火成岩中长石含量与其特征光谱间响应关系研究

杨长保1,高文博1*,侯光宇2,李星喆1,高曼婷1

吉林大学地球探测科学与技术学院,吉林长春 130026
 中国人民解放军 32023 部队,辽宁大连 116000

摘 要 长石是地表岩石最重要的造岩矿物,在地壳中的比例高达 60%,几乎是所有火成岩的主要矿物成 分。随着高光谱技术的发展,国内外众多学者研究主要造岩矿物含量与其特征光谱的响应关系,对遥感岩矿 识别以及矿化蚀变信息提取提供了多种可能性。该研究以 USGS 光谱库里 18 个火成岩样本为基础数据,研 究长石的特征光谱及其与含量之间的定量关系。通过原始光谱反射率及其变换(包括小波三层分解高频分 量、小波二层分解、去包络线后光谱、去包络线后小波三层分解高频分量及去包络线后小波二层分解),研 究其与长石的含量之间的相关关系,结果表明:(1)分析六种光谱反射率的变换,去包络线后小波三层分解 高频分量的光谱反射率与长石含量的相关关系最好,且相关系数正负不断变化,根据相关系数极值获得长 石的特征谱带为 431, 570, 972, 1 456, 1 856, 2 292. 9 和 2 481 nm; (2) 原始光谱反射率与长石含量的相关 性曲线趋势较为平缓, 而经小波分解得到的高频分量后, 趋势明显, 经去包络线及小波分解得到高频分量 后,相关性曲线的变化趋势愈加明显,由此可见,自变量的微小变化就会引起因变量变化,当岩石中长石的 含量极小时,小波分解处理能够提高模型的精度。将长石含量与特征光谱的关系量化,运用多元逐步线性回 归分析以及最小二乘法建模,建立6个线性回归模型和6个最小二乘法回归模型,结果表明:(1)去包络线 后的光谱比原始光谱建立的回归模型精度更高,经过小波二层分解后的低频分量建模的回归模型精度优于 未进行小波分解的光谱,其中去包络线后小波二层分解低频分量建立的回归模型效果最佳。(2)多元线性回 归建立的模型精度优于最小二乘法,同时筛选对因变量影响较大的自变量,972,1456,1856,2292.9以及 2 481 nm。因此选择去包络线后的光谱进行多元线性回归法进行分析长石含量与光谱反射率之间关系,考虑 到不同的特征吸收波段对长石含量的影响因子不同,可以利用长石的特征光谱定量反演某一区域内的长石 的含量,对识别矿物具有重要意义。

关键词 火成岩;长石;特征光谱;相关性分析;多元回归分析;小波分析 中图分类号: P585 文献标识码: A DOI: 10. 3964/j. issn. 1000-0593(2019)07-2077-06

引 言

高光谱遥感技术具有丰富的光谱信息,在可见光范围 内,不同的矿物具有不同的光谱特征谱带,利用这些特征谱 带,可以有效的识别岩石中的矿物成分,反演矿物在岩石中 的质量百分比。2012年,樊磊等通过高光谱数据对黑龙江省 北部地区进行岩石特征波段的提取,通过对应分析方法找出 与岩石相关的特征波段,解释其物理意义并将样品中岩石类 型进行分类,提供了岩石与成像光谱数据精细分类的依 据^[1];2015年杨长保等通过高光谱数据研究岩石中黑云母的 含量与光谱的吸收深度之间的关系,并用数学模型的方式定 量分析岩石的光谱吸收深度与矿物含量之间的关系^[2]。借鉴 前人的研究,发现岩类随着造岩矿物的不同,光谱会出现一 些系统的区域性变异变化,但也存在局部的稳定性和规律 性^[3-8],通过研究确定了岩石中主要造岩矿物长石的特征光 谱,并通过建立长石含量与光谱之间的定量关系来为遥感岩 矿识别以及矿化蚀变信息提取提供一种崭新的思路。

1 实验部分

1.1 样本数据

本次研究所用数据是以美国地质调查局提供的标准数字 光谱库为基础数据,选取火成岩样本18块,其中火山岩3块

收稿日期: 2018-12-17,修订日期: 2019-04-25

基金项目:国家自然科学基金项目(41371332),中国地质调查局"一带一路"地质矿产与资源环境遥感解译(DD20160117)资助 作者简介:杨长保,1972年生,吉林大学地球探测科学与技术学院副教授 *通讯联系人 e-mail:905920143@qq.com

