

粉尘环境下典型煤岩近红外光谱特征及识别方法

向阳, 王世博*, 葛世荣, 王赛亚, 周悦, 吕渊博, 杨恩

中国矿业大学机电工程学院, 江苏 徐州 221116

摘要 针对煤矿井下近红外煤岩识别中所存在的粉尘问题, 采用无烟煤与抑爆剂 9:1 混合的混合物模拟煤矿井下粉尘环境, 构建了粉尘环境煤岩光谱识别实验装置。为了研究粉尘环境对典型煤岩近红外光谱的影响, 从全国各地收集了页岩、砂岩、灰岩 3 类岩石样本及无烟煤、烟煤、褐煤 3 类煤类样本的原位典型煤岩试样 23 个, 采集无粉尘情况下的 23 个煤岩样本表面近红外波段 1000–2500nm 的反射光谱作为实验标准数据库, 分别从实验标准样本库中 3 类典型煤岩样本与 3 类典型岩石样本中随机选择 1 个样本作为实验样本, 分别采集测试样本在 600, 1 000, 1 500 和 3 000 $\text{mg} \cdot \text{m}^{-3}$ 粉尘浓度下的近红外波段的反射光谱数据, 结果显示: 粉尘的加入导致 1 000~1 200 nm 波段与 2 400~2 500 nm 波段的光谱图像信噪比降低; 随着粉尘浓度的增加, 粉尘中的无烟煤的不透明物质使得实验样本中的特征吸收谷减弱; 采用光谱角度匹配 SAM 以及皮尔逊相关系数对试样和标准样本库进行相关性分析, 无烟煤类样本、烟煤类样本、砂岩类样本、灰岩类样本在光谱角度匹配 SAM 匹配模型下有着较高的匹配度, 匹配度在各个粉尘浓度下均处于 0.9 以上; 相关系数匹配模型匹配度受粉尘的影响剧烈, 平均相关系数为 0.73; 实验标准数据库及实验样本经 SG 卷积和 SNV 标准正态预处理后, 预处理后的样本数据库与实验样本光谱角度匹配 SAM 匹配模型匹配度无明显变化, 相关系数匹配模型匹配度显著提升, 平均相关系数为 0.78; 除褐煤 2 号外, 所有的样本光谱相关系数平均提升 0.13, 无烟煤 2 号样本各个浓度平均相关系数提升 76.3%, 而样本 12 褐煤 2 号的光谱相关系数经光谱预处理降低。建立光谱角度匹配 SAM 以及皮尔逊相关系数煤岩识别模型, 二值化煤岩样本, 煤为“0”岩为“1”, 通过两种识别模型对不同浓度下的 6 个实验样本进行煤岩识别, 光谱角度匹配 SAM 的识别准确率 P 为 100%, 识别时间为 8 ms, 皮尔逊相关系数的识别准确率 P 为 87.5%, 识别时间为 852 ms。

关键词 煤岩识别; 典型煤岩; 近红外光谱; 粉尘环境

中图分类号: TD67 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2020)11-3430-08

引言

近红外光谱作为一种价格低廉、实时、信噪比高、非破坏性的光谱分析技术, 已经在煤质分析检测和煤岩分类^[1]等方面得到了应用, 其中, 聚类分析在近红外光谱识别分类中应用广泛, 国内外学者对煤岩聚类进行了许多的研究。杨恩^[2]等基于煤岩的灰分产率, 建立了支持向量煤岩灰分回归(SVR)以及向量煤岩分类(SVC)模型对煤岩进行远距离识别实验, 该模型具有较高的识别准确率。

上述文献报道以遥感的思路对煤岩光谱进行远距离的灰分产率预测, 未考虑到空气中的粉尘对光谱曲线的影响, 在

煤矿采掘现场中, 粉尘浓度较大, 刮板输送机上的煤流速度很快, 但目前还未见针对高粉尘, 高运载速度现场的近红外光谱研究, 导致识别的算法, 识别的模型大多不适应采掘现场。所以, 为了对井下实验提供参考, 以典型原位煤岩为实验样本, 在粉尘环境下分析样本近红外光谱特征以及基于标准样本库建立了两种煤岩快速识别模型。

1 实验部分

1.1 装置

图 1 为粉尘环境下近红外光谱采集装置。采用内部空间为 1 500 mm×500 mm×600 mm 密封箱形成密闭粉尘环境,

收稿日期: 2019-11-11, 修订日期: 2020-03-20

基金项目: 国家自然科学基金联合基金项目(U1610251, 51874279), 国家重点研发计划项目(2018YFC0604503), 江苏省高校优势学科建设工程项目(PAPD)资助

作者简介: 向阳, 1994 年生, 中国矿业大学机电工程学院硕士研究生 e-mail: xiangyang_cumt@qq.com

* 通讯联系人 e-mail: wangshb@cumt.edu.cn

根据设定的粉尘浓度投入定量粉尘颗粒。采用 SL-DC05 型无极变速鼓风机使粉尘处于分布均匀的悬浮状态。煤岩样品和光源分别位于密封箱两侧, 距离为 150 cm。为减小气溶胶消光系数对光源的影响, 使用惊鸿 2 000 W 透射式影视聚光灯作为光源, 用林上科技的 LS122A 红外辐射计测量光源辐射能量。使用荷兰 AvaSpec-NIR512 近红外光纤光谱仪进行数据采集, 设置光谱范围为 1 000~2 500 nm, 采集间隔为 3 nm; 使用直径为 3.5 cm 的准直镜为信号的采集装置, 准直

镜连接 Y 形光纤, 一端连接光谱仪, 另一端接入 650 nm 激光光源作为指示光源, 以便于能对准待测样本, 为避免粉尘对准直镜的污染, 在准直镜镜头处加装 JGS3 石英玻璃, 把准直镜固定于密封箱体上, 准直镜的镜头直径为 2.5 cm, 视场角为 0.1°, 准直镜与样本的探测距离保持 150 cm, 所以光谱仪在样本上的探测区域的直径为 2.76 cm, 面积大约为 6 cm²。在每组实验前用 15 cm×15 cm 材料为 PTFE 的标准白板在 150 cm 箱体对光源进行白校正。

