x_{41} - 172050.009x_{42} - 91481 \\ Z_4 = -8614.507x_1 + 194756.593x_2 - 321004.823x_3 - 242634.929x_4 \cdots + 164913.094x_{39} + \\ \quad 74570.081x_{40} - 150126.846x_{41} - 27624.688x_{42} - 29.128 \\ Z_5 = 78185.537x_1 - 143765.219x_2 + 46930.391x_3 + 30072.918x_4 \cdots + 10162.141x_{39} - \\ \quad 91539.982x_{40} + 107060.695x_{41} - 34485.355x_{42} - 21.6 \end{cases} \quad (10)$$

对于未知类型的甲板漆样本, 只需将其数据输入上述 5 个贝叶斯判别函数计算, 所得值最大的即为样本的类型。以函数 Z_1 、 Z_2 以及 Z_3 作为 XYZ 轴构建空间坐标系, M8 预处理条件下 5 种甲板漆样本在空间的分布如图 8 所示。

如图 8 所示, 5 种甲板漆样本完全分离, 表明该预处理条件下模型的识别率达到最优。其中 P1 和 P4 样本内部的分隔很密集, 表明各品牌生产该类型甲板漆的材料上区别不大; 而 P5 样本内部分离比较离散, 表明该类型的甲板漆在成品上存在一定的差异。P2 和 P3 样本被分离较远说明这两种甲板漆之间存在着明显的特征差异, 能够被该模型很好地识别; 而 P1、P4 以及 P5 样本虽然都被该模型所分开, 但是三个类型之间相隔比较近, 表明该模型对于 P1、P4 以及 P5 样本的识别率不如 P2 和 P3 样本, 原因可能是这三种样本之间的特征差异不是特别明显。对于未知的甲板漆样本, 将其对应的数值输入三个函数即可得到样本在空间中的位置。

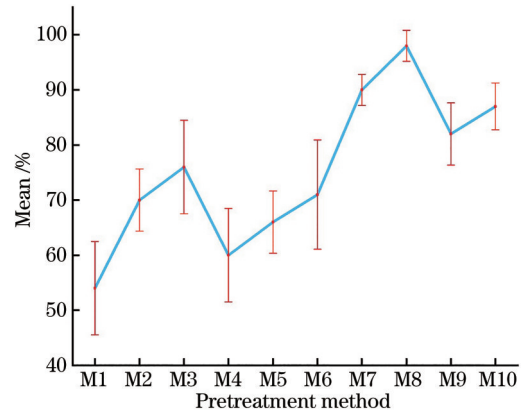


图 7 训练集和测试集在不同预处理方法下的误差棒
Fig. 7 Error bar of training set and test set under different pretreatment methods

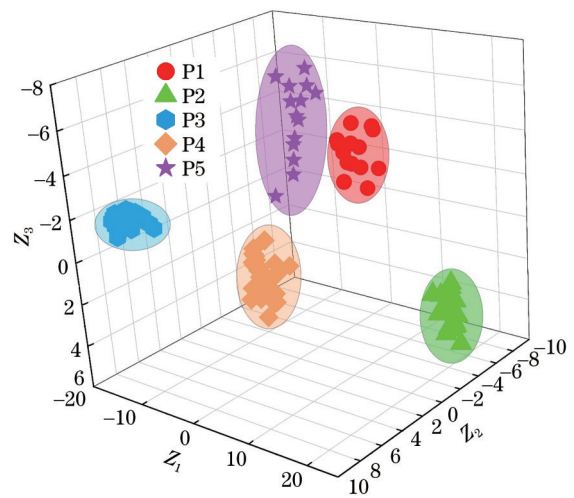


图 8 M8 条件下 5 种甲板漆样本在空间的分布
Fig. 8 Spatial distribution of five kinds of deck paint samples under M8 condition

5 结 论

本实验将甲板漆的红外原始谱图与单一或多

个红外预处理技术 (Savitzky-Golay 求导平滑、Norris 求导平滑和傅里叶自去卷积) 相结合, 利用 SVM 和贝叶斯判别分析对不同预处理组合处理后的甲板漆谱图进行识别, 根据 SVM 和贝叶斯判别分析的识别率来筛选甲板漆的预处理方法, 同时比较 SVM 和贝叶斯判别分析对于甲板漆谱图识别的适用性。结果表明, 贝叶斯判别分析要比 SVM 更适用于甲板漆的谱图识别, 贝叶斯判别分析对 M8 预处理条件的训练集识别率达到 100% 且验证集识别率也高达 96%, 是所有预处理组合中的最大值, 表明该预处理组合是处理甲板漆原始谱图的最优手段。通过该预处理方法结合贝叶斯判别分析能够实现甲板漆红外谱图的快速无损准确鉴别, 避免了主观判断和低效率的情况, 符合法庭科学对于甲板漆鉴别的要求, 为法庭科学对于甲板漆的鉴定提供一定的参考。

