

文章编号:1671-251X(2021)12-0093-05

DOI:10.13272/j.issn.1671-251x.17853

面向煤矸识别的近红外反射光谱数据预处理方法

丁震¹, 常博深²

(1. 国家能源集团 煤炭运输部, 北京 100013;
2. 中国矿业大学(北京) 机电与信息工程学院, 北京 100083)



扫码移动阅读

摘要:利用近红外反射光谱进行煤矸识别时,光谱采集装置距工作面的探测距离变化及粉尘干扰会对近红外反射光谱产生影响。为选取最佳的煤矸近红外反射光谱预处理方法,收集了外观相近的无烟煤和矸石样本,在实验室搭建了由近红外光谱仪、准直镜、卤素灯等组成的光谱采集装置,采集了不同探测距离(1.2, 1.5, 1.8 m)和粉尘浓度(200, 500, 800 mg/m³)下的煤矸近红外反射光谱。通过对煤矸近红外反射光谱特征分析,发现探测距离和粉尘浓度变化对煤矸近红外反射光谱曲线波形和吸收谷位置无明显影响,即不会改变光谱特征吸收波长点,但对煤矸近红外反射光谱的反射率产生明显影响,即光谱反射率随着探测距离和粉尘浓度的增大而减小,会造成煤矸近红外反射光谱漂移。为增强煤矸近红外反射光谱吸收特征,利用微分、标准正态变量变换和多项式平滑 3 种方法对光谱数据进行预处理,并将经过预处理的煤矸近红外反射光谱数据输入至粒子群优化 BP 神经网络模型进行煤矸识别。实验结果表明,微分预处理方法对探测距离和粉尘浓度变化下采集的煤矸近红外反射光谱数据的优化效果最佳,能有效消除探测距离和粉尘浓度变化对光谱反射率的影响。

关键词:煤矸识别; 近红外反射光谱; 数据预处理; 探测距离; 粉尘浓度
中图分类号:TD67 文献标志码:A

Near-infrared reflectance spectrum data preprocessing method for coal gangue identification

DING Zhen¹, CHANG Boshen²

(1. Coal Transportation Department, CHN Energy, Beijing 100013, China; 2. School of Mechanical Electronic and Information Engineering, China University of Mining and Technology(Beijing), Beijing 100083, China)

Abstract: When using near-infrared reflectance spectrum to identify coal gangue, the change of detection distance between spectrum acquisition device and working face and dust interference will affect near-infrared reflectance spectrum. In order to select the best pre-processing method for near infrared reflectance spectrum for coal gangue, samples of anthracite and gangue with similar appearance are collected. A spectrum acquisition device consisting of near infrared spectrometer, collimator and halogen lamp is set up in the laboratory to acquire near infrared reflectance spectrum of coal gangue at different detection distances (1.2, 1.5, 1.8 m) and dust concentrations (200, 500, 800 mg/m³). Through the analysis of near-infrared reflectance spectrum characteristics of coal gangue, it is found that the detection distance and dust concentration change have no obvious impact on the waveform of near-infrared reflectance spectrum curve and the position of absorption valley of coal gangue. The absorption wavelength point of spectral characteristics will not be changed. However, the reflectance of near-infrared reflectance spectrum

收稿日期:2021-10-20;修回日期:2021-12-12;责任编辑:盛男。
基金项目:国家自然科学基金项目(52104169)。
作者简介:丁震(1980—),男,山西万荣人,高级工程师,主要从事煤矿机电技术、煤矿智能化、露天卡车无人驾驶方面的研究及管理工作,
E-mail:zhen.ding@chnenergy.com.cn。
引用格式:丁震,常博深. 面向煤矸识别的近红外反射光谱数据预处理方法[J]. 工矿自动化,2021,47(12):93-97.
DING Zhen, CHANG Boshen. Near-infrared reflectance spectrum data preprocessing method for coal gangue identification[J].
Industry and Mine Automation,2021,47(12):93-97.

of coal gangue will be significantly affected. The spectral reflectance will decrease with the increase of detection distance and dust concentration, which will cause near-infrared reflectance spectrum drift of coal gangue. In order to enhance the absorption characteristics of near-infrared reflectance spectrum of coal gangue, the spectrum data are preprocessed by differential, standard normal variable transformation and polynomial smoothing methods. The preprocessed near-infrared reflectance spectrum data of coal gangue are input to the particle swarm optimization BP neural network model for coal gangue identification. The experimental results show that the differential preprocessing method has the best optimization effect on the near-infrared reflectance spectrum data of coal gangue collected under the change of detection distance and dust concentration, and can eliminate the impact of detection distance and dust concentration on the spectral reflectance effectively.

