

差分拉曼光谱结合化学计量学对口红的检验

张进¹, 姜红^{1*}, 段斌², 刘峰²

¹中国人民公安大学侦查学院, 北京 100038;

²南京简智仪器设备有限公司, 江苏 南京 210049

摘要 在法庭科学中, 口红被认为是重要的物证之一。为了建立一种新的检验口红的方法, 利用差分拉曼光谱结合化学计量学方法对 50 个口红样本进行了检验。从差分拉曼光谱中可以发现, 口红包含各种油、蜡和氧化物成分。通过解析谱图, 可以根据一个或多个拉曼特征峰, 区分不同品牌、同一品牌不同系列的口红样本, 区分能力 (DP) 值为 100%。利用聚类分析的方法将所有样本合理分成了 10 类, DP 值为 88.16%。利用距离判别法建立的基于聚类分析的分类模型, 经交叉验证后发现分类准确率可达 96.00%。结果表明, 所提方法能够实现无损检测, 且收集的谱图信噪比高, 不受荧光干扰。

关键词 光谱学; 口红; 差分拉曼光谱; 区分; 化学计量学; 法庭科学

中图分类号 DF795.2

文献标志码 A

doi: 10.3788/LOP202158.2430002

Differential Raman Spectroscopy Combined with Chemometrics for Lipstick Test

Zhang Jin¹, Jiang Hong^{1*}, Duan Bin², Liu Feng²

¹Criminal Investigation School, People's Public Security University of China, Beijing 100038, China;

²Nanjing Jianzhi Instrument and Equipment Co., Ltd., Nanjing, Jiangsu 210049, China

Abstract Lipstick is considered to be one of the most important physical evidence in forensic science. In order to establish a new method for detecting lipsticks, 50 lipstick samples are tested by using differential Raman spectroscopy combined with chemometrics method. It can be observed from the differential Raman spectrum that the lipstick contains various oils, waxes and oxides. By analyzing the spectrum, lipstick samples of different brands and different series of the same brand can be distinguished according to one or more Raman characteristic peaks, and the discrimination power (DP) value is 100%. All samples are reasonably divided into 10 categories by cluster analysis method, and the DP value is 88.16%. The classification model based on cluster analysis is established by distance discriminant method, the classification accuracy rate can reach 96.00% after cross-validation. The results show that the proposed method can realize non-destructive inspection, and the collected spectra have high signal-to-noise ratio and are not disturbed by fluorescence.

Key words spectroscopy; lipstick; differential Raman spectroscopy; differentiation; chemometrics; forensic science

OCIS codes 300.6170; 300.6450; 300.6330

1 引言

口红是犯罪现场重要的物证之一, 它可以作为

各种侵害妇女罪中的重要证据, 也可以作为犯罪时的书写工具^[1]。大多数的口红具有相似的成分, 主要由油(质量分数通常为 40%~70%)、蜡(质量分

收稿日期: 2020-12-01; 修回日期: 2020-12-30; 录用日期: 2021-01-06

基金项目: 拔尖创新人才培养经费支持硕士研究生项目(2020ssky023)、南京简智仪器设备有限公司技术合作项目(20191218)

通信作者: *jiangh2001@163.com

数约为 15%) 和着色剂(质量分数约为 8%) 组成, 并含有少量的抗氧化剂和香水类物质。目前, 主要利用红外光谱法^[2-6]、气相色谱-串联质谱法^[7] 和拉曼光谱法^[8-9] 等分析方法^[10-11] 来检验口红。其中, 拉曼光谱因其无需样本制备和无损检材的特点被广泛使用。然而, 在之前的研究中, 拉曼光谱会受到荧光的影响。Salahioglu 等^[12] 在激发波长为 632.8 nm 的条件下, 利用显微共聚焦拉曼光谱采集了 69 种不同颜色的口红的光谱, 其中超过 10% (口红数量占比) 的样品因为荧光干扰太强而完全识别不到特征峰。Gardner 等^[13] 在 532 nm 和 780 nm 的激发波长下, 用拉曼光谱对 80 个口红进行了检验研究。研究发现, 在 532 nm 激发波长下, 样本检验受到了荧光干扰的严重影响。在 780 nm 激发波长条件下, 可以区分出大部分口红样本, 但仍有荧光干扰。差分拉曼光谱基于 2 种波长有略微偏移的激发光, 通过物理和数学相结合的差分拉曼技术可以有效地消除荧光背景, 得到物质的纯净谱图。因此, 本研究的目的在于建立一种新的检验口红的方法, 克服以往其他工作中报道的口红荧光问题。结合化学计量学方法, 对光谱数据进行更深层次地挖掘, 借助区分能力 (DP, M_{DP}) 值评价了区分效果, 为今后法庭科学领域中该类物证的研究提供了一定的参考。