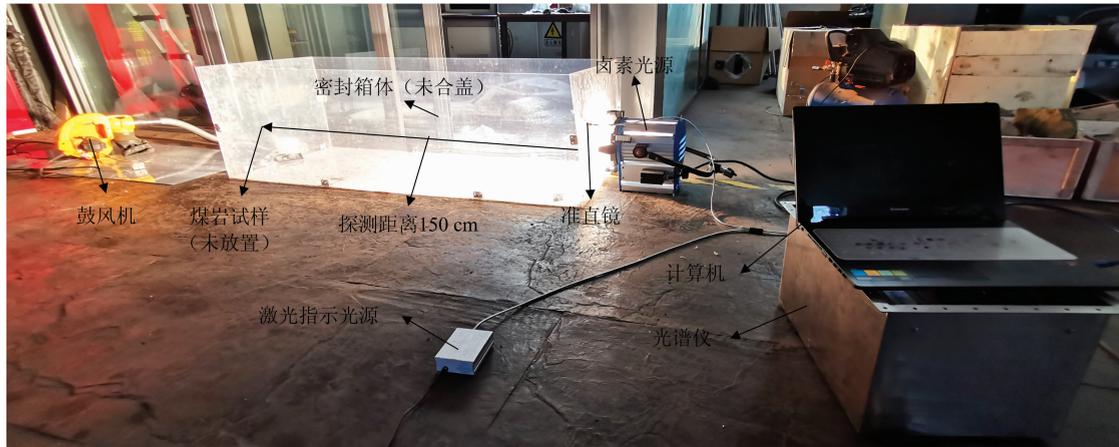


图 1 粉尘-近红外光谱实验装置

Fig. 1 Dust-near infrared spectroscopy experimental device

1.2 粉尘环境的设置

文献[3]通过对比尔定律变形, 得到了浓度 c (单位 $\text{mg} \cdot \text{m}^{-3}$)、吸收物质的厚度 L (单位为 m); 消光系数 k 、 I 和 I_0 分别为出射光光强和入射光光强 (单位为 $\text{mW} \cdot \text{cm}^{-2}$); 参量之间的关系式

$$I = I_0 e^{-kL} \quad (1)$$

同时还指出粉尘的消光系数与粉尘的密度、粉尘的中粒径成反比关系

$$k = \frac{\alpha}{\rho d_{50}} \quad (2)$$

式(2)中, α 为实验常数, 由实验获得; ρ 为粉尘密度; d_{50} 为粉尘中粒径。由上述两式可知在粉尘密度 ρ 、吸收物质的厚度 L 和中粒径 d_{50} 一定的情况下, 粉尘浓度越大, 试样所反射的光强 I_0 被粉尘吸收的越大, 为了保证准直镜所接收到的反射光强 I 只受到粉尘消光系数的单次影响, 在加入粉尘后, 用鼓风机吹拂粉尘 2 min, 使箱体内粉尘浓度均匀, 用红外辐射计从鼓风机风道处测试光源强度, 通过调整光源的发光强度, 保证测到的辐照度约为 $1\,200 \text{ W} \cdot \text{m}^{-2}$, 光照度约为 $20\,000 \text{ Lux}$ 。

根据文献[4]对综采工作面粉尘浓度的测定, 综采工作面的全尘的浓度和粒度为 $300 \sim 600 \text{ mg} \cdot \text{m}^{-3}$, 中粒径为 $22.23 \mu\text{m}$ 。在没有降尘措施的情况下, 井下粉尘的浓度可达 $3\,000 \text{ mg} \cdot \text{m}^{-3}$ 。为模拟煤矿综采工作面环境粉尘情况, 设置粉尘浓度分别为 $600, 1\,000, 1\,500$ 和 $3\,000 \text{ mg} \cdot \text{m}^{-3}$ 。通过控制注入密封箱体內的粉尘质量实现粉尘浓度设置, 颗粒质量与粉尘浓度的计算式为

$$c = \frac{1\,000 \times M}{V} \quad (3)$$

式(3)中, M 为注入粉尘的质量, 单位 g ; V 为箱体体积, 单位 m^3 ; c 为粉尘的浓度, 单位 $\text{mg} \cdot \text{m}^{-3}$ 。试验装置密闭箱体的体积为 0.48 m^3 , 由式(3)获得不同粉尘浓度下粉尘注入量, 列于表 1。

表 1 实验粉尘注入量

Table 1 Experimental dust injection

粉尘浓度/ $(\text{mg} \cdot \text{m}^{-3})$	粉尘注入量/ g
600	0.28
1 000	0.48
1 500	0.72
3 000	1.44

由于该实验中采用了大功率卤素灯来保持实验的光辐照度, 为保证煤粉不发生燃爆, 选用挥发分质量最低的无烟煤煤粉和抑爆剂碱式碳酸镁按 9:1 混合作为实验粉尘^[5]。为保证实验用粉尘符合实验要求, 把实验用粉尘通过 800 目标标准筛, 保证实验粉尘的中粒径在 $22 \mu\text{m}$ 以下。

1.3 煤岩样本

从全国各地收集了常见的无烟煤、烟煤、褐煤三个大类的 12 个典型煤样以及常见的页岩、砂岩、灰岩三大类沉积岩的 11 种典型煤系岩样, 为了保证实验更接近于采掘现场, 实验样本采用现场原位样本, 并用密封袋保存, 各个样本的编号、类别、来源如表 2 所示。

