

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

李 夏^{1,2} 黄梅珍^{1,2} 宋 彪¹ 邹 烨¹ 陈 婕¹ 汪 洋³

¹上海交通大学电子信息与电气工程学院仪器系, 上海 200240

²上海智能诊疗仪器工程技术研究中心, 上海 200240

³上海交通大学物理与天文系, 上海 200240

摘要 利用自行研制的便携式拉曼光谱仪(Hx-Spec)对高山石类寿山石(13组)和老挝石(22组)进行了分类判别分析。Hx-Spec拉曼光谱仪的激发波长785 nm,光谱范围200~2700 cm⁻¹,分辨率6 cm⁻¹。结合主成分分析(PCA)和线性判别分析(LDA)建立多元统计算法模型,将老挝石从高山石类寿山石中区别出的特异性和灵敏度分别为100%与86.7%,受试样品工作特征(ROC)曲线下面积为0.977。另外,加入8个新样品作为验证,判别正确率为87.5%。研究表明,自行研制的便携式拉曼光谱仪结合PCA-LDA统计分析方法有望用于区分高山石类寿山石和老挝石。

关键词 光谱学; 拉曼光谱; 无损鉴别; 主成分分析; 寿山石; 老挝石

中图分类号 P575.4 **文献标识码** A

doi: 10.3788/AOS201636.1230001

Nondestructive Discrimination of Shoushan Stone and Laos Stone Based on Raman Spectroscopy and Multivariate Methods

Li Xia^{1,2} Huang Meizhen^{1,2} Song Biao¹ Zou Ye¹ Chen Jie¹ Wang Yang³

¹Department of Instrument Science and Engineering, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

²Shanghai Research Center of Engineering and Technology for Intelligent Diagnosis and Treatment Instrumentation, Shanghai 200240, China

³Department of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China

Abstract The Raman spectra of 35 stone samples (13 Shoushan stones and 22 Laos stones) are measured by a self-developed portable Raman spectrometer Hx-Spec, which is excited by a 785 nm diode laser and the spectrum range is 200-2700 cm⁻¹ with a resolution of 6 cm⁻¹. A multivariate model is established by combining the principal component analysis (PCA) and the linear discriminant analysis (LDA), a discrimination sensitivity of 86.7% and a specificity of 100% are achieved in separating Laos stone from Shoushan stone, while the performance of corresponding receiver operating characteristic (ROC) curve is 0.977. Besides, a discriminant accuracy achieves 87.5% after 8 new samples are added as a validation. These results demonstrate that the utility of portable Raman spectrometer combined with chemometrics provides a new strategy to discriminate the Shoushan and Laos stones.

Key words spectroscopy; Raman spectroscopy; nondestructive identification; principal component analysis; Shoushan stone; Laos stone

OCIS codes 300.6360; 300.6450; 300.6490

收稿日期: 2016-07-04; **收到修改稿日期:** 2016-08-09

基金项目: 国家自然科学基金(31178083)、国家重大仪器设备开发专项(2012YQ180132)

作者简介: 李 夏(1991—),女,硕士研究生,主要从事光检测和仪器技术方面的研究。E-mail: lixia00@sjtu.edu.cn

导师简介: 黄梅珍(1966—),女,博士,研究员,博士生导师,主要从事光检测和仪器技术、生物医学光子技术等方面的研究。E-mail: mzhuang@sjtu.edu.cn(通信联系人)。

1 引 言

寿山石,作为中国传统“四大印章石”之首,在 2002 年正式被中国宝玉石协会命名为“中国国石”。寿山石主要分布在福州市晋安区宦溪镇、寿山乡、日溪乡及连江县小沧畲族乡、蓼沿乡等现辖行政区域。根据福建省技术监督局 2012 年颁布的福建省地方标准——《地理标志产品 寿山石》(DB35/T 419—2012),寿山石按其矿物组成、理化特征和产出环境不同分为田黄石类、高山石类、芙蓉石类和汶洋石类四大类。其中田黄石类主要矿物为地开石,少量为珍珠陶石和伊利石,高山石类主要矿物由高岭石族组成,如地开石、高岭石、珍珠陶石等,芙蓉石类主要矿物为叶腊石,汶洋石类主要矿物为伊利石^[1]。近两年来,和寿山石高度相似的大量产于老挝的老挝石涌入各大寿山石原石和成品市场,给寿山石市场造成较大的冲击。如何准确快速无损鉴别价格昂贵的寿山石与外来的老挝石目前仍然是一个迫切需要解决的难题。