(andesite1, andesite2, basalt), 侵入岩 15 块。岩石样本包括 各类矿物的百分含量以及用 ASD 地物光谱仪测量的 0.35~ 2.5 μm 波长范围内的反射率。表 1 是 18 块岩石主要造岩矿 物的百分含量。

表 1 18 个岩石样本主要矿物含量说明 Table 1 Main mineral content description of 18 rock samples

岩石类型	长石 含量/%	石英 含量/%	黑云母 /%	磁铁石 /%	辉石 /%
andesite1	26.75			1.05	
andesite2	30.05			1.65	
basalt	31				
diabase1	48		0.5		
diabase2	46		1.2		
diorite	51		3.4		
gabbro	66				28
granite2	60.1	38.5	1.1	0.3	
granite3	73.7	23.2	1.9	0.9	
granite5	62.5	35.2	1.4	0.8	
granodiorite1	67	19.7	6.3		
granodiorite2	72.3	8.6		4	
monzonite	70	18	1.2		18
norite1	30	2			
norite2	38				
syenite1	90				5
syenite2	71		1.5		
tonalite	53	21	15		

1.2 光谱数据去包络线处理

岩石的光谱测量值包含有背景信息以及所需要的特征光 谱信息,光谱数据的去包络线处理目的是为了抑制背景信 息,从而放大光谱上的吸收特征。"包络线"定义为逐点直线 连接局部波谱的最大值,并使出现在峰值点上的外角大于 180°,再以原始光谱上的值除以包络线上对应的值,即得到 去除包络线的光谱值,如式(1)^[9]

$$R'(\lambda_i) = \frac{R(\lambda_i)}{R_c(\lambda_i)} \tag{1}$$

其中, $R'(\lambda_i)$ 为去包络线后的光谱反射率, $R(\lambda_i)$ 为某波段在 吸收谷处的反射率, $R_{\epsilon}(\lambda_i)$ 为对应波长处包络线上的值。

1.3 光谱数据的小波分解重构

在离散小波变换(DWT)中,在空间 $V_j = V_{J-1} + W_{J-1}$ 上表示信号,也就是对于每一个在 V_j 上表示的信号x(t)能用上面提到的两个空间中的基函数来表示。

$$x(t) = \sum_{k} cA_{0}(k)\phi_{j,k}(t) = \sum_{k} cA_{1}(k)\phi_{j-1,k}(t) + \sum_{i} cD_{1}(k)\omega_{j-1,k}(t)$$
(2)

其中, $A_1(k)$ 和 $D_1(k)$ 是尺度度量空间 j-1 的系数, j-1空 间是由尺度度量空间 j 的系数 $A_0(k)$ 分解而来。类似的, 可 以用 $A_1(k)$ 和 $D_1(k)$ 重构 $A_0(k)$ 。

在尺度度量空间 j 对系数 $A_0(k)$ 进行分解得到在尺度度 量空间 j-1 的两个系数 $A_1(k)$ 和 $D_1(k)$ 通过重构得到系数 $A_0(k)$ 。

$$\xrightarrow{A_0(k)} fightarrow A_1(k) \xrightarrow{A_0(k)} fightarrow A_0(k)$$

如上图的分解与重构,可以通过小波变换算法计算,当小波 与尺度在空间内是正交的,可以用内积计算公式得到系数 *c*A₁(*k*)和 *D*₁(*k*)

$$cA_{1}(k) = \langle x(t), \phi_{j-1,k}(t) \rangle = \langle \sum_{n} cA_{0}(n)_{j,n}(t), \phi_{j-1,k}(t) \rangle = \sum_{n} cA_{0}(n) \langle \phi_{j,n}(t), \phi_{j-1,k}(t) \rangle$$
(3)