表 2 实验样本表

Table 2 Experimental samples table

样本编号	样本类型	样本类别	产地及煤矿
1	无烟煤	无烟煤 1 号	宁夏汝其沟煤矿
2	无烟煤	无烟煤 2 号	山西新景煤矿
3	烟煤	贫煤	山东东风煤矿
4	烟煤	贫瘦煤	山东东风煤矿
5	烟煤	瘦煤	山东埠村煤矿
6	烟煤	焦煤	山西马兰煤矿
7	烟煤	肥煤	山东曹庄煤矿
8	烟煤	1/3 焦煤	山东八一煤矿
9	烟煤	气肥煤	山东茅庄煤矿
10	烟煤	气煤	山东兴隆庄煤矿
11	褐煤	褐煤 1 号	山东五图煤矿
12	褐煤	褐煤 2 号	吉林舒兰煤矿
13	页岩	碳质页岩	山西马兰煤矿
14	页岩	黑色页岩 1 号	山西马兰煤矿
15	页岩	黑色页岩 2 号	山西新景煤矿
16	页岩	沙质页岩 1 号	山东东风煤矿
17	页岩	沙质页岩 2 号	山东兴隆庄煤矿
18	砂岩	中粒砂岩	山东东风煤矿
19	砂岩	细砂岩	山东东风煤矿
20	砂岩	粉砂岩 1 号	山西马兰煤矿
21	砂岩	粉砂岩 2 号	山东兴隆庄煤矿
22	灰岩	泥质灰岩 1 号	山东兴隆庄煤矿
23	灰岩	泥质灰岩 2 号	山东兴隆庄煤矿

2 结果与讨论

2.1 标准光谱数据库的建立

从近红外反射光谱可以分析了解在晶体场作用下由于离子能级的跃迁引起的吸收特征变化,而这些吸收谷含有丰富的物质成分信息,是物质鉴别的关键所在。

由图 2 可知,典型煤系岩石主要在 1 400, 1 900 和 2 200 nm 出现特征吸收谷,其中, 1 400 nm 波长的吸收谷主要是由矿物中结构水的 OH^- 引起的, 1 900 nm 波长处的吸收谷主要是由石英包体水(乳石英)和样本的吸附水引起的, 2 200 nm 波长处的吸收谷主要是由钠长石、高岭石中 $\text{Al}-\text{OH}$ 引起的;由于典型高煤阶煤样无烟煤、烟煤含有大量的不透明碳质物质,这些不透明碳质物质遮蔽结构水、石英包体水、 $\text{Al}-\text{OH}$ 在 1 400, 1 900 和 2 200 nm 波长处的吸收谷特征,所以高阶煤煤样在全近红外波段上无明显的吸收特征,而低阶煤褐煤中,由于灰分产率最高,其性状接近岩石,所以在褐煤类样本中出现了明显的吸收特征^[6]。

从试样表 23 个试样中从每个样本类型中随机取出 1 个样本作为测试试样,其中煤样 3 个,岩样 3 个,把 23 个试样的光谱曲线作为标准光谱数据库。

2.2 粉尘环境下的煤岩光谱

把表 2 中的样本分别在前述粉尘浓度下进行近红外光谱的采集,通过图 3 和图 4 对比可知,粉尘的加入导致波长 1 000~1 200 nm 与波长 2 400~2 500 nm 的光谱出现异常波

动,光谱信噪比降低,其产生的原因为卤钨光源的辐照度曲线为凸型,其最大辐照度所处的波长为 1 400 nm,所以前述两波段的光能量较小,受粉尘消光系数影响大,使得光谱信噪比降低。

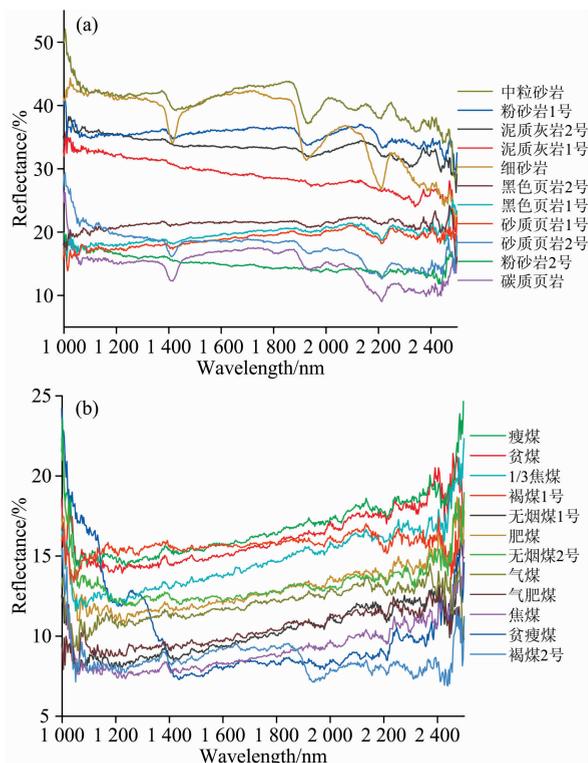


图 2 23 个煤岩试样的光谱曲线

(a): 11 个典型煤系岩石试样光谱曲线;

(b): 12 个典型煤试样光谱曲线

Fig. 2 Spectra of 23 coal and rock samples

(a): 11 typical coal measures samples

(b): 12 typical coal samples

由图 3 可知,在煤样本中,煤的反射率偏低,由于样本 2 无烟煤 2 号与样本 10 气煤为高阶煤,灰分产率较低,其基质(主要是 C, H 和 O 元素组成的有机物)和固定碳含量较多,试样中含有的不透明碳质物质较多,所以光谱较为平滑没有明显的吸收谷,其光谱反射率较小并趋于水平;由于样本 12 褐煤 2 号为低阶煤,灰分产率较高,试样中的不透明物质较低,所以石英胞体水的振动以及有机物分子团的振动在 1 900 和 2 300 nm 附近有明显的吸收谷,由于粉尘主要物质为无烟煤,随着粉尘浓度的增加,无烟煤中不透明物质增多,遮蔽作用使得褐煤 2 号特征吸收谷减弱。

