



结合多尺度分割和随机森林的变质矿物提取

唐淑兰 孟勇 王国强 卜涛

Extraction of metamorphic minerals by multiscale segmentation combined with random forest

TANG Shu-lan, MENG Yong, WANG Guo-qiang, BU Tao

引用本文:

唐淑兰, 孟勇, 王国强, 卜涛. 结合多尺度分割和随机森林的变质矿物提取[J]. *工程科学学报*, 2022, 44(2): 170–179. doi: 10.13374/j.issn2095-9389.2020.09.08.004

TANG Shu-lan, MENG Yong, WANG Guo-qiang, BU Tao. Extraction of metamorphic minerals by multiscale segmentation combined with random forest[J]. *Chinese Journal of Engineering*, 2022, 44(2): 170–179. doi: 10.13374/j.issn2095-9389.2020.09.08.004

在线阅读 View online: <https://doi.org/10.13374/j.issn2095-9389.2020.09.08.004>

您可能感兴趣的其他文章

Articles you may be interested in

基于监督学习的前列腺MR/TRUS图像分割和配准方法

Prostate MR/TRUS image segmentation and registration methods based on supervised learning
工程科学学报. 2020, 42(10): 1362 <https://doi.org/10.13374/j.issn2095-9389.2019.10.10.006>

基于改进CV模型的金相图像分割

Segmentation of metallographic images based on improved CV model
工程科学学报. 2017, 39(12): 1866 <https://doi.org/10.13374/j.issn2095-9389.2017.12.013>

基于自动多种子区域生长的遥感影像面向对象分割方法

Object-oriented remote sensing image segmentation based on automatic multiseed region growing algorithm
工程科学学报. 2017, 39(11): 1735 <https://doi.org/10.13374/j.issn2095-9389.2017.11.017>

一种基于鲁棒随机向量函数链接网络的磨矿粒度集成建模方法

Grinding process particle size modeling method using robust RVFLN-based ensemble learning
工程科学学报. 2019, 41(1): 67 <https://doi.org/10.13374/j.issn2095-9389.2019.01.007>

基于周期势系统随机共振的轴承故障诊断

Bearing fault diagnosis by stochastic resonance method in periodical potential system
工程科学学报. 2018, 40(8): 989 <https://doi.org/10.13374/j.issn2095-9389.2018.08.013>

陡脉冲干扰下的心电信号滤波及QRS提取

ECG filtering and QRS extraction under steep pulse interference
工程科学学报. 2020, 42(5): 654 <https://doi.org/10.13374/j.issn2095-9389.2019.06.20.004>

结合多尺度分割和随机森林的变质矿物提取

唐淑兰¹⁾, 孟 勇²⁾✉, 王国强²⁾, 卜 涛²⁾

1) 西安财经大学管理学院, 西安 710100 2) 中国地质调查局西安地质调查中心, 西安 710054
✉通信作者, E-mail: 16392800@qq.com

摘 要 为提高遥感影像变质矿物提取精度, 提升变质带的识别效果, 以甘肃北山 ASTER 影像为研究区, 结合了比值运算、多尺度分割、随机森林分类法进行变质矿物提取。首先, 通过矿物特征性光谱特征构造比值运算公式、进行影像增强; 然后, 对增强影像进行基于光谱及变差函数的多尺度分割; 接着, 采用随机森林法提取目标矿物; 最后, 通过野外勘查、采样、薄片鉴定进行精度评价。结果表明, 黑云母、白云母、角闪石在 ASTER 影像上具有鉴定性特征, 提取精度分别为 85.4088%、84.7640% 和 85.7308%; 其他含量较少的变质矿物提取精度可达到 60% 以上。多尺度分割能充分利用矿物的丛集特征; 变差函数纹理能增强形态特征对矿物的区分能力; 随机森林分类法对矿物混合引起的噪声不敏感、提取结果稳定。

关键词 变差函数; 多尺度分割; ASTER; 矿物提取; 随机森林

分类号

Extraction of metamorphic minerals by multiscale segmentation combined with random forest

TANG Shu-lan¹⁾, MENG Yong²⁾✉, WANG Guo-qiang²⁾, BU Tao²⁾

1) School of Management, Xi'an University of Finance and Economics, Xi'an 710100, China
2) Xi'an Center of Geological Survey, China Geological Survey, Xi'an 710054, China

