

基于可见光波段的燃烧与未燃烧矽石分类方法研究

宋 亮, 刘善军*, 毛亚纯, 王 东, 虞茉莉

东北大学资源与土木工程学院, 辽宁 沈阳 110819

摘 要 我国煤矿数量众多,分布广泛,大量堆积的煤矽石对矿区环境造成严重影响,其中部分煤矽石处理不当可能引发自燃和爆炸,对矿区安全构成直接威胁。根据煤矽石的燃烧状态可以分为燃烧矽石和未燃烧矽石两类,其存在的安全隐患和对环境的危害性有所不同,同时其综合利用的途径亦不相同。因此,对煤矽石进行燃烧矽石和未燃烧矽石的分类识别与监测就显得尤为重要。目前的监测方法主要为实地勘查调研,其效率低、成本高,难以满足煤矽石监测的实际需求。选择辽宁省铁法矿区作为研究区,首先从矿区矽石山现场采集典型的煤矽石样本106个;然后,利用SVC HR1024光谱仪测试其可见光-近红外光谱,分析燃烧和未燃烧矽石的光谱特征,并基于可见光波段构建光谱指数NDGI,用于识别燃烧矽石和未燃烧矽石。选择实验室测试的光谱数据和实际卫星遥感数据对该指数进行了验证,并与随机森林法进行对比。结果显示:在350~760 nm燃烧矽石光谱曲线斜率整体较高,在550~630 nm反射率存在陡升现象,而未燃烧矽石在整个可见光波段光谱曲线斜率较低;以0.25作为NDGI指数阈值,可以很好地将燃烧矽石和未燃烧矽石区分开来,实验室样本验证结果显示,NDGI指数的分类精度可达99.1%,高于随机森林分类法的95.2%;现场的验证结果表明,使用铁法矿区的landsat8 OLI数据,并基于NDGI指数对矿区内的矽石山进行燃烧和未燃烧区域识别划分,所提取的燃烧和未燃烧矽石在形态和大小上与Google Earth具有很好的一致性,表明该指数对于矽石的燃烧状态具有很好识别效果。在上述研究基础上,分别取燃烧和未燃烧矽石进行矿物鉴定,通过对比矽石燃烧前后矿物种类的变化,分析造成燃烧和未燃烧矽石的光谱特征差异的原因。结果表明:燃烧使矽石中的 Fe^{2+} 被氧化为 Fe^{3+} 。 Fe^{3+} 的大量增加造成光谱曲线在550nm处形成明显的波谷特征,在整个燃烧过程中生成的玻璃质在750nm处形成高反射率,二者综合造成燃烧和未燃烧矽石的NDGI指数差异。研究结果为煤矿区燃烧和未燃烧矽石的区分识别提供了一种快速、高效、较为准确的实用方法。

关键词 遥感;可见光-近红外光谱;矽石;分类

中图分类号: P237 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2019)04-1148-06

引 言

矽石是煤矿开采过程中的主要固体废弃物。大量矽石的长期堆积可能会引发矽石山自燃和爆炸,对矿区环境和安全造成严重影响^[1-2]。对此采取的主要措施是加大对矽石资源的综合利用,不仅可以消耗大量堆积的煤矽石,有效缓解环境问题,同时也是矿山绿色开采的重要组成部分,将产生显著的经济效益和社会效益^[3-4]。

根据矽石的燃烧状态可以将其分成燃烧与未燃烧矽石两类。其中未燃烧矽石仍存在自燃和爆炸的隐患,而燃烧矽石相对安全。由于矽石自燃前后矿物的种类和含量会发生变

化,二者的利用途径也有所不同。未燃烧矽石主要用于矿井回填,其中部分矽石含有较低的碳质和有机物,具有一定热值,可以用于烧砖、发电等^[5-6]。燃烧矽石经过高温后生成大量玻璃质,使其机械强度大大增加,可以作为铺路材料^[7];磨细后由于具有火山灰活性可以作为生产水泥的原料等^[8]。无论从环保、煤矽石资源化还是提高经济利益的角度看,对燃烧和未燃烧矽石进行有效动态监测都显得非常重要。