参 考 文 献

- [1] He X L, Wang J F, Zhao B, et al. Nondestructive discrimination of ship deck paint using attenuated total reflection-Fourier transform infrared (ATR-FTIR) spectroscopy with chemometric analysis[J]. *Analytical Letters*, 2020, 53(17): 2761-2774.
- [2] 郑尚菊, 姚中栋, 奚建华, 等. 热分析法鉴别油漆的研究[J]. *化学世界*, 1996, 37(2): 77-79.
Zheng S J, Yao Z D, Xi J H, et al. Study on identification of paint by thermal analysis[J]. *Chemical World*, 1996, 37(2): 77-79.
- [3] Cozzarini L, Marsich L, Ferluga A, et al. Failure investigation of a protective epoxy coating by means of crosscheck between infrared spectroscopy and thermal analysis[J]. *Engineering Failure Analysis*, 2020, 107: 104201.
- [4] 崔春妮. 基于多层油漆检验中 SEM/EDS 的应用探究[J]. *化学工程与装备*, 2016(5): 237-238.
Cui C N. Research on the application of SEM/EDS in the inspection of multi-layer paint[J]. *Chemical Engineering & Equipment*, 2016(5): 237-238.
- [5] 门腾腾. 交通事故中机动车油漆的傅里叶红外光谱仪和扫描电镜/能谱仪检验[J]. *山东化工*, 2014, 43(11): 100-103.
Men T T. Examination of motor vehicle paint in traffic accidents by Fourier transform infrared spectrometer and scanning electron microscope/energy spectrometer[J]. *Shandong Chemical Industry*, 2014, 43(11): 100-103.
- [6] 裴茂清, 肖翔, 郭海荣, 等. SEM/EDS 在多层油漆检验中的应用[J]. *广东公安科技*, 2010, 18(1): 27-28.
Pei M Q, Xiao X, Guo H R, et al. Application of SEM/EDS in the inspection of multi-layer paint[J]. *Guangdong Gong'an Keji*, 2010, 18(1): 27-28.
- [7] 段萍, 余然, 杨家明. 电子探针分析法对车辆油漆物证检验的案例[J]. *电子技术与软件工程*, 2017(2): 254.
Duan P, Yu R, Yang J M. A case of vehicle paint evidence inspection by electron probe analysis[J]. *Electronic Technology & Software Engineering*, 2017(2): 254.
- [8] 肖庆, 魏书亚, 付迎春. 红外光谱定量分析古代漆膜的可行性研究[J]. *光谱学与光谱分析*, 2020, 40(9): 2962-2967.
Xiao Q, Wei S Y, Fu Y C. Feasibility study on quantitative analysis of ancient lacquer films by infrared spectroscopy[J]. *Spectroscopy and Spectral Analysis*, 2020, 40(9): 2962-2967.
- [9] Kwofie F, Perera U D N, Allen M D, et al. Application of infrared microscopy and alternating least squares to the forensic analysis of automotive paint chips[J]. *Journal of Chemometrics*, 2021, 35(1): e3277.
- [10] Lavine B, Almirall J, Muehlethaler C, et al. Criteria for comparing infrared spectra—a review of the forensic and analytical chemistry literature[J]. *Forensic Chemistry*, 2020, 18: 100224.
- [11] Ferreira K B, Oliveira A G G, Gonçalves A S, et al. Evaluation of hyperspectral imaging visible/near infrared spectroscopy as a forensic tool for automotive paint distinction[J]. *Forensic Chemistry*, 2017, 5: 46-52.
- [12] 谢有超, 彭黔荣, 杨敏, 等. 基于烟丝近红外光谱的卷烟品牌识别方法[J]. *烟草科技*, 2021, 54(3): 72-77.
Xie Y C, Peng Q R, Yang M, et al. Cigarette brand recognition based on NIR spectral data of cut filler[J]. *Tobacco Science & Technology*, 2021, 54(3): 72-77.
- [13] 余梅, 李尚科, 杨菲, 等. 基于近红外光谱技术与优化光谱预处理的陈皮产地鉴别研究[J]. *分析测试学报*, 2021, 40(1): 65-71.
Yu M, Li S K, Yang F, et al. Identification on different origins of citri reticulatae pericarpium using near infrared spectroscopy combined with optimized spectral pretreatments[J]. *Journal of Instrumental Analysis*, 2021, 40(1): 65-71.
- [14] Savitzky A, Golay M J E. Smoothing and