Key words: coal gangue identification; near-infrared reflectance spectrum; data preprocessing; detection distance; dust concentration

0 引言

高效精准的煤矸识别是煤炭开采领域公认的难题,是实现煤炭无人化、智能化开采的关键^[1]。众多学者提出了多种煤矸识别方法,如天然射线探测法^[2]、声波探测法^[3]、热红外分析法^[4]、图像分析法^[5]和振动探测法^[6]等,但上述方法因适用环境约束问题没有被广泛应用。近红外反射光谱技术是一种针对原位物质的光谱分析技术^[7-8],其识别机理是被测物质的特定化学成分和物质结构在特定波长处存在光谱反射吸收特性,因此可根据该特性对煤矸属性进行反演。杨恩等^[9]针对煤矸反射光谱特征,建立了基于支持向量机的煤矸识别方法。韦任等^[10]利用煤矸反射光谱数据在 2 130~2 250 nm 波段的吸收特征,采用随机森林和支持向量机算法进行煤矸识别。然而,在利用近红外反射光谱数据进行煤矸定性分析时发现,由于工作面煤矸的形状不规则且堆积高度不确定,即使光谱采集装置固定,装置距工作面的探测距离仍不断变化,会导致近红外反射光谱数据变化^[11],且在粉尘干扰的工况条件下采集的近红外反射光谱数据含有大量无关信息,所以,对近红外反射光谱数据进行预处理尤为重要。笔者从某矿井综采工作面顶底板处收集了外观相近的无烟煤和矸石样本共计 100 个(煤样 75 个,矸石样本 25 个),在实验室采集了不同探测距离和粉尘浓度下的煤矸近红外反射光谱,分别使用微分、标准正态变量变换和多项式平滑 3 种方法对近红外反射光谱数据进行预处理,从而确定最佳的近红外反射光谱数据预处理方法。

1 煤矸近红外反射光谱采集

使用透明 PVC 板材组成密闭空间,密闭空间内煤矸近红外反射光谱采集装置布置如图 1 所示。

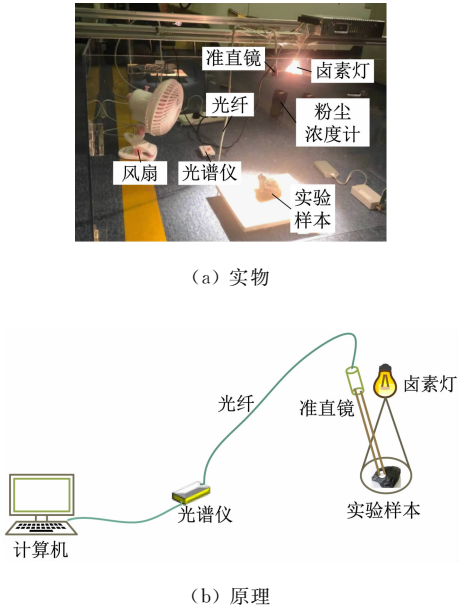


图 1 煤矸近红外反射光谱采集装置

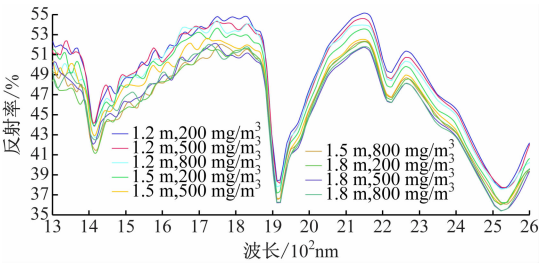
Fig. 1 Coal and gangue near-infrared reflectance spectrum acquisition device