目前,关于燃烧和未燃烧矽石的分类识别主要通过实地调研,其效率低、成本高,难以满足矿山的实际需求。因此,需要一种快速、准确、高效的方法对燃烧和未燃烧矽石进行区分。

近年来,遥感技术在矿山环境监测方面取得了快速发

收稿日期:2018-03-15,修订日期:2018-08-05

基金项目:国家自然科学基金项目(41771404)资助

作者简介:宋 亮,1989年生,东北大学资源与土木工程学院博士研究生 e-mail: neusongliang@163.com

* 通讯联系人 e-mail: liusjdr@126.com

展,特别是基于可见光-近红外光谱技术得到广泛应用,国内外学者做了大量的相关研究。Wang 等基于不同煤种的可见光-近红外光谱,结合 PLS 和 SVM 的方法对不同煤种进行分类和成分回归,预测能力有大幅提升^[9];Mao 等基于烟煤和褐煤的可见光-近红外光谱特征,建立煤归一化差异指数(NDCI),用于识别煤的类型,分类结果较好^[10]。宋亮等采用可见光-近红外和热红外光谱联合分析的方法对矿区的煤和矸石进行准确区分,有效地解决了含碳矸石在可见光-近红外波段和煤的“异物同谱”现象^[11]。此外,随着机器学习的发展,其在光谱分类研究和应用也很多。

本研究利用 SVC HR 1024 光谱仪测试燃烧和未燃烧矸石的可见光-近红外光谱,分析二者的光谱特征差异。在此基础上构建归一化差异矸石指数(NDGI),以 0.25 为阈值,将燃烧矸石和未燃烧矸石二者区分开来,并与随机森林分类法进行了对比。最后,利用该指数并基于 Landsat8 OLI 数据进行了实际煤矿的矸石分类提取,收到较好效果。

1 实验部分

1.1 样本采集

以铁法(煤)矿区为研究区,该矿区位于辽宁省北部,煤炭储量达 16.5 亿 t,是辽宁省重要的煤炭生产基地,煤的类型主要为烟煤。在研究区内共采集典型的矸石样本 106 块,其中未燃烧样本 83 块(编号为 1—83),燃烧样本 23 块(编号为 84—106),每块样本的大小约为 8 cm×8 cm×4 cm。

1.2 光谱测试

使用美国 Spectra Vista 公司的 SVC HR-1024 便携式地物光谱仪对样本进光谱测试。该仪器有 1 024 个通道,光谱测试范围为 350~2 500 nm。选择晴朗无云的天气进行光谱测试,时间在 10:00—14:00 之间。在数据采集的过程中,选择 4°镜头,镜头距离观测样本 35 cm。每个样品重复测试 2 次,取其平均值作为最终反射率光谱。

2 光谱特征分析

2.1 燃烧和未燃烧矸石光谱特征

106 块样本的光谱测试结果表明燃烧和未燃烧矸石的可见光-近红外光谱特征存在明显差异。由于样本数量较多,一次不便显示,因此选取部分典型样本进行图形显示。图 1 为选择的 4 个燃烧矸石和 8 个未燃烧矸石样本的光谱测试结果。图 1 清楚地反映了两类矸石的反射率光谱差异。主要光谱特征如下:

(1)燃烧矸石的反射率从 350 nm 开始缓慢增加,在 550~630 nm 反射率出现陡增,随后在 630~760 nm 上升速度变缓。进入近红外波段后,反射率缓慢上升,在 1 400~2 100 nm 基本保持不变,在 2 100~2 500 nm 反射率缓慢降低,并在 1 400 和 1 900 nm 处出现波谷特征,主要是由水汽和羟基引起。

(2)未燃烧矸石在波段 350~580 nm 反射率快速上升,在 580~2 100 nm 缓慢升高,在 2 100~2 500 nm 缓慢降低。同样

由于样本含水在 1 400 和 1 900 nm 附近有明显的波谷特征。

两类矸石光谱主要差别存在于:在 350~760 nm 波段,燃烧矸石的光谱曲线斜率较大,在 550~630 nm 存在陡增现象;而未燃烧矸石在该波段斜率较小。

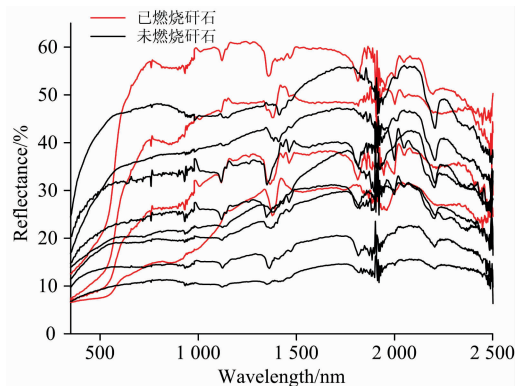


图 1 四个燃烧矸石和八个未燃烧矸石样本的可见光-近红外光谱

Fig. 1 Visible, near-infrared spectral curve of four burned gangue samples and eight unburned gangue samples