✉ Corresponding author, E-mail: 16392800@qq.com

ABSTRACT The identification of metamorphic minerals is the basis of metamorphic rock research. Extraction of mineral information by remote sensing technology has been widely used. Digital image processing technology is also effectively applied to remote sensing image processing. Results show that the band ratio of remote sensing images can enhance mineral information, while the variogram function can describe the spatial correlation and variability of image pixels and extract more detailed texture information. The metamorphic minerals are found to present a block or strip distribution. The object-oriented remote sensing image information extraction method can avoid the “salt and pepper phenomenon” based on pixel extraction. Meanwhile, the random forest classification method has a fast calculation speed and high parameter accuracy. It is not sensitive to the noise caused by more lithologic components and its classification effect is found to be stable. To improve the extraction accuracy of metamorphic minerals from remote sensing images and further improve the recognition effect of metamorphic zones, this paper combined the ratio operation, multiscale segmentation, and random forest classification to extract metamorphic mineral information from ASTER images in Beishan area in Gansu Province. Initially, the image was enhanced by the ratio formula of the characteristic spectral structure of the target mineral. Multiscale image segmentation was then performed based on the spectrum and variogram. Finally, the accuracy was evaluated by the thin film identification results of the field exploration samples after the extraction of the target mineral by random forest. Results show that biotite,

收稿日期: 2020-09-08

基金项目: 中国地质调查局资助项目 (DD20179403, DD20190364, DD20190812); 西安财经大学科学研究扶持计划资助项目 (21FCJH008); 陕西省自然科学基金基础研究计划资助项目 (2020JM-585)

muscovite, and amphibole have identification characteristics on the ASTER image with an extraction accuracy of 85.4088%, 84.7640%, and 85.7308%, respectively. The extraction accuracy of other metamorphic minerals with less content are found to reach more than 60%. Multiscale segmentation can make full use of the clustering features of minerals and the variogram texture can enhance the ability of morphological features to distinguish the minerals. Random forest is not sensitive to noise and the extraction results are observed to be stable.

KEY WORDS variogram; multiscale segmentation; ASTER; mineral extraction; random forest

变质矿物的识别是变质岩研究的基础, 造山带中的变质岩是造山带不同演化阶段动力学过程的直接记录, 其中, 中-低压变质带的分布是研究造山过程和地壳演化等地球科学问题的重要参考依据^[1].

遥感技术作为对传统实地调查的补充, 已广泛应用于矿产勘查和岩性识别中^[2-3]. 一些矿物提取方法得到了快速发展: 如比值法、主成分分析 (Principal component analysis, PCA)、支持向量机 (Support vector machine, SVM)、光谱角 (Spectral angle mapping, SAM)、混合像元分解、随机森林 (Random forest, RF) 等^[4-8]. 同时, 图像处理技术也被结合到遥感矿物提取中: 如多尺度分割、小波变换等^[9-10]. 具有较高分辨率的 ASTER 传感器通过 14 个波段的波谱数据, 提供了更加精细的矿物及岩石信息.

矿物信息在遥感影像上表现为弱信息, 利用反射带和吸收带的比值, 可增强各地质信息之间的光谱差异. 可根据各矿物的特征性光谱特征, 选择合适的 ASTER 波段进行比值增强. 基于高分辨率遥感影像的光谱特征提取矿物信息, 会出现“椒盐现象”^[11]; 而遥感影像上变质矿物呈现块状或条带状分布, 可使用面向对象思想, 通过多尺度分割得到相似的对象块, 利用影像的光谱、形状、大小、色调、纹理等特征^[12], 提高矿物提取精度. 为了使遥感影像矿物提取更加精确, 各种非参数监督分类法得到了发展, 应用较为广泛的包括神经网络 (Neural network, NN)、SVM 和 RF^[13-14]. RF 是一个集成分类器, 将复杂的分类过程分解为一系列决策过程 (树). 由于计算速度快、参数要求少、对训练数据的统计假设少、对噪声或过拟合的敏感性较低, RF 分类器在遥感信息提取中得到了越来越多的关注^[15]. 但是, 由于部分矿物的光谱相似性、变质矿物与主要造岩矿物的混合性, 仅基于光谱特征提取变质矿物精度有限.

本文以 ASTER 数据为研究对象, 利用波段比值增强各变质矿物的光谱特征; 接着, 结合光谱特征和变差函数纹理进行多尺度分割; 然后, 利用 RF 提取变质矿物; 最后, 通过野外调查验证精度.

本研究旨在提高变质矿物提取精度, 为变质带的遥感解译提供技术参考.