由图 4 可知,在岩石样本中样本 17 砂质页岩 2 号分别在 1 400, 1 900 和 2 200 nm 有特征吸收谷, 1 400 nm 主要由石英胞体水产生,样本中 1 400 nm 吸收谷主要由黏土矿物多硅白云母、高岭石中的 OH^- 振动产生, 1 900 nm 的吸收谷主要为吸附水产生, 2 200 nm 的吸收谷主要由钠长石中 $\text{Al}-\text{OH}$ 的振动产生,随着粉尘浓度的增加, 1 400, 1 900 和 2 200 nm 的特征吸收谷减小;样本 20 粉砂岩 1 号分别在

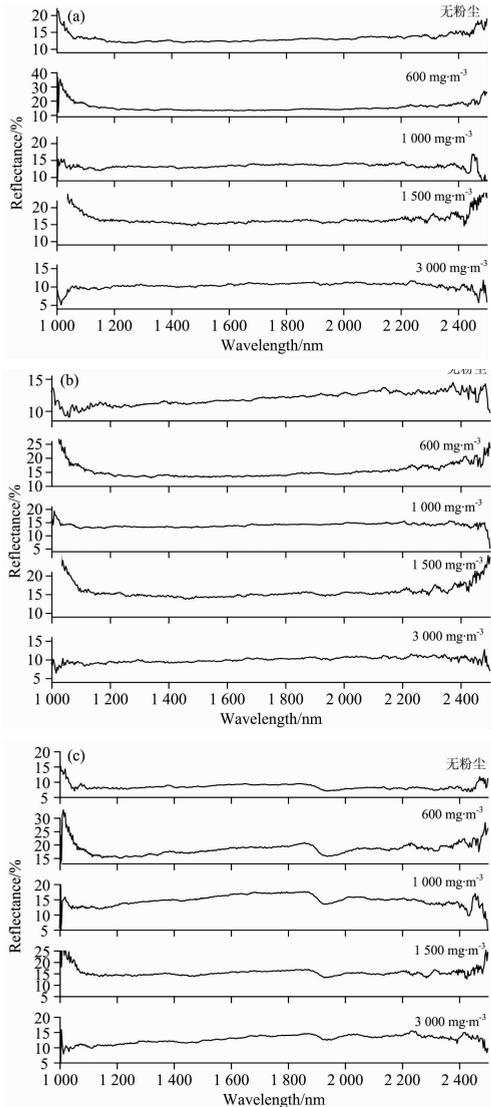


图 3 典型煤样本不同粉尘浓度的近红外光谱
(a): 样本 2 无烟煤 2 号; (b): 样本 10 气煤;
(c) 样本 12 褐煤 2 号

Fig. 3 Near-infrared spectroscopy of different dust concentrations in typical coal samples

- (a): Sample 2, anthracite No. 2;
- (b): Sample 10, gas coal;
- (c): Sample 12, lignite coal

1 400, 1 900, 2 200 和 2 350 nm 有明显的特征吸收谷, 其 2 200 nm 的吸收谷主要由钠长石中 Al—OH 的振动产生, 影响 1 400, 1 900, 2 350 nm 吸收峰的主要是高岭石、石英胞体水和方解石, 但高岭石、石英胞体水含量较小, 随着粉尘的增加使得粉尘对 1 400, 1 900 和 2 200 nm 的特征吸收谷减小; 样本 23 泥质灰岩 2 号在 2 350 nm 有明显的吸收谷, 这是由白云石中 CO₃²⁻ 的振动产生, 空气中的 H₂O 吸附使样本 1 900 nm 出现吸收谷, 随粉尘浓度增加, 泥质灰岩 2 号的特征吸收谷减弱。

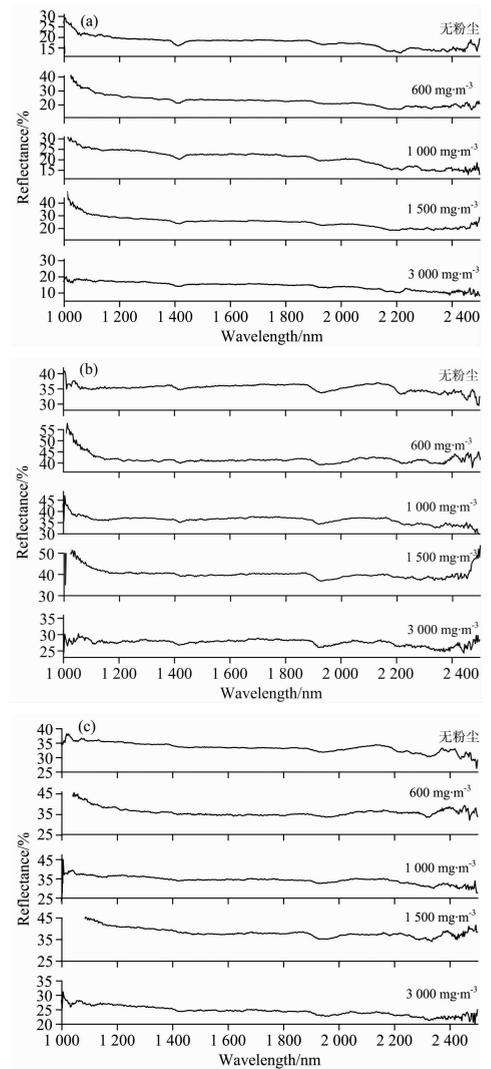


图 4 典型煤系岩样本不同粉尘浓度的近红外光谱
(a): 样本 17 砂质页岩 2 号; (b): 样本 20 粉砂岩 1 号;
(c): 样本 23 泥质灰岩 2 号