2.2 矿区其他地物的光谱

研究中还测试了矿区其他典型地物的光谱,包括煤堆、建筑物、植被和水体,结果如图 2 所示。

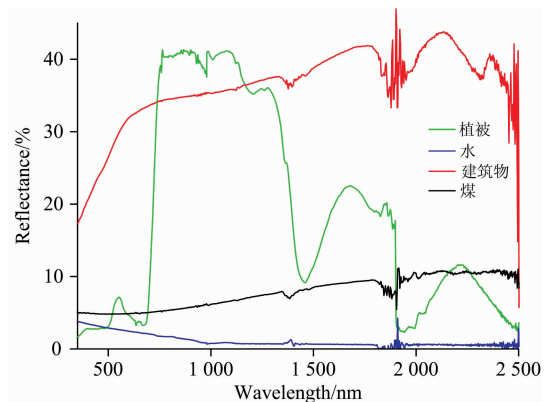


图 2 矿区其他地物的可见光-近红外光谱

Fig. 2 Visible, near-infrared spectral curve of other object in coal time

由图 2 可以看出,与矸石的光谱特征相比,植被(草坪)的光谱特征是:在 350~680 nm 发射率较低且变化较小,之后快速上升,进入近红外波段达到最大,在 1 100 nm 又呈下降趋势,且在 1 400 和 1 900 nm 处存在较强的波谷特征。根据 NDVI 指数可以将其与矸石区分开来。水体的发射光谱非常弱,在 950 nm 后几乎为 0,易于与矸石区别。烟煤的总体特征是:在波段 350~900 nm 反射率变化平缓,在 900 nm 后逐渐升高,在 2 000 nm 后不存在下降趋势,光谱曲线在整个波段变化较为平缓,最大与最小反射率之差小于 10%。建筑物的光谱在 350~750 nm 显著增强,在近红外波段(750~2 500 nm)缓慢增强。与未燃烧矸石的光谱曲线相似,二者不易区分。

3 遥感模型建立与验证

3.1 矽石遥感模型建立

从上述可知, 燃烧矽石和未燃烧矽石的光谱在 350~760 nm 存在反射率斜率的明显差异。据此可以利用 Landsat 8 卫星的 OLI 多光谱数据建立遥感模型进行区分。OLI 成像仪在 350~2 500 nm 之间有 9 个波段(B1—B9), 其中波段 2 的波长为 450~515 nm, 波段 3 的波长为 525~600 nm, 波段 4 的波长为 630~680 nm, 波段 5 的波长为 845~885 nm。除波段 8 的空间分辨率为 15 m 外, 其余的 8 个波段的分辨率为 30 m。将 1—7 和 9 波段标注在图 1 中, 如图 3 所示。

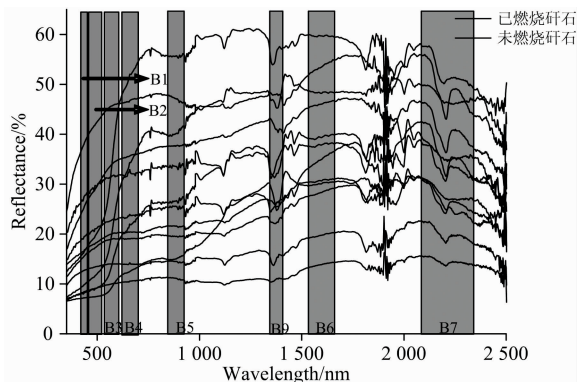


图 3 OLI 八个波段在矽石样本光谱上的位置显示

Fig. 3 The location of eight bands of OLI on the spectral curve of gangue samples