下面是内基计算方法的具体计算公式 $\langle \phi_{j,n}(t), \phi_{j-1,k}(t) \rangle = \int_{-\infty}^{+\infty} \sqrt{2^{j}} \phi(2^{j}-n) \sqrt{2^{j-1}} \phi(2^{j-1}t-k) dt$ $= \int_{-\infty}^{+\infty} \sqrt{2^{2j-1}} \phi(2^{j}-n) \sqrt{2^{j-1}} \phi(2^{j-1}t-k) dt$ (substitutes = $2^{j-1}-k) = \int_{-\infty}^{+\infty} \sqrt{2} \phi(2s+sk-n) \phi(s) ds$ (Use the 2 - scale equation for $\phi(s) = \int_{-\infty}^{+\infty} \sqrt{2} \phi(2s+sk-n) \sum_{m} h_{0}(m) \sqrt{2} \phi(2s-m) ds$ $m) ds = \sum_{m} h_{0}(m) \int_{-\infty}^{+\infty} \phi(2s+sk-n) \phi(2s-m) 2ds$ (integral is 0 unless m = n - 2k) = $h_{0}(n - 2k)$ (4) 具体系数计算过程如下

$$cA_{1}(k) = \sum_{n} h_{0}(n-2k)cA_{0}(n)$$

$$cD_{1}(k) = \langle x(t), \omega_{j-1,k}(t) \rangle =$$

$$\langle \sum_{n} cA_{0}(n)\phi_{j,n}(t), \omega_{j-1,k}(t) \rangle =$$

$$\sum_{n} cA_{0}(n)\langle \phi_{j,n}(t), \omega_{j-1,k}(t) \rangle$$
(5)
$$cD_{1}(k) = \sum h_{1}(n-2k)cA_{1}(n)$$
(6)

在 MATLAB 中实现小波分解重构的算法处理。具体操 作为利用小波分解进行多层分解,得到细节特征有效的各光 谱曲线。依据光谱曲线,可得原始光谱中的噪声以及微小的 吸收特征是影响高频系数的主要因素。利用 sym5 小波基对 原始光谱曲线进行三层小波分解得出其低频分量以及高频分 量。图 1 和图 2 分别是原始光谱曲线和去包络线后进行小波 分解重构的结果。图 1 是某个样本的原始光谱反射率进行小 波分解重构后的结果。(a)是样本的原始光谱曲线,(b)为样 本光谱曲线小波分解低频分量光谱,(c)为样本光谱曲线小 波分解高频分量光谱,(d)为样本光谱曲线小波分解重构后 保留的特征光谱。图 2 为样本反射率光谱进行包络线去除后 进行小波分解重构的结果。其中(a)是某样本去包络线后的 光谱曲线,(b)是某样本包络线去除后小波分解低频分量光 谱,(c)是样本包络线去除后小波分解高频分量光谱,(d)是 样本包络线去除后小波分解重构后保留的特征光谱。

2 结果与讨论

在光谱库中筛选出 18 个样本,对样本光谱进行去包络 线,小波分解等预处理,并运用多元逐步回归法进行建模。



图 1 原始光谱曲线小波分解结果





Fig. 2 De-envelope curve de-spectrum curve wavelet decomposition results

2.1 光谱数据的相关性分析

将岩石中长石的含量与光谱的反射率,小波分解高频波 段,小波二层分解,去包络线,去包络线后小波分解高频分 量以及去包络线后小波二层分解后的光谱反射率分别做双变 量相关性分析,随着波长的变化,长石的含量与各种算法下 的光谱反射率的相关性系数也不同,相关系数曲线趋势图如 图 3 和图 4 所示。从表 2 可以看出,岩石中长石的含量与岩 石的反射率的相关系数为 0.407,呈正相关关系。但对原始 光谱进行小波二层分解后,提高了相关系数为 0.595,同时 岩石中长石含量与岩石的反射率相关系数呈正负不断变化, 这能有效的区分不同波段对长石含量的影响。这为后面建立 回归方程所需的特征波段提供理论依据。