Fig. 4 Near-infrared spectra of different dust concentrations in typical rock samples

- (a): Sample 17, sandy shale No. 2;
- (b): Sample 20, siltstone No. 1;
- (c): Sample 23, argillaceous limestone No. 2

综上, 随着粉尘浓度的增加, 粉尘中的无烟煤的不透明物质含量增加, 使得实验样本的特征吸收谷减弱。

为了评价不同浓度的粉尘对光谱的影响, 对所选取的实验样本按不同的粉尘浓度的数据与该样本的标准样本库数据进行相似系数的求解, 相似系数越大说明粉尘对煤岩近红外反射光谱曲线的影响越小。光谱相似系数通常用光谱角度匹配

(1) 光谱角度匹配(SAM)

$$\cos\theta = \frac{\sum_{k=1}^n x_{ik} y_{ik}}{\sqrt{(\sum_{k=1}^n x_{ik}^2)(\sum_{k=1}^n y_{ik}^2)}} \quad (1)$$

式(1)中, x_{ik} 表示实验样本中第 i 个样本的第 k 个波长的反射率; y_{ik} 为第 i 个样本的第 k 个波长在标准样本库中的样本数据; $\cos\theta$ 是实验样本数据与标准样本库数据广义夹角余弦, $\cos\theta$ 越大光谱数据的相似程度越大。

图 5 为实验样本在不同浓度下所采集的近红外反射光谱数据与标准样本库数据的夹角余弦, 由图 5 可知, 煤样本和岩样本在不同粉尘浓度下所采集的曲线和标准光谱库中的曲线的余弦夹角均处于 0.9 以上, 有着较大的匹配度; 其中, 煤类样本中, 褐煤类样本 12 和褐煤 2 号的光谱曲线对粉尘浓度的变化较为敏感, 由于样本 12 在粉尘浓度 $3\ 000\ \text{mg}\cdot\text{m}^{-3}$ 下新增了一个波长为 $2\ 200\ \text{nm}$ 较大的吸收谷, 导致光谱匹配度下降; 其余样本在各个粉尘浓度下没有出现新的较大吸收谷, 所以测试样本与样本库中样本的余弦夹角值接近 1, 其光谱曲线高度匹配。

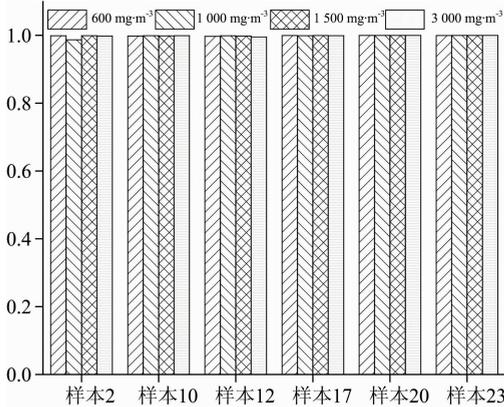


图 5 实验样本与标准样本库的夹角余弦

Fig. 5 Angle cosine between experimental samples and standard samples

(2) 皮尔逊相关系数

$$r_{xy} = \frac{\sum_{k=1}^n (x_{ik} - \bar{x}_i)(y_k - \bar{y})}{\sqrt{\sum_{k=1}^n (x_{ik} - \bar{x}_i)^2 \sum_{k=1}^n (y_k - \bar{y})^2}} \quad (2)$$

式(2)中, n 为样本数量; x_{ik} , y_k 为波长 k 的光谱数据; x_i 为第 i 个实验样本的数据; y 为标准样本库中的样本数据; \bar{x}_i , \bar{y} 为第 i 个实验样本和该样本的标注数据的所有光谱数据均值。 r_{xy} 为第 i 个实验样本数据与标准数据的相关系数; r_{xy} 越大, 所测的两条光谱曲线的相似程度越高。

图 6 为实验样本在不同浓度的条件下所采集的近红外光谱数据与标准样本库数据的皮尔逊相关系数。对比图 5、图 6 可看出, 相比于光谱角度匹配法(SAM)来说, 由于该算法与光谱数据中的均值有关, 所以不仅新增较大的吸收谷对相关系数结果有影响, 而且不同浓度下反射率的变化结果也较为敏感。煤类样本中, 由于无烟煤类样本 2 号为高阶煤样, 反射率最低, 所以在粉尘加入后使得光谱曲线在 $2\ 200\ \text{nm}$ 剧烈震荡, 导致相关系数降低; 烟煤类样本 10 的反射率变化较低, 且无明显吸收谷的产生, 所有相关系数均保持在 0.7 以上, 浓度对其影响较小; 褐煤类样本 12 和褐煤 2 号由于反射

率的变化较大, 光谱曲线的斜率由负变为了正, 所以导致相关系数偏低; 岩类样本中, 除灰岩类样本 23 在 $600\ \text{mg}\cdot\text{m}^{-3}$ 相关系数为 0.4 外, 由于样本为灰白色, 反射率较高, 信噪比比煤类样本高, 所以均有着相比煤类样本有着较大的相关系数; 通过计算各个实验样本的平均相关系数, 各个样本浓度的平均相关系数为 0.72。

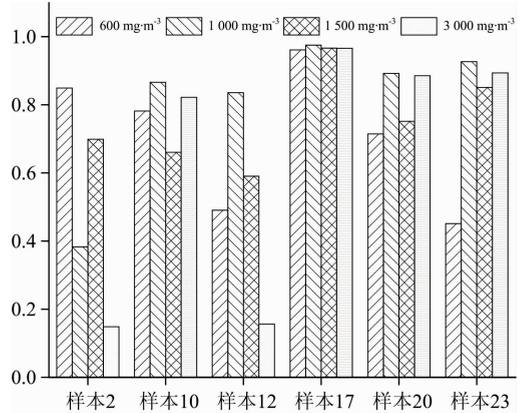


图 6 实验样本与标准样本库的相关系数

Fig. 6 Correlation between experimental samples and standard samples

2.3 试样光谱曲线的处理

由于光谱采集过程受仪器性能和粉尘的影响, 导致大部分试样的光谱在 $1\ 000\sim 1\ 200$ 和 $2\ 400\sim 2\ 500\ \text{nm}$ 出现极大的噪声, 为提高数据的信噪比, 舍去各个浓度的样本 $1\ 000\sim 1\ 200$ 和 $2\ 400\sim 2\ 500\ \text{nm}$ 的光谱数据并进行光谱预处理。