为有效区分两种矽石, 首先构建一种归一化差异矽石指数(normalized difference gangue index, NDGI)

$$NDGI = \frac{B_j - B_i}{B_j + B_i} \quad (1)$$

式(1)中, B_i 与 B_j 分别表示 OLI 的第 i 波段和第 j 波段的反射率。NDGI 的大小反映了光谱曲线在 i 和 j 两个波段反射率的差异。NDGI 越大, 两个波段反射率的差值越大, 光谱曲线在这两个波段间的斜率也越大。

为寻找最优的 NDGI 指数, 分别计算了 2—3 波段、2—4 波段、2—5 波段之间的 NDGI 指数, 对燃烧矽石和未燃烧矽石进行区分。最终发现, 2—5 波段之间的 NDGI 指数对于区分燃烧和未燃烧矽石效果最好。因此, 最终确定铁法矿区归一化差异矽石指数如式(2)

$$NDGI = \frac{B_5 - B_2}{B_5 + B_2} \quad (2)$$

分别计算 106 块燃烧矽石和未燃烧矽石样本的 NDGI 值, 结果如图 4 所示。

从图 4 可以看出燃烧矽石的 NDGI 值都高于未燃烧矽石, 83 个未燃烧矽石中 82 个 NDGI 值小于 0.23, 23 个燃烧矽石的 NDGI 值都高于 0.26。以 0.25 为阈值可以将二者区分开来, 106 个样本的分类准确率为 99.06%。

3.3 与随机森林分类结果对比

随机森林算法作为一种新型高效的分类算法, 是目前数

据挖掘最热门的算法之一, 该算法是 Leo Breiman 在 2001 年提出的一种集成学习算法, 其实质是对决策树算法的一种改进^[12-13]。

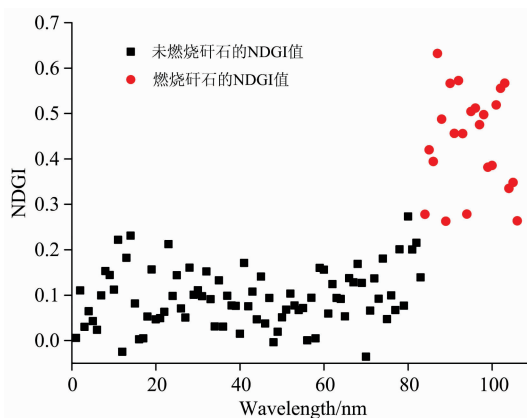


图 4 燃烧矽石和未燃烧矽石的 NDGI 值

Fig. 4 The NDGI values of burned gangue samples and unburned gangue samples

通过数据分析发现, 由光谱仪测试的光谱曲线在 350~2 500 nm 之间有 973 个波段, 各个波段之间存在高度的相关性。因此, 首先采用主成分分析法对波段进行降维。

主成分分析是一种将多指标转换为少数几个综合指标的多元统计分析方法。实验中, 选取前 5 个主成分来表征原来的光谱数据信息, 可代表 99.5% 以上的信息。前 5 个主成分累积贡献率如表 1 所示。

表 1 主成分分析的方差贡献率

Table 1 The variance contribution rate calculated by the principle component analysis

主成分个数	方差贡献率/%	累积贡献率/%
PC1	87.22	87.22
PC2	10.30	97.52
PC3	1.59	99.11
PC4	0.29	99.40
PC5	0.23	99.63

随机森林算法的分类精度取决于参数 N (树的个数) 和 m (特征变量的个数)。Breiman 建议设定随机森林特征变量的个数 m 等于特征变量总数的平方根, 能得到最优化结果。经过主成分分析后特征变量总数为 5, 选择特征变量 m 为 2。树的个数 N 选择 100。

采用十折十次交叉验证的方法来检验其分类精度^[14]。即: 将 106 个样本随机等分成 10 个子样本。依次用其中的 9 个子样本作为训练样本, 剩余的一个子样本作为测试样本。循环完十次后, 所有样本都被预测且仅被预测了一次, 将分类结果的均值作为一次十折的分类精度评价。为了能够更准确地确定其分类精度, 重复上述过程十次, 将十次十折的平均分类精度作为最终分类结果。本文所有算法均由 MatlabR2009(a) 编程实现。实验所用电脑配置为 CPU: Core (TM)-i3, 主频 3.4 GHz, 4.00 GB 内存, Win7 操作系统,