2 实验部分

2.1 谱图采集及预处理

实验收集了 50 个不同品牌、不同系列的口红样本, 使用南京简智仪器有限公司生产的便携式差分拉曼光谱仪采集样本的光谱信号, 选用有略微偏移的中心波长为 784 nm 和 785 nm 的激发光, 测试波数范围为 280~2600 cm^{-1} , 激光功率设置为 220 mW, 积分时间设置为 40 s。

对两种激发光在同一条件下采集的拉曼谱图进行处理, 采用差分拉曼技术和反向传播 (BP) 神经网络算法将差分信号从基线偏离中分离出来。同时, 利用差分解调和去噪算法进行归一化、重建以及基线矫正, 将光谱的基线对齐, 并将对齐后光谱相减以得到纯净的拉曼光谱, 从而达到抑制荧光的目的^[14]。图 1 为 24 号样本在中心波长为 784 nm 和 785 nm 激发光下的拉曼谱图和处理后的差分拉曼谱图。可以发现, 在中心波长为 784 nm 和 785 nm 激发光下的光谱信噪比较低, 且特征峰不明显。在经过处理后, 可以清晰地识别出多个特征峰, 进而成功地提取了纯净的拉曼光谱。

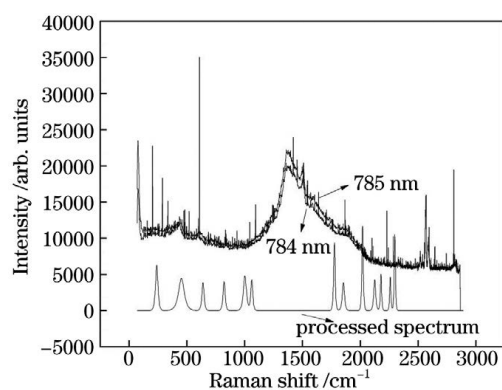


图 1 24 号样本在中心波长为 784 nm 和 785 nm 激发光下的拉曼谱图和处理后的差分拉曼谱图

Fig. 1 Raman spectra and differential Raman spectra after processing of sample 24 under excitation light with central wavelengths of 784 nm and 785 nm

2.2 理论基础

2.2.1 区分能力值

DP 最早由 Smalldon 等^[15] 提出, DP 值越大, 表明方法对样本的区分效果越好。计算公式为

$$M_{DP} = \frac{A}{B} \times 100\% = \frac{2A}{X(X-1)} \times 100\%, \quad (1)$$

式中: A 是可区分的样本对数; B 是样本总对数; X 是样本总数。

2.2.2 系统聚类分析

系统聚类分析^[16-18] 是一种无监督地对研究分类问题进行统计分析的方法。在该方法中, 首先将每个样本单独看成一类, 通过计算类间距离矩阵, 将距离最近的两类合并成一个新的类别。然后, 计算新类和其他类的距离, 并重复合并。随着类间距离的增加, 聚类数目不断减少, 在所有样本凝聚成一类后, 停止计算。

2.2.3 判别分析

判别分析^[19] 是一种判别个体所属类别的统计分析方法。在分类明确的情况下, 其根据研究对象的观测指标和所属类别建立判别准则, 以判别未知样本的归属。判别分析有多种判别方法, 包括 Fisher 判别法、Bayes 判别法、距离判别法等。其中, 距离判别法是根据已知的分类数据, 计算不同类别的重心。新类别距离哪一类样本的重心距离更近, 则将其归属到该类别中, 该方法对样本的分布没有特殊要求。

3 分析与讨论

3.1 典型的单幅拉曼谱图

选择具有代表性的 7 号样本的差分拉曼谱图进

行分析,如图 2 所示。可以发现,该谱图主要由两个区域组成,第一个区域的波数范围为 $280\sim 750\text{ cm}^{-1}$,其中 294 cm^{-1} 处的峰是由氧化铁(赤铁矿)引起的^[20], $395, 513, 636\text{ cm}^{-1}$ 处的峰是由二氧化钛(锐钛矿)引起的。第二个区域的波数范围为 $750\sim 1700\text{ cm}^{-1}$,其中大部分的峰是由油和蜡引起的,如 $1060, 1131, 1296\text{ cm}^{-1}$ 处的峰与蜂蜡有关; $1269\text{ cm}^{-1}, 1440\text{ cm}^{-1}, 1656\text{ cm}^{-1}$ 的峰是由蓖麻油引起的。此外,其他占比较少的成分也可能会造成特征峰的出现,如香水、防腐剂和润肤剂^[21]。口红是一种成分较为复杂的混合物,不同品牌、系列的口红样本的成分和比例都有可能不同,这为解析图谱可实现样本区分提供了依据。