目前对寿山石进行分类和鉴定主要采用 X 射线衍射(XRD)、红外光谱等分析手段^[2-3],文献[4-5]采用 X 射线粉末衍射、中红外光谱(MIR)和近红外光谱(NIR)等方法对老挝石矿物组成进行了研究。研究表明,老挝石主要由 Al 及 Si 组成,其含量与高岭石族矿物的理论化学组成较为一致^[6],这与高山石类寿山石高度相似。现有的主流分析鉴别手段为 XRD,但需对被测样品进行研磨至粉品的表面破坏,难以从外观及常规宝石学性质方面进行名石鉴别。

拉曼光谱检测样品具有无损,无需对样品进行前期制样,可以同时测定多个样品的优点,已被应用于翡翠、透闪石,红宝石,鸡血石,寿山石^[7-11]等宝玉石的鉴别方面。不同的珠宝玉石由于分子组成不同、结构不同,其拉曼光谱的特征峰也不同,据此可以对测试样品的组成成分和晶体结构进行表征,获取有用信息,以达到测定、鉴别宝玉石的目的。对于分子信息相近的两类样品,其拉曼光谱在光谱形状,更多地在谱峰相对强度上会有所不同,利用主成分分析(PCA)方法提取出具有显著性差异的变量可以更加直观和准确地作出判别。

鉴于拉曼光谱结合主成分分析方法在医学^[12-14]、食品^[15-17]等领域的应用,本文初步研究了其在寿山石和老挝石区分中的可行性。首先,采用自行研制的便携式拉曼光谱仪对高山石类寿山石和老挝石进行了测试,并对其拉曼峰进行分析和归属。然后,利用主成分分析结合线性判别(LDA)方法初步区分高山石类寿山石和老挝石,由此建立基于 PCA-LDA 的老挝石与寿山石的鉴别模型,并采用已知种类的高山石类寿山石和老挝石对该模型进行验证。初步研究结果表明,便携式拉曼光谱仪结合多变量统计分析方法能够对高山石类寿山石和老挝石进行快速、无损、现场化的初步区分。

2 测试仪器与方法

2.1 测试仪器

自行研制的 Hx-Spec 便携式拉曼光谱仪^[18-19],包括激发光源、光纤探头和光谱仪三个主要部分。其中,激发光源采用半导体激光器,中心波长为 785 nm;光纤拉曼探头激发端将激光传导、纯化并会聚至样品,接收端收集样品的拉曼散射信号,经过滤光后传导到光谱仪;光谱仪采用 Czerny-Turner (C-T)光路结构,经过消彗差和消像散设计及优化,光谱范围 200~2700 cm^{-1} ,分辨率 6 cm^{-1} 。

2.2 测试方法

使用自行研制的便携式拉曼光谱仪对寿山石和老挝石样品进行了拉曼信号的测试。寿山石样品主要有高山石、鸡母窝、马背石、旗降石和善伯洞石等 13 块,老挝石样品由民间收藏家提供 22 块。其中取高山石类寿山石 11 块,老挝石 16 块作为建模集样品,另取高山石类寿山石 2 块和老挝石 6 块作为验证集样品。拉曼光谱测量过程无损,样品可直接进行测试,不需要进行前期制样。测试条件为:激光波长 785 nm,激光输出功率为 300 mW,积分时间 10~20 s,对样品的同一测试点采集 3 次光谱数据,取其平均值作为最终光谱。为保证测试条件的一致性,将光纤探头固定在支架上并使用定焦套筒,同时测试过程全程在暗室中进行以避免杂散光对实验的影响。