GPU 为 NVIDIA GeForceGT620。该方法的分类结果如表 2 所示。

表 2 随机森林法的分类结果

Table 2 The classification results calculated by the random forest algorithm

检验次数	误样样本数	分类准确率/%
第一次	6	94.3
第二次	5	95.3
第三次	5	95.3
第四次	6	94.3
第五次	5	95.3
第六次	4	96.2
第七次	5	95.3
第八次	6	94.3
第九次	5	95.3
第十次	4	96.2

从表 2 看出，随机森林算法的平均分类准确率为 95.2%，低于 NDGI 指数法的 99.1%。在分类时间上，NDGI 指数法耗时 0.87 ms，而随机森林法耗时 94.8 ms，是 NDGI 指数法的 108.9 倍，表明 NDGI 指数法在大面积矸石分类时存在明显优势。

3.2 模型应用

为了进一步验证模型的实用性，使用 2016 年 8 月 28 日铁法矿区的 Landsat 8 OLI 数据，首先对其进行辐射校正和 FLAASH 大气校正，得到地物的反射率。然后，对矿区内的大兴矿和大隆矸石山区域进行燃烧和未燃烧区域识别划分。

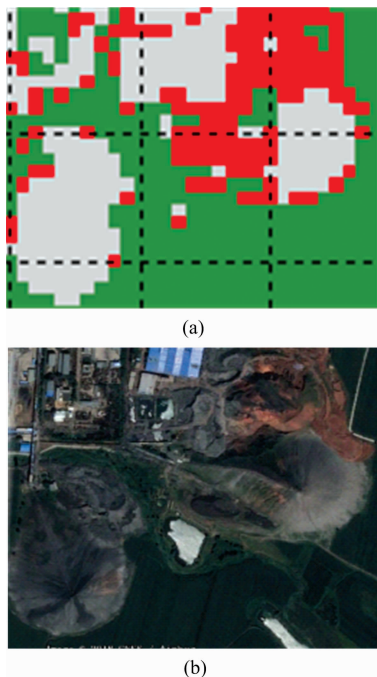


图 5 大兴矿矸石山分类结果与 Google Earth 效果对比图
Fig. 5 The comparison results between the classification on coal mine waste in Daxing mine and the Google Earth

由于影像获取日期为夏季，矿区周围的植被覆盖度较高，选取归一化植被指数 NDVI 并以 0.25 作为阈值将整个矸石山区域划分为植被区和非植被区两部分；然后非植被区主要为矸石山，根据前文中用 NDGI 作为指标并以 0.25 为阈值进行燃烧和未燃烧矸石提取，结果如图 5 和图 6 所示。其中图 5(a)和图 6(a)是两个矿区矸石区域的提取结果，图 5(b)和图 6(b)是两个区域的 Google Earth 高分辨率影像。

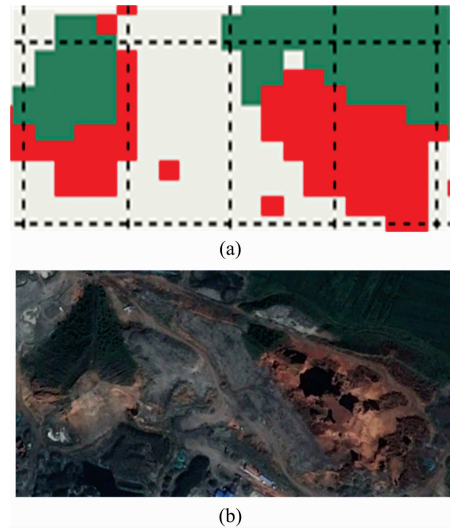


图 6 大隆矿矸石山分类结果 (a) 与 Google Earth 效果 (b) 对比图
Fig. 6 The comparison results between the classification on coal mine waste in Dalong mine (a) and the Google Earth (b)

图 5(a)和图 6(a)的分类结果中绿色代表植被，红色代表燃烧矸石，灰色代表未燃烧矸石。图 5(b)和图 6(b)的 Google Earth 影像中绿色为植被，砖红色为燃烧矸石，灰色代表未燃烧矸石。从结果可以看出，所提取的燃烧矸石和未燃烧矸石在形态和大小上基本与 Google Earth 一致，表明模型具有很好的实用性。图 5 中的少量建筑物误分为未燃烧矸石，主要由于二者光谱特征相似所致(见 2.2 部分)。