Fig. 3 Correlation coefficient of the original spectrum and its wavelet decomposition



following its wavelet decomposition

综合图 3 和图 4 可以看出,对原始光谱去包络线后小波 分解高频的光谱与长石含量的相关性最好,相关系数范围为 [-0.7,0.7],其次是原始光谱经小波分解得到高频分量后 的光谱,其与长石含量的相关性范围是[-0.7,0.6]。研究 发现经小波分解得到高频分量后,光谱反射率与长石含量相 关性系数曲线趋势有较为明显的变化,经去包络线及小波分 解得到高频分量后,相关性系数曲线的变化趋势愈加明显, 这说明:自变量的微小变化就会引起因变量变化,当岩石中 长石的含量极小时,通过小波分解处理能够提高模型的精 度。由表 2 以及图 3 和图 4 等相关系数趋势图可知,当波长 处于 431,570,972,1 456,1 856,2 292.9 以及 2 481 nm 处 相关系数达到极显著水平,说明在这 7 个波段处,长石的含 量与光谱反射率之间的相关关系最好,故岩石光谱在这7处 存在特征吸收波段,可见火成岩中长石的含量与岩石光谱之 间的关系复杂,具有多个特征吸收谱带。

表 2 长石含量与最大相关波段分析 Table 2 Analysis of feldspar content and maximum correlation band

变换形式	最大正 相关波 长/nm	相关 系数	最大负 相关波 长/nm	相关 系数
原始光谱	944	0.407		
小波分解高频	$2 \ 214$	0.595	591	-0.667
小波二层分解	944	0.406		
去包络线	972	0.638	431	-0.402
去包络线后小波分解高频	2 269	0.683	2 159	-0.714
去包络线后小波二层分解	968	0.637	433	-0.347

2.2 长石含量与光谱反射率的回归方程建立

将以上分析结果中的波段作为自变量,并设为 $X_1 - X_6$, 长石的含量作为因变量进行回归分析,选用 enter 法和两阶 最小二乘法建立回归方程,并使用判断系数 R²、调整后的判 断系数 R² 以及 F 统计量来判断模型的稳定性。分别用原始 的反射率光谱,小波分解后高频分量,小波二层分解,去包 络线,小波分解后的高频分量以及去包络线后的小波二层分 解作为自变量,长石的含量作为因变量,进行多元回归分 析,得到12个方程。通过对表3和表4所列的12个方程分 析可得,(1)线性回归分析模型的精度要优于最小二乘法回 归模型,在线性回归分析模型中,去包络线小波重构分解得 到方程的 F 统计量最大,为 43.827,且调整后的 R^2 最大为 0.922;在最小二乘法回归模型中,原始光谱小波分解得到 的方程中, F 统计量最大为 16.623, 调整后的 R^2 统计量最 大为 0.859; (2)在线性回归分析中,经过小波二层分解得到 的方程的精度比小波三层分解高频得到的方程精度高,可见 在小波分解中, 三层分解的高频信号携带大部分噪声, 不建 议使用小波分三层分解进行反演;(3)在线性回归分析中, 去包络线光谱与去包络线光谱小波二层分解得到的方程能够 自动筛选对因变量作用明显的波段,并将对因变量影响微小 的波段予以剔除,其中 972,1 456,1 856,2 292.9 以及 2 481 nm 这些波段是对长石含量影响最大; (4)最小二乘法 回归分析中,原始光谱的小波三级分解、去包络线光谱、去 包络线光谱的小波三级分解以及去包络线光谱的小波二级分 解的 F 统计量均小于 2.9, 调整后的 R^2 统计量均小于 0.5, 不具有明显的统计意义,故在反演模型中将其剔除。

3 结 论

在研究岩石中长石的含量与光谱反射率之间的关系时发现,去包络线处理后的光谱反射率与长石的相关性比未经过 反射率处理的相关性强;经过小波分解后,相关性均有显著 提高,但去包络线后再小波分解得到的相关系数明显更高, 范围为[-0.7,0.7],据此结合相关系数趋势图,得到在

Table 3 Multivariate linear regression analysis equations for feldspar content

光谱变形	方程	R^2	调整后 R ²	F 统计量
原始光谱	$y = 496.334x_1 - 447.278x_2 + 0.128x_3 + 261.325x_4 + 162.841x_5 + 264.489x_6 - 496.175x_7$	0.908	0.849	15.45
原始光谱小波 分解-高频	$y=54.436-2325.655x_1+4244.450x_2+748.242x_3-3055.676x_4+$ 1704.493 $x_5+13227.795x_6+19115.041x_7$	0.616	0.348	2.294
原始光谱 小波分解重构	$y = 434.\ 200x_1 - 336.\ 361x_2 - 53.\ 057x_3 + 230.\ 569x_4 + 268.\ 837x_5 + 161.\ 587x_6 - 469.\ 764x_7$	0.909	0.851	15.700
去包络线光谱	$y = 100.470x_3 - 101.298x_4 - 13.26x_5 + 8.584x_6 + 70.506x_7$	0.940	0.920	42.194
去包络线小波 分解-高频	$y = -251.046x_1 + 867.661x_2 - 5 340.435x_3 - 5 216.931x_4 - 2 862.081x_5 - 150.942x_6 - 20 386.709x_7$	0.608	0.358	2.433
去包络线小波 分解重构	$y = 96.881x_3 - 76.941x_4 - 18.617x_5 + 57.338x_6 + 7.282x_7$	0.944	0.922	43.827