采用 Savitzky-Golay 平滑算法对数据进行去噪平滑。Savitzky-Golay 卷积平滑算法具体如式(3)

$$x_{i,\lambda}^* = \frac{\sum_{j=-r}^r x_{i,j} n_j}{\sum_{j=-r}^r n_j} \quad (3)$$

式(3)中, $x_{i,\lambda}$ 为波长为 λ 点的光谱数据; n 为滤波窗口宽度, 取 $n=2r+1$; $x_{i,\lambda}^*$ 为 Savitzky-Golay 卷积平滑后的数据。在建立标准光谱数据库时, 首先运用该算法对 $1\ 200\sim 2\ 400\ \text{nm}$ 的光谱数据进行去噪处理, 滤波窗口为 3。

为了校正因光谱散射引起的光谱差异, 消除基线漂移, 使用 SNV 算法对光谱数据库进行进一步处理, SNV 算法为

$$X_{i,\text{SNV}} = \frac{X_{i,k} - \bar{X}_i}{\sqrt{\frac{\sum_{k=1}^m (X_{i,k} - \bar{X}_i)^2}{S_i}}} \quad (4)$$

$$\bar{X}_i = \frac{1}{m} \sum_{k=1}^m X_{i,k} \quad (5)$$

式中 $X_{i,\lambda}$ 为第 i 个样本在 k 个波长点的光谱吸光度; m 为光谱的维数; $X_{i,\text{SNV}}$ 为经校正后的光谱吸光度如图 7 所示。

采用 SG 去噪和 SNV 预处理后的各个浓度的典型煤岩样本数据和经同样算法处理后的标准样本库进行 SAM 光谱角度匹配及皮尔逊相关系数的求解, 如图 8 所示, 其中由于 SAM 光谱角度在未经预处理就有极高的光谱匹配度, 所以

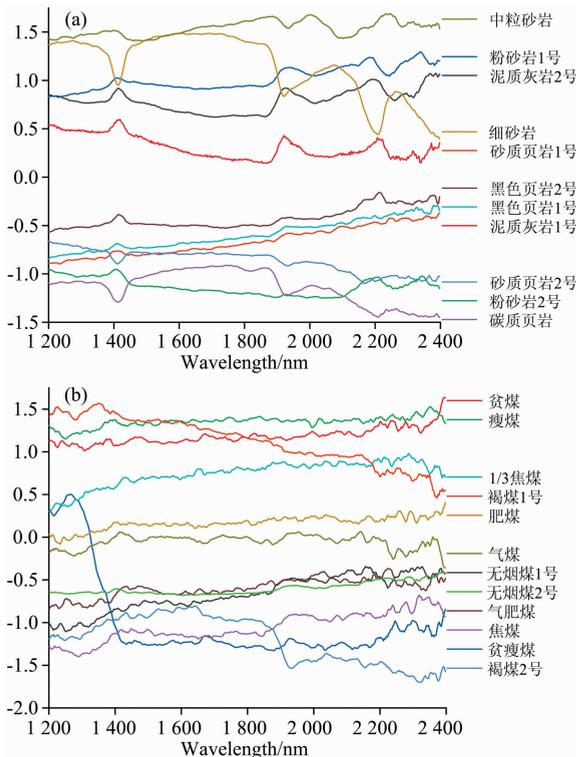


图 7 SG+SNV 处理后的标准光谱库

(a): 典型煤系岩石样本数据库;
(b): 典型煤样本数据库

Fig. 7 Processed standard spectral database by SG+SNV

(a): Typical coal measures samples;
(b): Typical coal samples

在经过 SG+SNV 预处理后, 各个粉尘浓度下的典型煤岩光谱曲线与标准光谱数据库匹配程度变化不大, 除 17 号样本外匹配度均在 0.9 以上, 17 号样本砂质页岩预处理后出现异常降低, 可能是因为 17 号样本的光谱数据特征吸收谷吸收较低, 经去噪处理后吸收谷消失; 对比图 6 与图 9 可知, 样本 12 褐煤 2 号相关性较低, 主要原因为样本 12 的信号强度是在所有信号里最低的[由图 2(b)图 7(b)可知], 经 SNV 算法后, 光谱曲线产生较大偏差, 所以只采用经过 SG 预处理算法的样本 12 实验光谱曲线与标准光谱曲线进行相关系数的计算, 在 600, 1 000, 1 500 和 3 000 $\text{mg} \cdot \text{m}^{-3}$ 粉尘环境下的相关系数分别为: 0.492, 0.837 7, 0.605 9 和 0.157 6 每个粉尘浓度的相关系数相比于未经预处理的均有上涨; 除样本 12 以外, 其余所有样本和标准样本库的样本经 SG+SNV 预处理后光谱相关系数都显著增加, 平均相关系数为 0.78, 平均提升 0.13, 样本 2 无烟煤 2 号样本相关系数提升 76.3%。

2.4 不同浓度试样光谱曲线的聚类分析

以光谱角度匹配(SAM)和皮尔逊相关系数算法为基础, 对不同浓度条件下的实验样本进行聚类分析。根据前述 23 个煤岩样本的煤岩特征, 把样本库分为煤样 12 个, 岩样 11 个, 二值化煤岩样本, 把煤样本设为“0”, 岩样本设为“1”; 选择 SG+SNV 预处理后的样本数据库为标准数据库, 并把