3 数据分析和结果

3.1 拉曼光谱分析

图 1 反映出高山石类寿山石和老挝石的特征光谱信息在 $250\sim 1250\text{ cm}^{-1}$ 内,所以分析主要在此光谱范围。进行光谱分析之前,为了消除噪声和基线对实验分析的影响,提高实验分析的准确性,首先对原始光谱采用 5 阶多项式拟合去除背景,然后对光谱选用移动窗口最小二乘多项式平滑处理,处理后的光谱噪声有一定的消除^[20]。另外,由于光谱中的信息主要来自谱峰的相对强度和光谱形状,而其受激光功率变动、聚集情况等具体的实验条件影响,所以为了减小实验条件对实验结果的影响,需要先对样品信息主要所在的 $250\sim 1250\text{ cm}^{-1}$ 范围光谱进行面积归一化处理。图 2 上方谱线对比了 11 个高山石类寿山石面积归一化平均拉曼光谱(蓝色线)与 16 个老挝石面积归一化平均拉曼光谱(红色线),下方谱线的绿色线代表高山石类寿山石和老挝石的差分谱(寿山石减去老挝石)。从图 2 的平均拉曼谱图可见,主要拉曼特征峰在 $264, 329, 430, 456, 742, 789, 910\text{ cm}^{-1}$ 处,寿山石在 264 cm^{-1} 的谱峰强度比相对应的老挝石的谱峰强;而在 $329, 430, 456, 742, 789, 910\text{ cm}^{-1}$ 等处老挝石的谱峰强度比较大,相应的振动键的归属如表 1 所示。本文中的寿山石和老挝石均以地开石为基质,所以在谱峰位置形状上差异不明显,若直接以谱峰位置为判据并不能区分它们。但由于两者的产地不同,其矿物质元素的相对含量不同,故在归一化后的拉曼谱峰相对强度上存在差异。之后利用 PCA-LDA 的统计方法,对高山石类寿山石和老挝石在光谱上的差异进行进一步的分析与讨论。

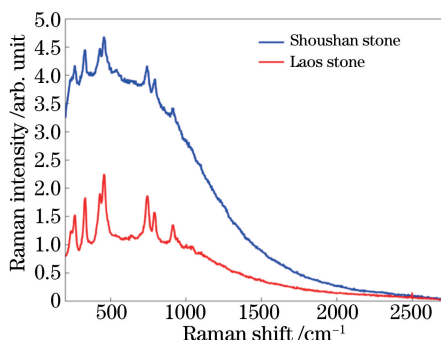


图 1 寿山石和老挝石的原始拉曼图谱

Fig. 1 Raman spectra of Shoushan stone and Laos stone

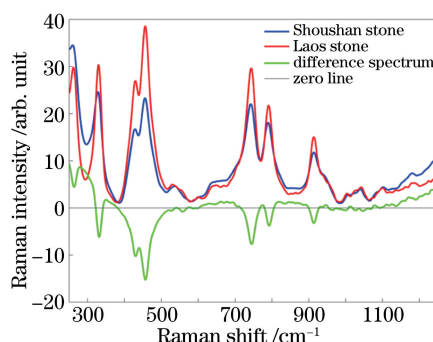


图 2 经平滑和归一化处理后的寿山石、老挝石平均拉曼光谱及差分谱

Fig. 2 Difference spectrum of the normalized mean spectra of Shoushan and Laos stone

表 1 拉曼位移与谱带归属^[11,21-23]

Table 1 Raman shift and their respective tentative assignments based on literature data^[11,21-23]

Peak position/ cm^{-1}	Suggested assignment
264	O—H—O triangle
329	Si—O bending vibration
430	Si—O bending vibration
456	Si—O bending vibration
742	Si—O—Al translation
789	Al—OH translation
910	OH deformation

3.2 判别分析

从拉曼光谱上不能直观地判断出高山石类寿山石与老挝石,需借助多变量统计方法来实现。利用 PCA 结合 LDA 进行统计分析。主成分分析是利用降维的思想,在保留原始变量尽可能多的信息的前提下把复杂的多维数据转化成几个能代表两组数据之间差异的变量的多元统计方法,在光谱分析中得到应用^[24-25]。LDA 根据同组方差最小和组间方差最大的原则划分组类。以代表整个光谱主要信息的第一主成分 PC1 和第二主成分 PC2 分别为 y 轴和 x 轴所得的分类结果如图 3 所示,其中第一和第二主成分包含变量信息的