3.3 机理分析

图 7 是铁法矿区典型矸石的照片，从照片中可以看出燃烧矸石整体呈现砖红色或土黄色，而未燃烧矸石为灰白色或灰黑色，二者颜色差异较大。

为了进一步了解燃烧矸石和未燃烧矸石的矿物成分差异。分别对图 7 中的样本送与地质服务公司进行矿物鉴定。结果表明，未燃烧矸石中主要包括粉砂级碎屑(石英、长石、云母等)、黏土矿物和铁质等。其中铁质呈黑色、红褐色，成细条带状，有定向性，分布不均匀。燃烧矸石主要包括粉砂级碎屑(主要为石英，少量云母和长石)、玻璃质和铁质三大类。其中玻璃质在单偏光下为红褐色，不显光性。铁质主要为红褐色，呈条带状。

随着矸石的自燃，矿物种类和含量都发生变化。未燃烧矸石中的含铁矿物主要为黏土矿物和黄铁矿，在高温燃烧的

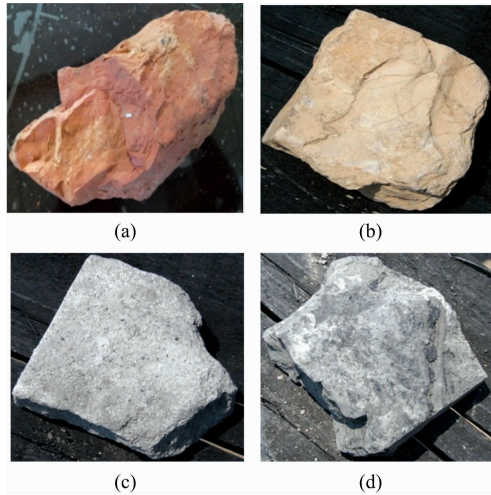


图 7 (a), (b) 为燃烧矽石的照片;
(c), (d) 为未燃烧矽石的照片

Fig. 7 (a), (b) The image of burned gangue samples;
(c), (d) The image of unburned gangue samples

过程中 Fe^{2+} 被氧化为 Fe^{3+} 。 Fe^{3+} 在 550 nm 处会出现一个明显的波谷特征, 这一特征在燃烧矽石光谱中得到了验证。另外, 由于氧化铁的生成使矽石颜色由黑色、灰色变为砖红色

(橙色)和土黄色, 导致在黄色-橙色波段(550~630 nm)反射率显著上升。

另外, 未燃烧矽石中的部分石英和黏土矿物在高温环境下变为新的物相, 主要为玻璃质, 形成了燃烧矽石在 750 nm 处的高反射率。基于上述特征, 构建 NDGI 指数可以很好地将燃烧矽石和未燃烧矽石区分开。

4 结 论

针对燃烧矽石和未燃矽石的可见光-近红外光谱特征差异, 构建了光谱指数 NDGI, 用于燃烧矽石和未燃烧矽石的区分, 得到以下结论:

(1) 燃烧矽石和未燃矽石在可见光波段光谱特征差异明显, 燃烧矽石在 350~760 nm 间光谱曲线斜率整体较高, 在 550~630 nm 反射率存在陡升现象, 而未燃烧矽石则无上述特征。

(2) 基于燃烧和未燃烧矽石的光谱特征差异构建了 NDGI 指数, 用于燃烧和未燃烧矽石的区分。室内实验结果证明, 该指数识别燃烧和未燃烧矽石的分类精度达到了 99.1%, 高于随机森林分类法的 95.2%。现场的 Landsat8 卫星监测结果表明, 基于 NDGI 指数可以很好地将矽石山中的燃烧和未燃烧区域区分开来, 具有很好地实用性。