准直镜(视场角 $\theta=0.12^\circ$, 镜头直径 $d=2.5\text{ cm}$)和卤素灯并列放置;准直镜与 NeoSpectra 近红外光谱仪(波长范围为 1 000~2 500 nm,分辨率为 8 nm)之间使用石英光纤连接;计算机通过 USB3.0 与近红外光谱仪连接,实时显示煤矸样本近红外反射光谱曲线。为减小被测煤矸样本双向反射特性的影响,使用 100 W 的卤素灯以竖直方向照射样本表面中心,形成光照度约为 20 000 lx 的圆形光斑。使用近红外光谱仪进行光谱采集,采集每条光谱曲线耗时约为 50 ms。为限制光谱采集时光的散射现象,使用准直镜对光束进行准直,调节准直镜使其轴线垂直对准样本表面的光斑中心,准直镜与光斑中心之间的距离为 L 。近红外光谱仪所采集光谱的范围为准直镜视场角在样本表面所形成圆面,底面圆的直径 D 、面积 S 分别为

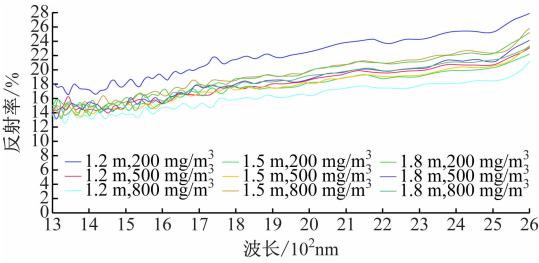
$$D = d + 2L \tan \frac{\theta}{2} \tag{1}$$

$$S = \frac{\pi D^2}{4} \tag{2}$$

设置准直镜与煤矸样本之间的距离分别为 1.2, 1.5, 1.8 m, 粉尘浓度分别为 200, 500, 800 mg/m³。同一煤矸样本在不同探测距离和粉尘浓度下的近红外反射光谱曲线如图 2 所示。



(a) 矸石近红外反射光谱曲线



(b) 煤样近红外反射光谱曲线

图 2 不同探测距离和粉尘浓度下煤矸样本近红外反射光谱曲线

Fig. 2 Near-infrared reflectance spectra of coal and gangue samples under different detection distances and dust concentrations

从图 2 可看出,矸石近红外反射光谱曲线在整个波段存在较为明显的波动,特别是在 1 400, 1 900, 2 500 nm 附近出现了明显的吸收谷;煤样近红外反射光谱反射率较低且变化较为平缓,且在 1 400, 1 900 nm 附近存在微弱吸收谷;随着探测距离和粉尘浓度增大,煤矸近红外反射光谱的反射率减小,但光谱特征吸收波长点不受影响。

2 煤矸近红外光谱数据预处理

由于实验场地自然光强度不断变化,且不能及时使用定标白板进行背景反射光谱标定,所以会存在背景偏移现象,且煤和矸石均富含多种化学成分,其成分的差异性和不均匀性会形成光的散射,使得采集的煤矸近红外反射光谱数据存在大量噪声。为增强煤矸近红外反射光谱吸收特征,消除探测距离和粉尘浓度变化对光谱反射率的影响,利用微分、标准正态变量变换和多项式平滑 3 种方法对煤矸近红外反射光谱数据进行预处理^[12-15]。3 种预处理方法

的核心求解算法如下。

(1) 一阶微分和二阶微分:

$$y_{\text{fir}} = \frac{-2y_j - 2 - y_j - 1 + y_j + 1 + y_j + 2}{10\lambda} \tag{3}$$
$$y_{\text{sec}} = \frac{-2y_j - 2 - y_j - 1 + 2y_j + y_j + 1 + y_j + 2}{7\lambda^2} \tag{4}$$

式中: y_{fir} , y_{sec} 分别为经一阶微分、二阶微分处理后的反射率; y_j 为第 j 个波长点的反射率; λ 为波长间隔。

(2) 标准正态变量变换:

$$y_{\text{SNV}} = \frac{y_{ij} - \bar{y}_i}{\sqrt{\frac{1}{p-1} \sum_{j=1}^p (y_{ij} - \bar{y}_i)^2}} \tag{5}$$

$$\bar{y}_i = \frac{1}{p} \sum_{j=1}^p y_{ij} \tag{6}$$

式中: y_{SNV} 为经标准正态变量变换处理后的反射率; y_{ij} 为第 i 个样本第 j 个波长点的反射率; $\bar{y}_i$ 为第 i 个样本的平均反射率; p 为光谱波长点数。