92.84%。从图 3(a)可以看出,寿山石和老挝石能得到比较好的分类。经过 PCA 分析后,对高山类寿山石和老挝石两组数据的 PC1,PC2 做进一步 LDA。结果通过 LDA 建立的分割线将高山类寿山石从老挝石中区分出来的灵敏度为 86.7%,特异性为 100%。

为了鉴定 PCA-LDA 模型的准确度,利用受试样品的工作特征(ROC)曲线对所构建的算法进行评价。ROC 曲线反映了分析方法的灵敏度和特异性之间的关系,是判断算法准确性的代表方法。对 PCA-LDA 模型得到的判别函数得分进行 ROC 曲线分析,如图 3(b)所示,ROC 曲线下的积分区域为 0.977,表明实验所构建的 PCA-LDA 算法具有较高的可靠性。

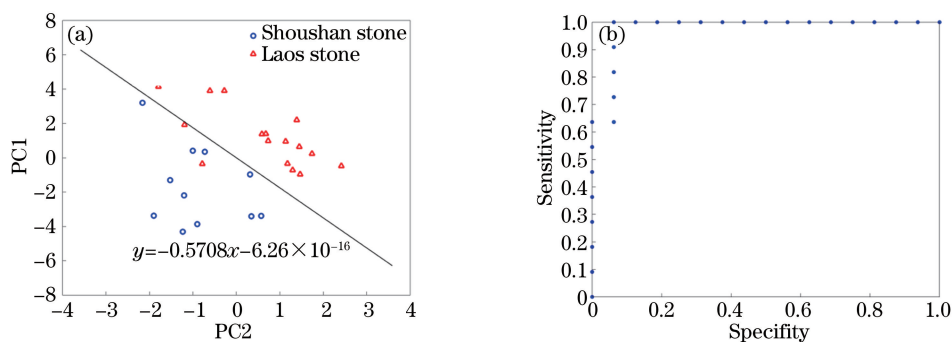


图 3 (a) 以 PC1,PC2 为横纵坐标的老挝石和寿山石散点图及分割线; (b) ROC 曲线

Fig. 3 (a) Dividing line as a diagnostic algorithm separating the two groups; (b) ROC curve

对上述 PCA-LDA 模型所用的高山类寿山石和老挝石的光谱数据做建模集,新加入 2 个高山石类寿山石和 6 个老挝石样品作为验证,采用与上文相同的方法对原始光谱进行背景噪声扣除、基线校正、面积归一化等处理后,对验证集进行验证,结果如图 4 所示。验证的正确判别率 87.5%。所得样品数量有限,但初步可判别高山石类寿山石和老挝石。

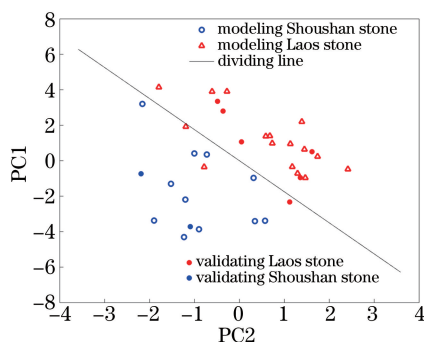


图 4 加入 8 个验证样品所得分类结果

Fig. 4 Classification result after adding eight validation samples

4 结 论

传统鉴别同一类别不同产地的宝石时多采用 X 射线衍射能谱分析法,这是一种有损的分析方法。本文利用拉曼光谱结合 PCA-LDA 统计分析建模,将高山石类寿山石从老挝石中区分出来的特异性与灵敏度分别为 100%和 86.7%,ROC 曲线积分面积为 0.977。实验结果表明,拉曼光谱结合 PCA-LDA 统计分析的方法可望用于对寿山石和老挝石做初步的快速无损的区分。