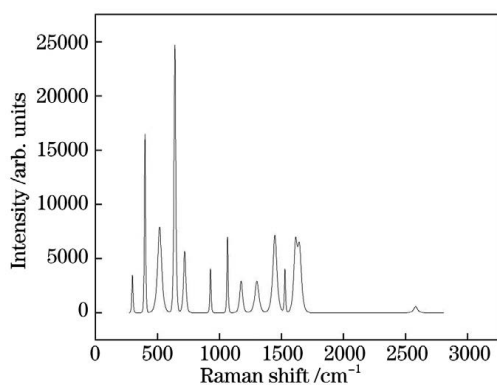


图 2 7 号样本的差分拉曼谱图

Fig. 2 Differential Raman spectrum of sample 7

3.2 谱图区分

首先,对不同品牌的样本的拉曼谱图进行了研究,以 13 号样本(2B197,美宝莲)、17 号样本(RUBY WOO, MAC)和 40 号样本(687, Dior)为例,三者的拉曼谱图(L1、L2、L3)如图 3 所示,其中圆圈的作用是标注用于比较的特征峰,箭头的作用是标注两个特征峰之间的联系。可以看出,虽然三

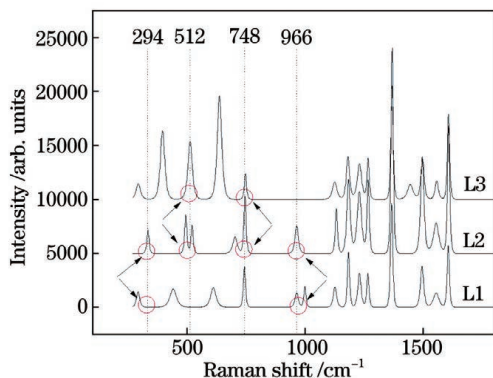


图 3 三种不同品牌样本的拉曼光谱

Fig. 3 Raman spectra of samples from three different brands

个样本在 $1000\sim 1700\text{ cm}^{-1}$ 波数范围内的峰位和峰数基本一致,但仍可以利用 $280\sim 1000\text{ cm}^{-1}$ 波数范围内的特征峰进行区分。例如,在 17 号样本和 40 号样本的谱图中,17 号样本在波数为 512 cm^{-1} 和 748 cm^{-1} 处的特征峰是双峰,而 40 号样本在相应位置处的特征峰是单峰。

然后,对同一品牌、不同系列的样本的拉曼谱图进行了研究,以 4 号样本(BE310,欧珀莱)、5 号样本(RD610,欧珀莱)和 11 号样本(PK700,欧珀莱)为例,三者的拉曼谱图(L4、L5、L6)如图 4 所示,其中圆圈的作用是标注用于比较的特征峰,箭头的作用是标注两个特征峰之间的联系。可以发现,4 号样本和 5 号样本在波数为 730 cm^{-1} 附近虽然都有一个双峰,但两个峰的相对峰高比差距很大。出现该现象的原因可能是两个样本含有某个相同的成分,但配比不同。因此,可以将两个样本区分开。5 号样本和 11 号样本在波数为 $750\sim 1500\text{ cm}^{-1}$ 的区域内存在多个不同的特征峰。通过对所有的样本谱图进行分析可以发现,并不存在某种单个峰或多个峰的组合是品牌、系列所特有的,并且制造商和产品配方也在不断变化,因此创建口红样本数据库非常困难。然而,所有样本均可因一个或多个峰的不同,与其他被检验的样本区分开来,即 DP 值为 100%。虽然通过解析谱图可以将所有样本区分开,但是这种方法费时费力,且不可避免人工误差的影响^[22],故引入化学计量学方法对口红样本进行多元区分。