(3) 多项式平滑:

$$y = \mathbf{Y} \cdot \mathbf{A} + \mathbf{E} \tag{7}$$

式中: y 为经多项式平滑处理后的反射率; $\mathbf{Y}$ 为以波长点 m 为中心的反射率矩阵, $\mathbf{Y} = [y_{m-2\lambda} \ y_{m-\lambda} \ y_m \ y_{m+\lambda} \ y_{m+2\lambda}]$; $\mathbf{A}$ 为拟合矩阵; $\mathbf{E}$ 为平滑矩阵。

3 实验结果及分析

为比较 3 种预处理方法的效果,将经过不同方法预处理后的煤矸近红外反射光谱曲线以 3:1 的比例划分为训练集和测试集,并输入粒子群优化 BP 神经网络模型进行煤矸识别。粒子群优化 BP 神经网络模型参数设置:神经网络隐藏层数为 3,损失函数为 MSE,学习因子为 2,学习率为 0.001。

以整体煤矸识别准确率为评价指标,不同预处理方法下煤矸识别结果见表 1。

从表 1 可看出:经微分预处理后的煤矸识别效果最佳,仅在探测距离为 1.8 m、粉尘浓度为 800 mg/m³ 情况下煤矸识别准确率为 88%,其他条件下识别准确率均在 90% 以上,整体煤矸识别正确率约为 98%;经标准正态变量变换预处理后,仅在探测距离为 1.8 m、粉尘浓度大于 200 mg/m³ 情况下煤矸识别准确率低于 90%,整体煤矸识别正确率约为 95%,这是由于该方法最突出的作用是修正样本个体物质差异造成的光谱数据突变,但不能有效优化探测距离和粉尘浓度对光谱数据的影响;经多项式平滑预处理后的煤矸识别准确率最低,约为

表 1 不同预处理方法下煤矸识别结果

Table 1 Identification results of coal and gangue under different pretreatment methods

粉尘浓度/ (mg·m ⁻³)	识别准确率/%								
	探测距离 1.2 m			探测距离 1.5 m			探测距离 1.8 m		
	微分	标准正态 变量变换	多项式 平滑	微分	标准正态 变量变换	多项式 平滑	微分	标准正态 变量变换	多项式 平滑
200	100	100	100	100	100	100	100	100	92
500	100	100	92	100	100	92	92	88	88
800	100	92	80	100	92	80	88	84	80

89%，这是由于该方法在剔除光谱数据突变的噪声点时，探测距离和粉尘浓度变化使得光谱数据产生的波动及因煤矸本身特性产生的吸收谷特征被过滤掉，造成特定选择的煤矸光谱波段的特征信息部分丢失。

4 结论

(1) 探测距离和粉尘浓度的变化对煤矸近红外反射光谱曲线波形和吸收谷位置无明显影响，即探测距离和粉尘浓度的变化不会影响煤矸特定光谱特征吸收波长点；但对煤矸近红外反射光谱的反射率产生明显影响，即光谱反射率随着探测距离和粉尘浓度的增大而减小，存在煤矸近红外反射光谱漂移现象。

(2) 与标准正态变量变换和多项式平滑预处理方法相比，微分预处理方法对探测距离和粉尘浓度变化下采集的煤矸近红外反射光谱数据的优化效果最佳，可有效消除探测距离和粉尘浓度变化对光谱数据的影响。

参考文献 (References):

[1] 曾翰林. 基于图像处理的煤矸识别技术研究[D]. 唐山:华北理工大学,2015.
ZENG Hanlin. Research on recognition technology of coal and gangue based on image processing [D]. Tangshan: North China University of Science and Technology,2015.

[2] 郭永存,何磊,刘普壮,等. 煤矸双能 X 射线图像多维度分析识别方法 [J]. 煤炭学报, 2021, 46 (1): 300-309.
GUO Yongcun, HE Lei, LIU Puzhuang, et al. Multi-dimensional analysis and recognition method of coal and gangue dual-energy X-ray images[J]. Journal of China Coal Society,2021,46(1):300-309.

[3] 马瑞. 基于声波信号的煤矸界面识别研究[D]. 济南:山东大学,2010.
MA Rui. Study on the coal-rock interface recognition

based on acoustic wave technique [D]. Jinan: Shandong University,2010.

[4] 王中举,毛馨凯,孙江. 基于视频解析的智能煤矸分选技术研究 [J]. 工矿自动化, 2021, 47 (增刊 1): 122-125.
WANG Zhongju, MAO Xinkai, SUN Jiang. Research on intelligent coal-gangue separation technology based on video analysis[J]. Industry and Mine Automation, 2021,47(S1):122-125.