表 4 长石含量最小二乘法回归方程

Table 5 The least square reg	gression equ	lation for t	eldspar	content
------------------------------	--------------	--------------	---------	---------

光谱变形	方程	R^2	调整后 R ²	F 统计量
原始光谱	$y = 496.334x_1 - 447.278x_2 + 0.128x_3 + 261.325x_4 + 162.841x_5 + 264.489x_6 - 496.175x_7$	0.912	0.857	16.359
原始光谱小波 分解-高频	$y=54.436-2325.655x_1+4244.450x_2+748.242x_3-3055.676x_4+$ 1704.493 $x_5+13227.795x_6+19115.041x_7$	0.616	0.348	2.294
原始光谱小波 分解重构	$y = 434.\ 200x_1 - 336.\ 361x_2 - 53.\ 057x_3 + 230.\ 569x_4 + 268.\ 837x_5 + 161.\ 587x_6 - 469.\ 764x_7$	0.914	0.859	16.623
去包络线光谱	$y = -1 \ 080.\ 859 \ -369.\ 728x_1 \ +145.\ 034x_2 \ +41.\ 107x_3 \ -71.\ 995x_4 \ +\\35.\ 013x_5 \ +8.\ 545x_6 \ +1\ 357.\ 777x_7$	0.460	0.082	1.217
去包络线小波 分解-高频	$y = 54.060 - 477.096x_1 + 771.844x_2 + 1 304.994x_3 - 1 067.964x_4 + 180.462x_5 + 4 832.981x_6 + 7 238.556x_7$	0.668	0.453	2.870
去包络线小波 分解重构	$y=4$ 750. 501 - 781. 908 x_1 + 124. 377 x_2 + 73. 142 x_3 - 109. 824 x_4 + 86. 74 x_5 + 199. 134 x_6 - 4 293. 171 x_7	0.570	0.269	1.893

431,570,972,1456,1856,2292.9以及2481 nm 处长石 的含量与光谱反射率的相关性系数达到极值,岩石在这七处 有特征光谱。

研究长石含量与长石特征光谱之间的回归分析时,运用 多元逐步线性回归分析以及最小二乘法建模,建模精度通过 调整后的判定系数(Adjusted R²)和F统计量的大小判断,采 用 enter 线性回归分析以及最小二乘法得到12个回归方程, 分析可得如下结论:(1)在线性回归分析中小波二级分解比 小波三级分解高频信号建立的模型精度高,F统计量大,且 调整后的 R²大,可见小波三级分解高频信号携带大量噪声, 不建议使用;(2)在线性回归分析中,去包络线光谱与去包 络线光谱小波二层分解得到的方程能够自动筛选对因变量作 用明显的波段,并将对因变量影响微小的波段予以剔除; (3)线性回归分析所建立的模型精度要优于最小二乘法所建 立的模型精度;因此选择线性回归分析的方法,对去包络线 后在小波二级分解的光谱建立的模型,是精度最优解。将长 石的含量与光谱的反射率进行线性回归,结合不同特征波段 的影响权重可以利用长石的特征光谱定量反演某区域内长石 的含量并能有效识别矿物成分。