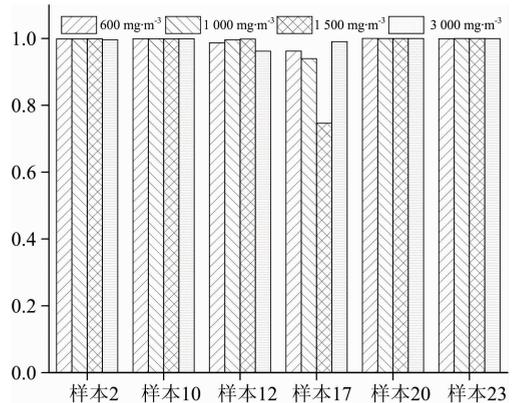


图 8 实验样本与标准样本库的夹角余弦

Fig. 8 Angle cosine between experiment samples and standard samples

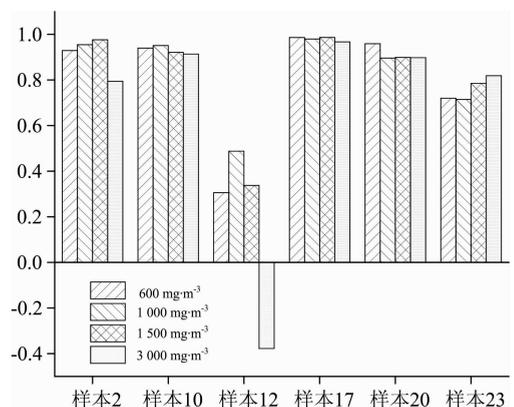


图 9 实验样本与标准样本库的相关系数

Fig. 9 Correlation coefficient between experimental samples and standard samples

数据库中的样本以实样样本的方式进行二值化处理, 取前述随机选取的不同浓度的实验样本作为煤岩类型的测试样本, 并对其进行实验编号如表 3 所示。

分别计算各个不同浓度的实验样本的光谱数据与每个标准样本库数据的余弦夹角与相关系数, 对标准样本库的遍历, 把实验样本数据与标准样本库数据的余弦夹角和相关系数大小最大的归为同一类进行定性分析, 两种识别模型的识别结果如图 10 所示。

由图 9 和 10(b)可知, 样本 2 在 1 000 及 3 000 $\text{mg} \cdot \text{m}^{-3}$ 粉尘条件和具有高相关系数情况下均出现了错误识别, 其主要的原因: 皮尔逊相关系数对波形相似、特别是纵坐标相差很小的两个样本曲线会出现较大误差, 由于样本 14 黑色页岩含碳量较高, 样本 2 在经过 SG+SNV 预处理后其波形与数据库中的样本 14 黑色页岩 1 号相近, 通过相关系数算法识别模型求解导致实验样本的错误识别。

采用识别时间(T)以及识别精度(P)对两种识别模型进行评价。光谱角度匹配 SAM 和皮尔逊相关系数对 24 个测试样本的煤岩类型判定时间(T)为 Matlab 识别程序运行时间, 识别精度(P)的计算方法如式(6)

表 3 煤岩聚类分析样本表

Table 3 Sample table for coal rock cluster analysis

样本类型	样本序号	实验浓度/(mg · m ⁻³)
	1	600
样本 2	2	1 000
无烟煤 2 号	3	1 500
	4	3 000
	5	600
样本 10 气煤	6	1 000
	7	1 500
	8	3 000
	9	600
样本 12 褐煤 2 号	10	1 000
	11	1 500
	12	3 000
	13	600
样本 17 砂质页岩 2 号	14	1 000
	15	1 500
	16	3 000
	17	600
样本 20 粉砂岩 1 号	18	1 000
	19	1 500
	20	3 000
	21	600
样本 23 泥质灰岩 2 号	22	1 000
	23	1 500
	24	3 000

$$P = \frac{N}{24} \times 100\% \quad (6)$$

式(6)中, N 为测试样本的样本类型被正确识别的个数, P 为识别算法的识别精度。

由表 4 可知, 两种不同识别算法识别精度均达到 87% 以上, 相比于皮尔逊相关系数, 光谱角度匹配有着更高的识别效率和识别准确度, 所以光谱角度匹配 SAM 的方法对不同浓度的试样的煤岩定性识别更为精确。

3 结 论

(1) 粉尘的加入使样本在近红外波段 1 000~1 200 和 2 400~2 500 nm 的光谱数据信噪比降低; 随着粉尘浓度的增加, 粉尘中的无烟煤的不透明物质使得实验样本中的特征吸收谷减弱。

(2) 光谱角度匹配法(SAM)在预处理前后均具有极高的光谱匹配度, 除砂质页岩 2 号在粉尘浓度为 1 500 mg · m⁻³ 的光谱匹配度为 0.7 外, 其他所有样本的光谱匹配度在 0.9

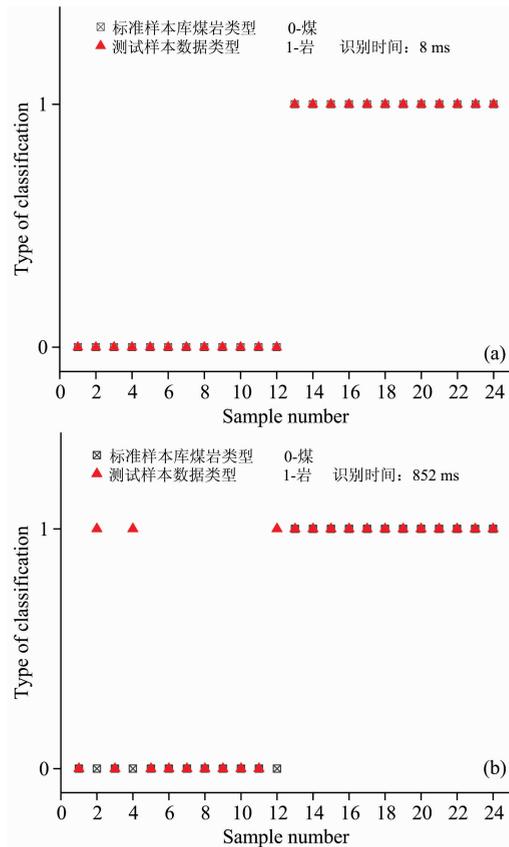


图 10 两种识别模型对实验样本的煤岩识别结果
(a): 光谱角度匹配 SAM; (b): 皮尔逊相关系数

Fig. 10 Coal and rock identification results of experimental samples by two recognition models