[5] 余乐,郑力新,杜永兆,等. 采用部分灰度压缩扩阶共生矩阵的煤和煤矸石图像识别[J]. 华侨大学学报(自然科学版),2018,39(6):906-912.
YU Le, ZHENG Lixin, DU Yongzhao, et al. Image recognition method of coal and coal gangue based on partial grayscale compression extended coexistence matrix[J]. Journal of Huaqiao University (Natural Science),2018,39(6):906-912.

[6] 窦希杰,王世博,刘后广,等. 基于 EMD 特征提取与随机森林的煤矸识别方法 [J]. 工矿自动化, 2021, 47(3):60-65.
DOU Xijie, WANG Shibo, LIU Houguang, et al. Coal and gangue identification method based on EMD feature extraction and random forest[J]. Industry and Mine Automation,2021,47(3):60-65.

[7] YANG En, GE Shirong, WANG Shibo. Characterization and identification of coal and carbonaceous shale using visible and near-infrared reflectance spectroscopy[J]. Journal of Spectroscopy, 2018(2):1-13.

[8] 宋亮,刘善军,毛亚纯,等. 基于可见光-近红外光谱的煤种分类方法 [J]. 东北大学学报(自然科学版), 2017,38(10):1473-1476.
SONG Liang, LIU Shanjun, MAO Yachun, et al. Coal classification based on visible and near-infrared spectrum [J]. Journal of Northeastern University (Natural Science),2017,38(10):1473-1476.

[9] 杨恩,王世博,葛世荣,等. 基于反射光谱的煤岩感知实验研究[J]. 煤炭学报,2019,44(12):3912-3920.
YANG EN, WANG Shibo, GE Shirong, et al.

Experimental study on coal-rock perception based on reflectance spectroscopy[J]. Journal of China Coal Society,2019,44(12):3912-3920.

[10] 韦任,徐良骥,孟雪莹,等. 基于高光谱特征吸收峰的煤岩识别方法[J]. 光谱学与光谱分析,2021,41(6):1942-1948.

WEI Ren,XU Liangji,MENG Xueying,et al. Coal and rock identification method based on hyper spectral feature absorption peak[J]. Spectroscopy and Spectral Analysis,2021,41(6):1942-1948.

[11] 周悦,王世博,葛世荣,等. 不同探测距离与角度下典型煤岩近红外光谱特征与定性分析[J]. 光谱学与光谱分析,2020,40(9):2737-2742.

ZHOU Yue,WANG Shibo,GE Shirong,et al. Near infrared spectral characteristics and qualitative analysis of typical coal-rock under different detection distances and angles[J]. Spectroscopy and Spectral Analysis,2020,40(9):2737-2742.

[12] 王雅圣,杨梦,骆志远,等. 基于置信学习机与近红外光谱的煤种快速分类方法[J]. 光谱学与光谱分析,2016,36(6):1685-1689.

WANG Yasheng,YANG Meng,LUO Zhiyuan,et al. Rapid coal classification based on confidence machine and near infrared spectroscopy[J]. Spectroscopy and Spectral Analysis,2016,36(6):1685-1689.

[13] 张守祥,张学亮,刘帅,等. 智能化放顶煤开采的精确放煤控制技术[J]. 煤炭学报,2020,45(6):2008-2020.

ZHANG Shouxiang,ZHANG Xueliang,LIU Shuai,et al. Intelligent precise control technology of fully mechanized top coal caving face[J]. Journal of China Coal Society,2020,45(6):2008-2020.

[14] 郭靖雯,先怡衡,肖薇,等. 红外光谱的煤精类文物材质判别方法[J]. 光谱学与光谱分析,2021,41(5):1424-1429.

GUO Jingwen,XIAN Yiheng,XIAO Wei,et al. Study on the method of distinguishing the material of jet-like cultural relics based on infrared spectra[J]. Spectroscopy and Spectral Analysis,2021,41(5):1424-1429.

[15] 贾桂红,张建军,郑海明. 基于离散频谱校正的差分吸收光谱数据处理方法研究[J]. 计量学报,2021,42(8):1081-1086.

JIA Guihong,ZHANG Jianjun,ZHENG Haiming. Study on differential optical absorption spectroscopy data processing based on discrete spectrum correction[J]. Acta Metrologica Sinica,2021,42(8):1081-1086.