参 考 文 献

- Huang Zhiming, Tang Deping, Lin Hui, *et al.* Near infrared spectroscopy: A new tool to identify different types of larderite[J]. *Journal of Gems and Gemmology*, 2014, 16(5): 34-43.
黄志明, 汤德平, 林 辉, 等. 近红外光谱分析在寿山石鉴别中的应用[J]. *宝石和宝石学杂志*, 2014, 16(5): 34-43.
- Liu Weidong, Jiang Guojian, Xu Jiayue, *et al.* A study of the composition and classification of Shoushan stone by XRD

- method[J]. Journal of Shanghai Institute of Technology (Natural Science), 2009, 9(3): 175-178.
- 刘卫东, 江国健, 徐家跃, 等. 寿山石矿物组成及其分类的 XRD 研究[J]. 上海应用技术学院学报(自然科学版), 2009, 9(3): 175-178.
- 3 Wu Bin, Feng Ansheng. Identification method of X-ray diffractometry in Shoushan stone[J], Conservation and Utilization of Mineral Resources, 2010(6): 56-57.
- 吴 彬, 冯安生. 寿山石的 X 射线衍射鉴定法[J]. 矿产保护与利用, 2010(6): 56-57.
- 4 Tang Deping, Zheng Danwei, Huang Shanshan. Mineral components and identification of Laos stones [J]. Journal of Gems and Gemmology, 2015, 17(1): 1-9.
- 汤德平, 郑丹威, 黄珊珊. “老挝石”的矿物组成及鉴别研究[J]. 宝石和宝石学杂志, 2015, 17(1): 1-9.
- 5 Wang Han, Zhou Zhengyu, Zhong Qian, *et al.* Study on petrological and mineralogical characteristics of Laos stone by EPMA-XRD-SEM[J]. Rock and Mineral Analysis, 2016, 35(1): 56-61.
- 王 含, 周征宇, 钟 倩, 等. 电子探针-X 射线衍射-扫描电镜研究老挝石岩石矿物学特征[J]. 岩矿测试, 2016, 35(1): 56-61.
- 6 杨雅秀, 张乃娴. 中国粘土矿物[M]. 北京: 地质出版社, 1994: 21-22.
- 7 Zu Endong, Chen Dapeng, Zhang Pengxiang. Identification of B jade by Raman spectroscopy [J]. Spectroscopy and Spectral Analysis, 2003, 23(1): 64-66.
- 祖恩东, 陈大鹏, 张鹏翔. 翡翠 B 货的拉曼光谱鉴别[J]. 光谱学与光谱分析, 2003, 23(1): 64-66.
- 8 Yang Xiaodan, Shi Guanghai, Liu Yan. Vibrational spectra of black species of hetiannephrite (tremolite jade) and its color genesis[J]. Spectroscopy and Spectral Analysis, 2012, 32(3): 681-685.
- 杨晓丹, 施光海, 刘 琰. 新疆和田黑色透闪石质软玉振动光谱特征及颜色成因[J]. 光谱学与光谱分析, 2012, 32(3): 681-685.
- 9 Fan Jianliang, Guo Shouguo, Liu Xueliang, *et al.* Application of Raman spectroscopy in ruby identification[J]. Applied Laser, 2008, 28(2): 150-154.
- 范建良, 郭守国, 刘学良, 等. 拉曼光谱在红宝石检测中的应用研究[J]. 应用激光, 2008, 28(2): 150-154.
- 10 Han Xiaozhen, Guo Zhengye, Kang Yan, *et al.* Application of Raman spectroscopy in certification of chicken-blood stones [J]. Acta Optica Sinica, 2015, 35(1): 0130003.
- 韩孝朕, 郭正也, 康 燕, 等. 拉曼光谱在鸡血石鉴定中的应用[J]. 光学学报, 2015, 35(1): 0130003.
- 11 Sun Zhenhua, Huang Meizhen, Yu Zhengang, *et al.* Portable Raman spectrometer and its application in Shoushan stone detection[J]. Journal of Optoelectronics • Laser, 2015, 26(6): 1152-1156.
- 孙振华, 黄梅珍, 余镇岗, 等. 便携式拉曼光谱仪在寿山石检测中的应用[J]. 光电子 • 激光, 2015, 26(6): 1152-1156.
- 12 Feng Shangyuan, Pan Jianji, Wu Yan'an, *et al.* Study on gastric cancer blood plasma based on surface-enhanced Raman spectroscopy combined with multivariate analysis[J]. Scientia Sinica Vitae, 2011, 41(7): 550-557.
- 冯尚源, 潘建基, 伍严安, 等. 基于 SERS 技术结合多变量统计分析胃癌患者血浆拉曼光谱[J]. 中国科学生命科学, 2011, 41(7): 550-557.
- 13 Chen L, Wang Y, Liu Nenrong, *et al.* Near-infrared confocal micro-Raman spectroscopy combined with PCA-LDA multivariate analysis for detection of esophageal cancer[J]. Laser Physics, 2013, 23(6): 065601.
- 14 Guo Li, Zhang Yi, Lü Jinyan, *et al.* Multivariate statistical analysis of serum from breast cancer patients using surface enhanced Raman spectrum[J]. Spectroscopy and Spectral Analysis, 2013, 33(6): 1553-1556.
- 郭 丽, 张 毅, 吕金燕, 等. 用多元分析方法研究乳腺癌血清的表面增强拉曼光谱[J]. 光谱学与光谱分析, 2013, 33(6): 1553-1556.
- 15 Júnior P H R, Oliveira K D S, de Almeida C E, *et al.* FT-Raman and chemometric tools for rapid determination of quality parameters in milk powder: Classification of samples for the presence of lactose and fraud detection by addition of maltodextrin[J]. Food Chemistry, 2016, 196: 584-588.
- 16 Almeida M R, Oliveira K D S, Stephani R, *et al.* Fourier-transform Raman analysis of milk powder: A potential method for rapid quality screening[J]. Journal of Raman Spectroscopy, 2011, 42(7): 1548-1552.
- 17 Rubayiza A B, Meurens M. Chemical discrimination of arabica and robusta coffees by Fourier transform Raman spectroscopy[J]. Journal of Agricultural & Food Chemistry, 2005, 53(12): 4654-4659.
- 18 Sun Zhenhua, Yu Zhengang, Huang Meizhen, *et al.* Optimal design of a portable Raman spectrometer and the preliminary applications[J]. Journal of Optoelectronics • Laser, 2015, 26(6): 1132-1137.
- 孙振华, 余镇岗, 黄梅珍, 等. 小型化拉曼光谱仪的优化设计及应用[J]. 光电子 • 激光, 2015, 26(6): 1132-1137.
- 19 Yu Z G, Huang M Z, Wang Y, *et al.* Fast wavelength calibration method for spectrometers based on waveguide comb