References

- [1] Pone J D, Hein K A, Stracher G B, et al. International Journal of Coal Geology, 2007, 72: 124.
- [2] Querol X, Izquierdo M, Monfort E, et al. International Journal of Coal Geology, 2008, 75: 93.
- [3] Liu H B, Liu Z L. Resources, Conservation and Recycling, 2010, 54: 1331.
- [4] Fan G W, Zhang D S, Wang X F. Resources, Conservation and Recycling, 2014, 83: 24.
- [5] Yassine Taha, Mostafa Benzaazoua, Rachid Hakkou, et al. Minerals Engineering, 2017, 107: 123.
- [6] Chugh Y P, Patwardhan A. Resources, Conservation and Recycling, 2004, 40: 225.
- [7] Zhang J X, Deng X J, Zhao F J, et al. International Journal of Mining, Eclamation and Environment, 2017, 31(2): 91.
- [8] Frias M, Villa R, Rojas M, et al. Journal of America Ceramic Society, 2012, 95(1): 386.
- [9] Wang Y S, Yang M, Wei G, et al. Sensors and Actuators B: Chemical, 2014, 193: 723.
- [10] Mao Y C, Liu S J, Wu L X, et al. Canadian Journal of Remote Sensing, 2014, 40(5): 327.
- [11] SONG Liang, LIU Shan-jun, YU Mo-li, et al(宋 亮, 刘善军, 虞茉莉, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2017, 37(2): 416.
- [12] Leo Breiman. Machine Learning, 2001, 45(1): 5.
- [13] Daniel R G, Wade A W, Andrew J F, et al. Journal of Terramechanics, 2016, 65: 85.
- [14] Diamantidis N A, Karlis D, Giakoumakis E A. Artificial Intelligence, 2000, 116: 1.

A Classification Method Based on the Visible Spectrum for Burned and Unburned Gangue Distinguishment

SONG Liang, LIU Shan-jun*, MAO Ya-chun, WANG Dong, YU Mo-li

College of Resources and Civil Engineering, Northeastern University, Shenyang 110819, China

Abstract A large number of coal mines are widely distributed over China. Bulk coal gangue deposits seriously affect the mining area environment, and some mishandling of coal gangue may cause spontaneous combustion and explosion, which poses a direct threat to mine safety. The comprehensive utilization of coal gangue can effectively alleviate this problem, and it is of significance to the ecological safety and sustainable development of mine. Depending on the burning state, coal gangue is divided into two types - burned and unburned gangue, whose hidden dangers of security and harm to the environment are different, as well as ways of comprehensive utilization. Therefore, it is very important to do the classification, recognition and monitoring of the coal gangue. The current monitoring methods are mainly the field investigation with low efficiency and high cost, almost impossible for meeting the actual demand of coal gangue monitoring. Tiefa mine in Liaoning Province was chosen as the study area. Firstly, a total of 106 typical coal gangue samples were collected from waste dump in mining areas. Then, SVC HR1024 spectrometer was used to test the visible and near infrared spectrum of samples, and a differential spectral index NDGI was constructed to identify the burned and unburned gangue based on the difference of spectral characteristics of the burned and unburned gangue. Finally, the laboratory spectral data and the corresponding satellite remote sensing images were utilized for verifying the index. The random forest classification method was used as a contrast to the results of the laboratory spectrum treatment. The results showed that the slope of the spectral curves of burned gangue samples was higher ranging from 350 to 750 nm, and the reflectance within range of 550~630 nm increased sharply, while the slope of the unburned gangue in the whole visible bands of spectrum was lower. The threshold of the NDGI index was set as 0.25 to distinguish the burned and unburned gangue. The laboratory spectral data showed that the classification accuracy of the NDGI index is up to 99.1%, higher than that of 95.2% of the random forest classification method. The Field results showed burned and unburned areas of waste dump were distinguished and classified in Landsat8 OLI images based on the NDGI index, and the burned and unburned coal gangue areas were in good agreement with the Google Earth on the morphology and size. The overall results showed that the index can effectively distinguish the combustion states of gangue. In addition, burned and unburned gangue samples were taken for mineral identification respectively. By comparing the changes of mineral species before and after combustion, the cause of spectral difference was analyzed between the burned and unburned gangue. The results showed the oxidation from Fe^{2+} to Fe^{3+} of gangue in the process of combustion. A large increase in Fe^{3+} caused the formation of an obvious spectral valley characteristic at 550 nm band, and a highly reflectance appeared at 750 nm band due to the glass quality generated during combustion. The above conditions cause differences in NDGI index between the burned and unburned gangue. In this paper, the results provide a fast, efficient and accurate model and method for burned and unburned gangue distinguishment in coal mine.

Keywords Remote sensing; Visible and near-infrared; Gangue; Classification

(Received Mar. 15, 2018; accepted Aug. 5, 2018)

* Corresponding author