References

- [1] FAN Lei, ZHAO Wen-ji, GONG Zhao-ning, et al(樊 磊,赵文吉,宫兆宁,等). Journal of Jilin University・Earth Science Edition(吉 林大学学报・地球科学版), 2012, 42(2): 575.
- [2] YANG Chang-bao, ZHANG Chen-xi, LIU Fang, et al(杨长保,张晨曦,刘 舫,等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2015, 35(9): 2583.
- [3] XU Jing, JIANG Ping-an, WU Hong-qi(胥 静, 蒋平安, 武红旗). Hubei Agricultural Science(湖北农业科学), 2015, 54(1): 43.
- [4] SUN Ya-qin, TIAN Shu-fang, WANG Xing-zhen, et al(孙娅琴,田淑芳,王兴振,等). Geoscience(现代地质), 2016, 30(1): 239.
- [5] QIN Kai, CHEN Jian-ping, ZHAO Ying-jun, et al(秦 凯,陈建平,赵英俊,等). Uranium Geology(铀矿地质), 2018, 34(3): 174.
- [6] GUO Bang-jie, ZHANG Jie-lin, WU Ding(郭帮杰,张杰林,武 鼎). Science Technology and Engineering(科学技术与工程), 2018, 18

(17): 125.

- [7] YIN Zi-yao, LIU Tang, WANG Zhen(殷子瑶,刘 唐,王 震). Beijing Surveying and Mapping(北京测绘), 2018, 32(7): 788.
- [8] Gomez C, Adeline K, Bacha S, et al. Remote Sensing of Environment, 2018, 204: 18.
- [9] Elsy Ibrahim, Pierre Barnabé, Erick Ramanaidou, et al. International Journal of Applied Earth Observations and Geoinformation, 2018, 73: 653.

Response Relationship between Feldspar Content and Characteristic Spectra in Igneous Rocks

YANG Chang-bao1, GAO Wen-bo1*, HOU Guang-yu2, LI Xing-zhe1, GAO Man-ting1

1. College of Earth Exploration Science and Technology, Jilin University, Changchun 130026, China

2. Unit 32023 of the People's Liberation Army of China, Dalian 116000, China

Abstract Feldspar is the most important rock-forming mineral in surface rocks, accounting for up to 60% in the earth's crust. With the development of high-spectral technology, many scholars at home and abroad have been studying the response of the main building of rock and mineral content and the spectrum of the features, and it offers a variety of possibilities for remote sensing, mineralization, and mineralization. It is based on data of 18 igneous samples in the USGS spectral library to study the quantitative relation between the characteristic spectra and the content of the feldspar. Through the original spectral reflectivity and the transformation (including the three layers of the small wave to break down the high frequencies, the little wave layers, the spectrum of the wave to the back of the line, and then the little wave layer after the cable, and then the little wave layer after the cable, and then the small wave layer of the wave, and then the second layer of the cable) study the correlation between the high and the long rocks, and it turns out: (1) to analyze the transformation of six spectral reflectivity, the relation between the spectral reflectivity of the high frequency and the long rock content of the small wave three layers of the envelope, and the correlation coefficients are the best, and the correlation coefficients are constantly changing, and based on the relative coefficients of the relative coefficients, the high value of the long rock is 4, 31, 570, 972, 1 456, 1 856, 2 292. 9, 2 481 nm; (2) the correlation between the original spectral reflectance and feldspar content curve trend is relatively flat, and after the wavelet decomposed high frequency component, through to the envelope and wavelet decomposition for the high frequency component, the change trend of correlation curve is becoming ever more obvious, therefore, the independent variable of the small changes will cause changes in the dependent variable, very hour when the content of the rock feldspar, wavelet decomposition processing can improve the accuracy of the model. The relation between the content of feldspar and the characteristic spectrum is quantified, using a multi-element stepwise linear regression analysis and a least-square method of modeling, establishing six linear regression models and six least-squares regression models, and the results show that: (1) the spectrum of the spectrum behind the envelope is more accurate than the original spectrum, and the low-frequency portion of the lower wave part of the wave is better than the one that's not done with the small wave decomposition, which is the best way to get to the back model of the low-frequency portion of the small wave after the envelope. (2) the multilinear regression model is better than the least-squares, and the variables that affect the variables that affect the larger variables, 972, 1 456, 1 856, 2 292. 9 and 2 481 nm. In view of that relation between the content of feldspar and the spectral reflectance of the long stone and the influence factors on the content of feldspar in different absorption band, it is important to use the characteristic spectrum of feldspar to quantitatively invert the content of feldspar in a region.

Keywords Igneous rock; feldspar; Characteristic spectrum; Correlation analysis; Multiple regression analysis; wavelet analysis

(Received Dec. 17, 2018; accepted Apr. 25, 2019)

* Corresponding author