(a): Spectral angle matching (SAM);
(b): Pearson correlation coefficient

表 4 两种识别模型的煤岩识别结果

Table 4 Coal and rock identification results of two models

识别算法	T/ms	P/%
光谱角度匹配	8	100
相关系数	852	87.5

以上; 相关系数经预处理有显著提升, 预处理前的平均相关系数为 0.72, 预处理后的平均相关系数为 0.78; 除褐煤 2 号外, 所有的样本光谱相关系数平均提升 0.13, 无烟煤 2 号样本各个浓度平均相关系数提升 76.3%, 褐煤 2 号的光谱相关系数经光谱预处理降低。

(3) 光谱角度匹配法(SAM)在粉尘环境中对典型煤岩样本的定性分析效果比皮尔逊相关系数法好。光谱角度匹配法(SAM)煤岩识别精度为 100%, 识别时间为 8 ms, 皮尔逊相关系数煤岩识别精度为 87.5%, 识别时间为 852 ms。

References

- [1] Zaini N, van der Meer F, van der Werff H. *Remote Sensing*, 2014, 6(5): 4149.
- [2] YANG En, WANG Shi-bo, GE Shi-rong, et al(杨恩, 王世博, 葛世荣, 等). *Journal of China Coal Society(煤炭学报)*, 2019, 44(12): 3912.
- [3] WANG Zi-liang(王自亮). *Journal of China Coal Society(煤炭学报)*, 2000, 25(4): 404.
- [4] JIAN Jie, XIE Jian-lin, GUO Yong-yi(菅洁, 谢建林, 郭勇义). *Journal of Taiyuan University of Technology(太原理工大学学报)*, 2017, 48(4): 592.
- [5] GAO Qiang-jian, HE Yun-feng, LIU Zhao, et al(高强健, 何耘丰, 刘翌, 等). *Iron & Steel(钢铁)*, 2018, 53(12): 30.
- [6] YAN Shou-xun, ZHANG Bing, ZHAO Yong-chao, et al(燕守勋, 张兵, 赵永超, 等). *Remote Sensing Technology and Application(遥感技术与应用)*, 2003, 18(4): 191.

Study on Near-Infrared Spectrum Features and Identification Methods of Typical Coal-Rock in Dust Environment

XIANG Yang, WANG Shi-bo*, GE Shi-rong, WANG Sai-ya, ZHOU Yue, LÜ Yuan-bo, YANG En

School of Mechanical and Electrical Engineering, China University of Mining and Technology (Xuzhou), Xuzhou 221116, China

Abstract In order to study the dust problem existing in the identification of near-infrared coal and rock in the underground coal mine, the mixture of anthracite and the anti-explosive agent was used to simulate the dust environment of coal mine underground so that the experimental device for spectral identification of coal and rock in dust environment was constructed. 23 samples of in-situ typical coal samples of shale, sandstone and limestone samples and anthracite, bituminous coal and lignite coal samples were collected from all over the country. 23 coal and stone samples reflection spectrum collecting without the dust of the near-infrared band (1 000~2 500 nm) was used as an experimental standard database. 1 sample was randomly selected from the three typical coal samples and three typical rock samples in the experimental standard the sample library as experimental samples, and the test samples' reflectance spectra of the near-infrared bands at 600, 1 000, 1 500 and 3 000 $\text{mg} \cdot \text{m}^{-3}$ dust concentrations were collected. The results showed that the addition of dust led to a decrease in the signal-to-noise ratio of the spectral image between 1 000~1 200 and 2 400~2 500 nm; With the increase of dust concentration, The opaque substance of anthracite in the dust made the characteristic absorption valley in the experimental sample weaken; Correlation analysis between sample and standard sample library was carried out by spectral angle matching SAM and Pearson correlation coefficient. Anthracite samples, bituminous coal samples, sandstone samples and limestone samples had a high matching degree under SAM matching model. The cosine angle was above 0.9 at each dust concentration; Correlation coefficient matching model matching degree was strongly affected by dust, and the average correlation coefficient was 0.73; After the experimental standard database and special envoy samples were normal preprocessed by SG convolution and SNV standard, the matching degree of SAM matching model did not change significantly. Correlation coefficient matching model matching degree was significantly improved, the average correlation coefficient was 0.78; The correlation coefficient matching model excepted for lignite No. 2, the spectral correlation coefficient of all samples increased by 0.13, anthracite No. 2 The sample correlation coefficient increased by 76.3%, while the spectral correlation coefficient of sample 12 lignite No. 2 was reduced by spectral pretreatment. The spectral angle matching SAM and Pearson correlation coefficient coal and rock identification model was established. The two models were used to identify coal samples under different concentrations and binarized coal rock sample, the coals were "0", and the rocks were "1". Coal rock identification was performed on 6 experimental samples at different concentrations by two recognition models. The identification accuracy of SAM was 100%, and the recognition time was 8 ms. Pearson correlation recognition accuracy of the coefficient P was 87.5%, and the recognition time was 852 ms.

Keywords Coal and rock identification; Typical coal rock; Near infrared spectroscopy; Dust environment

* Corresponding author

(Received Nov. 11, 2019; accepted Mar. 20, 2020)