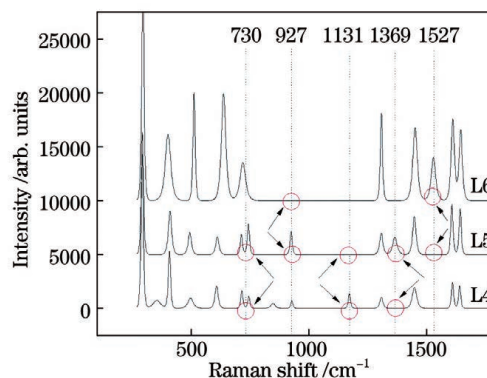


图 4 三种同一品牌、不同系列样本的拉曼谱图

Fig. 4 Raman spectra of three samples from same brand and different series

3.3 多元区分

利用系统聚类分析方法对样本的光谱数据进行处理,选择的聚类方法为 Ward 法,并计算数据点间的欧氏距离的平方(d),得到样本的聚类树状图,如图 5 所示。可以看出,当距离最小时,样本被划分为

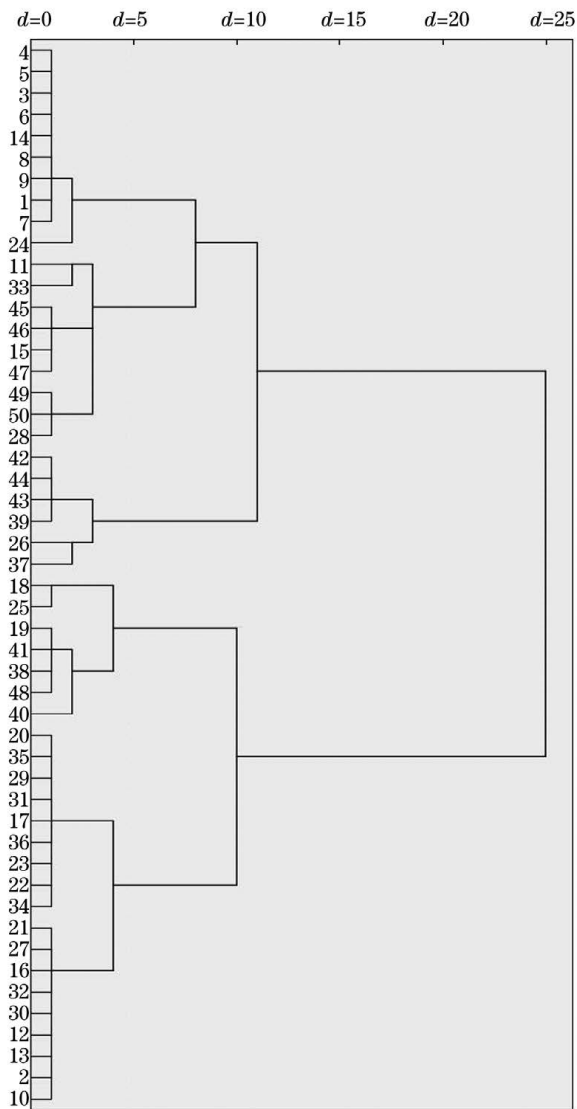


图 5 聚类树状图

Fig. 5 Clustering dendrogram

14 类。当距离为 5 时,样本被分为 5 类。同时,随着距离的增大,样本被划分的类数逐渐减少,当距离大于 25 时,样本停止聚类,即所有样本被分为一类。

考察聚类效果的方法有很多^[23]。集中计划是输出聚类的过程表,包括每一步被合并的类或者观测量、类或者观测量之间的距离和新生成类的信息等。利用集中计划中的聚类数和系数可以绘制一张碎石图,以确定合理的聚类数目,如图 6 所示。可以看出,当聚类数目小于 40 时曲线上升缓慢,当聚类数目大于 40 时曲线骤升。此时,上升缓慢的曲线部分和骤升的曲线部分构成了一个“肘部”。因此,可以确定合理的聚类数目为 10,结果见表 1。可以发现,不同类别中的样本可以被区分,而同一类中的样本不能被区分,以第 4 类为例,不能被区分的样本对对数为 3,分别是 28 号-49 号,28 号-50 号和 49 号-50 号。统计各个类别中不能被区分的样本总数,最终可计算得到聚类结果的 DP 值为 88.16%。