- optical filter[J]. *Review of Scientific Instruments*, 2015, 86(4): 043103.
- 20 Mokrane A, Friant-Michel P, Cartier A, *et al.* Scaled semiempirical method for the calculation of vibrational spectra molecular vibrational frequencies of monosaccharides and disaccharides by PM3 method [J]. *Journal of Molecular Structure*, 1997, 395: 71-80.
- 21 Frost R L. Fourier transform Raman spectroscopy of kaolinite, dickite and halloysite[J]. *Clays and Clay Minerals*, 1995, 43(2): 191-195.
- 22 Frost R L, Tran T H, Rintoul L, *et al.* Raman microscopy of dickite, kaolinite and their intercalates[J]. *Analyst*, 1998, 123(4): 611-616.
- 23 Johnston C T, Helsen J, Schoonheydt R A, *et al.* Single-crystal Raman spectroscopic study of dickite[J]. *American Mineralogist*, 1998, 83(1): 75-84.
- 24 Haaland D M, Thomas E V. Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information[J]. *Analytical Chemistry*, 1988, 60(11): 1193-1202.
- 25 Holmes E, Nicholson J K, Nicholls A W, *et al.* The identification of novel biomarkers of renal toxicity using automatic data reduction techniques and PCA of proton NMR spectra of urine [J]. *Chemometrics and Intelligent Laboratory Systems*, 1998, 44(1): 245-255.