- differentiation of data by simplified least squares procedures[J]. *Analytical Chemistry*, 1964, 36(8): 1627-1639.
- [15] Yang Y H, Pan T, Zhang J. Global optimization of Norris derivative filtering with application for near-infrared analysis of serum urea nitrogen[J]. *American Journal of Analytical Chemistry*, 2019, 10(5): 143-152.
- [16] Vazhnova T, Lukyanov D B. Fourier self-deconvolution of the IR spectra as a tool for investigation of distinct functional groups in porous materials: Brønsted acid sites in zeolites[J]. *Analytical Chemistry*, 2013, 85(23): 11291-11296.
- [17] Ghasemi J B, Heidari Z, Jabbari A. Toward a continuous wavelet transform-based search method for feature selection for classification of spectroscopic data[J]. *Chemometrics and Intelligent Laboratory Systems*, 2013, 127: 185-194.
- [18] 王淑贤, 肖航, 杨振发, 等. 香精掺假普洱茶的近红外光谱检测[J]. *激光与光电子学进展*, 2020, 57(23): 223005.
Wang S X, Xiao H, Yang Z F, et al. Detection of flavor adulterated Pu'er tea by near-infrared spectroscopy[J]. *Laser & Optoelectronics Progress*, 2020, 57(23): 223005.
- [19] Liu Y, Liu Y L, Chen Y Y, et al. The influence of spectral pretreatment on the selection of representative calibration samples for soil organic matter estimation using vis-NIR reflectance spectroscopy[J]. *Remote Sensing*, 2019, 11(4): 450-466.
- [20] 何欣龙, 王继芬. 牛顿插值多项式-导数光谱无损检测车用保险杠[J]. *激光技术*, 2020, 44(3): 333-337.
He X L, Wang J F. The identification about the automotive bumper based on Newton interpolation polynomial-infrared derivative spectroscopy[J]. *Laser Technology*, 2020, 44(3): 333-337.
- [21] 何欣龙, 陈利波, 王继芬, 等. 基于K近邻算法的塑钢窗拉曼光谱分析[J]. *激光与光电子学进展*, 2018, 55(5): 053001.
He X L, Chen L B, Wang J F, et al. Raman spectroscopy analysis of plastic steel window based on K nearest neighbors algorithm[J]. *Laser & Optoelectronics Progress*, 2018, 55(5): 053001.
- [22] 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.
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.
- [23] Rouchi M B, Khorrami M K, Garmarudi A B, et al. Application of infrared spectroscopy as Process Analytics Technology (PAT) approach in biodiesel production process utilizing Multivariate Curve Resolution Alternative Least Square (MCR-ALS)[J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2019, 213: 347-353.
- [24] Zhang F, Tang X J, Tong A X, et al. Baseline correction for infrared spectra using adaptive smoothness parameter penalized least squares method[J]. *Spectroscopy Letters*, 2020, 53(3): 222-233.
- [25] 沙云菲, 黄雯, 王亮, 等. 中红外和近红外数据融合的香型风格判别[J]. *光谱学与光谱分析*, 2021, 41(2): 473-477.
Sha Y F, Huang W, Wang L, et al. Merging MIR and NIR spectral data for flavor style determination [J]. *Spectroscopy and Spectral Analysis*, 2021, 41(2): 473-477.
- [26] 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.
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.
- [27] 李斌, 张敏, 周恒, 等. 基于小波包分析和支持向量机的光时域反射仪光缆故障识别[J]. *激光与光电子学进展*, 2019, 56(2): 021205.
Li B, Zhang M, Zhou H, et al. Identifying optical cable faults in OTDR based on wavelet packet analysis and support vector machine[J]. *Laser & Optoelectronics Progress*, 2019, 56(2): 021205.
- [28] 王利花, 杨智凯, 杨洋, 等. 基于聚类分析和贝叶斯判别的储层分类评价方法[J]. *石油化工应用*, 2019, 38(9): 67-70, 74.
Wang L H, Yang Z K, Yang Y, et al. Reservoir classification and evaluation method based on cluster analysis and Bayesian discrimination[J]. *Petrochemical Industry Application*, 2019, 38(9): 67-70, 74.
- [29] 杜倩, 万茶艳, 程存归. 薏苡仁及其伪品衰减全反射傅里叶变换红外光谱及傅里叶自去卷积的鉴别[J]. *医药导报*, 2015, 34(1): 104-107.
Du Q, Wan C Y, Cheng C G. Identification of coix seed and its counterfeit by attenuated total reflection Fourier transform infrared spectroscopy and Fourier self-deconvolution[J]. *Herald of Medicine*, 2015, 34(1): 104-107.