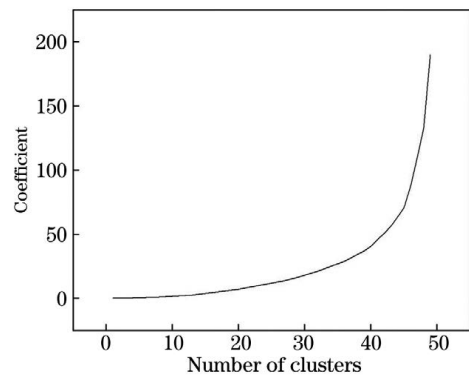


图 6 利用集中计划绘制的碎石图

Fig. 6 Gravel map drawn by agglomeration schedule

表 1 系统聚类结果

Table 1 Results of system clustering

Category	Sample number	Category	Sample number
1	1, 3, 4, 5, 6, 7, 8, 9, 14, 24	6	26, 37
2	11, 33	7	18, 25
3	15, 45, 46, 47	8	19, 38, 40, 41, 48
4	28, 49, 50	9	17, 20, 22, 23, 29, 31, 34, 35, 36
5	39, 42, 43, 44	10	2, 10, 12, 13, 16, 21, 27, 30, 32

系统聚类分析可以将样本分为不同的类别,但不能为确定未知样本的归属提供帮助。因此,利用判别分析法对样本进行分析,探究并建立了未知样本的预测模型。利用距离判别法判别并分析提取了 9 个特征函数,表 2 列出了前 3 个特征函数的函数

摘要,其中典范相关系数代表了特征函数与各类样本的相关程度;Wilks' lambda 值越小,判别函数对模型的影响越显著;方差百分比代表了函数对样本分类的解释能力,前 2 个函数的累计方差贡献率达 80.10%;M 表示函数检验的显著性,其值小于

表 2 前三个特征函数的函数摘要

Table 2 Function summaries of the first three characteristic functions

Function	Eigenvalue	Variance / %	Cumulative variance / %	Canonical correlation coefficient	Wilks' lambda	M
1	343.644	46.6	46.6	0.999	0.000	0.000
2	246.784	33.5	80.1	0.999	0.000	0.000
3	59.702	8.1	88.2	0.992	0.000	0.000

0.05 代表判别函数是成立的。因此,可以选择前 2 个函数来绘制样本的联合分布图,如图 7 所示,其中正方形表示每一类样本的重心,圆形表示单个样本。图 7 所示的联合分布图直观展示了样本的分类情况,可以发现,10 类样本的重心均可被有效地区分开。其中,第 2,4,5,6,7 类样本区分明显,第 1,3,8,9,10 类样本较为聚集。针对未知样本,将样本的光谱数据放入判别模型,考察样本的归属情况即可。利用留一法对判别分类结果进行交叉验证,最终可实现准确率为 96.00% 的分类,证明了聚类结果的合理性和判别模型的准确可靠性。

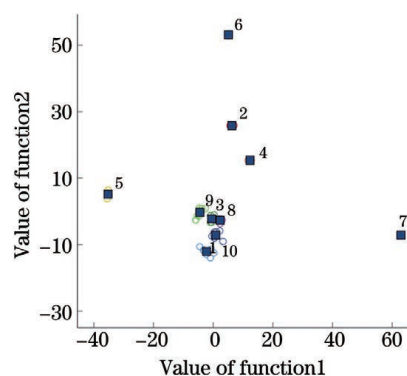


图 7 样本的联合分布图

Fig. 7 Joint distribution diagram of samples

4 结 论

基于差分拉曼光谱技术和化学计量学方法,提出了一种新的检验口红样本的方法。采集的谱图表明该方法可以有效克服口红样本强荧光的问题。通过对不同品牌、同一品牌不同系列的样本谱图的解析发现,所有样本均可因一个或多个峰的不同而被区分开,即 DP 值为 100.00%。聚类分析将样本分为了 10 类,不同类的样本可得到区分,计算出的 DP 值为 88.16%。利用距离判别法建立的基于聚类分析结果的判别模型,经交叉验证后准确率达 96.00%,实验结果较为理想,为法庭科学领域该类物证的区分检验提供了一定的参考。未来将进一步扩充样本容量,采用不同的算法对样本数据进行深

度挖掘。

参 考 文 献

- [1] Sharma S, Chophi R, Kumar R, et al. Differentiation of locally manufactured *Kajal* by Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy supported by chemometric analysis [J]. *Forensic Science International*, 2019, 303: 109930.
- [2] Gładysz M, Król M, Kościelniak P. Differentiation of red lipsticks using the attenuated total reflection technique supported by two chemometric methods [J]. *Forensic Science International*, 2017, 280: 130-138.
- [3] Chophi R, Sharma S, Singh R. Forensic analysis of red lipsticks using ATR-FTIR spectroscopy and chemometrics [J]. *Forensic Chemistry*, 2020, 17: 100209.
- [4] Wong J X W, Sauzier G, Lewis S W. Forensic discrimination of lipsticks using visible and attenuated total reflectance infrared spectroscopy [J]. *Forensic Science International*, 2019, 298: 88-96.
- [5] Sharma V, Bharti A, Kumar R. On the spectroscopic investigation of lipstick stains: forensic trace evidence [J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2019, 215: 48-57.
- [6] Pasieczna-Patkowska S, Olejnik T. Analysis of cosmetic products using different IR spectroscopy techniques [J]. *Annales UMCS Chemia*, 2014, 68 (1/2): 96-106.
- [7] Ma Q, Xi H W, Wang C, et al. Determination of ten volatile nitrosamines in cosmetics by gas chromatography tandem mass spectrometry [J]. *Chinese Journal of Analytical Chemistry*, 2011, 39 (8): 1201-1207.
- [8] Soares A R, Nascentes C C. Development of a simple method for the determination of lead in lipstick using alkaline solubilization and graphite furnace atomic absorption spectrometry [J]. *Talanta*, 2013, 105: 272-277.
- [9] Misra G, Mittal V K. Neutron activation analysis of

- lipsticks using γ -ray spectrometry [J]. *Journal of Applied Spectroscopy*, 2004, 71(2): 270-274.
- [10] López-López M, Özbek N, García-Ruiz C. Confocal Raman spectroscopy to trace lipstick with their smudges on different surfaces [J]. *Talanta*, 2014, 123: 135-139.
- [11] Salahioglu F, Went M J, Gibson S J. Application of Raman spectroscopy for the differentiation of lipstick traces [J]. *Analytical Methods*, 2013, 5(20): 5392-5401.
- [12] Salahioglu F, Went M J. Differentiation of lipsticks by Raman spectroscopy [J]. *Forensic Science International*, 2012, 223(1/2/3): 148-152.
- [13] Gardner P, Bertino M F, Weimer R, et al. Analysis of lipsticks using Raman spectroscopy [J]. *Forensic Science International*, 2013, 232(1/2/3): 67-72.
- [14] Fang G, Yin L, Liu F, et al. Application research of fluorescence suppression based on differential Raman technique [J]. *Laser Technology*, 2019, 43(3): 359-362.
方刚, 殷磊, 刘峰, 等. 基于差分喇曼技术在抑制荧光中的应用研究 [J]. *激光技术*, 2019, 43(3): 359-362.
- [15] Smalldon K W, Moffat A C. The calculation of discriminating power for a series of correlated attributes [J]. *Journal of the Forensic Science Society*, 1973, 13(4): 291-295.
- [16] Kumar R, Sharma V. Chemometrics in forensic science [J]. *Trends in Analytical Chemistry*, 2018, 105: 191-201.
- [17] Chen Z K, Guo R, Cheng P F. Application of LIF technology-based spectral feature extraction in oil detection [J]. *Laser & Optoelectronics Progress*, 2020, 57(13): 133002.
陈至坤, 郭蕊, 程朋飞. 基于 LIF 技术的光谱特征提取在油类检测中的应用 [J]. *激光与光电子学进展*, 2020, 57(13): 133002.
- [18] Zhu M, Zhou Z F, Jiang Y, et al. An accurate recognition method of pitaya plants based on visible light band UAV remote sensing [J]. *Laser & Optoelectronics Progress*, 2020, 57(14): 142801.
朱孟, 周忠发, 蒋翼, 等. 基于可见光波段无人机遥感的火龙果精准识别方法 [J]. *激光与光电子学进展*, 2020, 57(14): 142801.
- [19] He Y, Wang J F. Rapid nondestructive identification of wood lacquer using Raman spectroscopy based on characteristic-band-Fisher-K nearest neighbor [J]. *Laser & Optoelectronics Progress*, 2020, 57(1): 013001.
何亚, 王继芬. 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 [J]. *激光与光电子学进展*, 2020, 57(1): 013001.
- [20] Hanesch M. Raman spectroscopy of iron oxides and (oxy) hydroxides at low laser power and possible applications in environmental magnetic studies [J]. *Geophysical Journal International*, 2009, 177(3): 941-948.
- [21] Chopri R, Sharma S, Sharma S C, et al. Trends in the forensic analysis of cosmetic evidence [J]. *Forensic Chemistry*, 2019, 14: 100165.
- [22] 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.
- [23] Gautam R, Vanga S, Ariese F, et al. Review of multidimensional data processing approaches for Raman and infrared spectroscopy [J]. *EPJ Techniques and Instrumentation*, 2015, 2: 8.