1 地质背景

研究区 (图 1) 位于甘肃省玉门市地区, 大地构造隶属于北山造山带. 出露的主要地层有: 晚太古界、长城系、蓟县系、石炭系、二叠系、侏罗系、白垩系和第四系. 发育有石炭纪和二叠纪中酸性侵入岩. 其中, 晚太古界主要岩性组合为: (云母) 石英片岩、(石榴石、黑云母、白云母) 斜长片麻岩、(石榴石、黑云母) 斜长角闪片麻岩、二长片麻岩、斜长变粒岩、钠长绿帘绿泥片岩和大理岩等; 长城系主要岩性为: 大理岩、石英岩、长石石英岩和绿泥石石英千枚岩等; 蓟县系主要为大理岩, 局部见蛇纹石、透辉石和阳起石化; 石炭系和二叠系主要为浅变质碎屑岩, 局部可见绿泥石化, 石炭系碎屑岩中夹少量灰岩; 侏罗系和白垩系主要为未变质的陆源碎屑岩沉积.

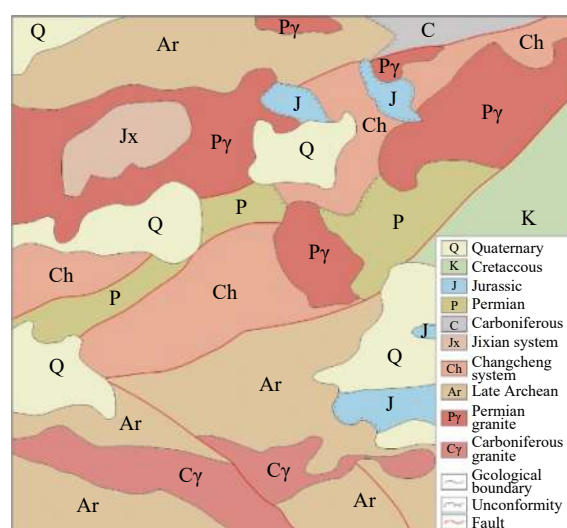


图 1 研究区地质简图

Fig.1 Geological sketch of the study area

2 数据与方法

2.1 数据

2.1.1 ASTER 数据

ASTER 具有 14 个波段的波谱数据, 近红外及可

见光 3 个波段, 可提取铁及稀土元素; 短波红外 6 个波段, 可提取含羟基及碳酸盐蚀变信息; 热红外 5 个波段, 可识别石榴石、黑云母及长石等. 本文选择甘肃中盐池地区 ASTER 数据进行变质矿物提取, 图像为拍摄于 2003-8-15 日的 L1T 级数据, 图像清晰, 无积雪、植被覆盖, 有云覆盖. 数据预处理包括串扰校正、辐射定标、大气定标及去云处理, 大气定标采用 FLAASH 模块进行, 去云处理是鉴于近红外波段的异常高值即为云覆盖, 进行掩膜运算.

2.1.2 变质矿物光谱特征

研究区需提取的标志性矿物为: 黑云母(Bi)、白云母(Mus)、角闪石(Am)、绿泥石(Chl)、石榴石(Gt)、阳起石(Act). 根据野外岩石样品光谱测试数据(图 2), 获得各变质矿物的吸收谱带与 ASTER 波段的对应关系(表 1).

2.2 方法

因矿物具有丛集特征, 本文基于面向对象思想, 采用多尺度分割和 RF 分类法提取变质矿物. 对预处理之后的影像通过比值运算, 得到各矿物增强影像; 选取矿物特征, 构造分类特征向量; 利用 RF 筛选特征并提取矿物; 野外采样、薄片鉴定, 通过鉴定结果对矿物提取结果进行精度评价. 技术流程如图 3 所示.

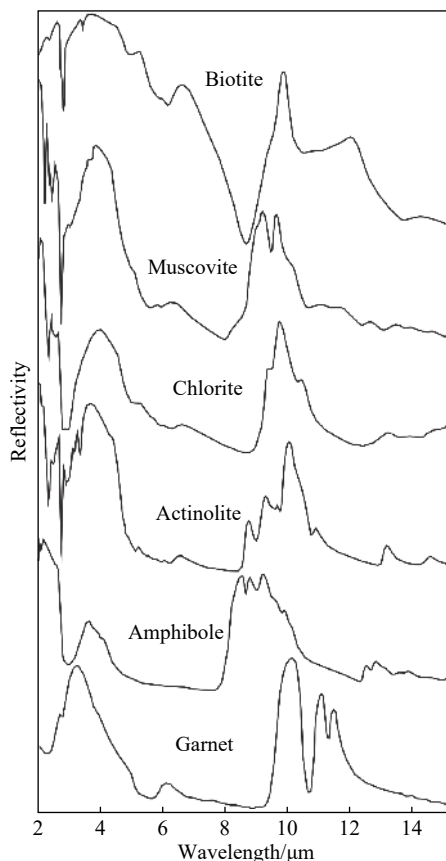


图 2 标志性矿物反射率曲线

Fig.2 Reflectance curve of marker minerals

表 1 矿物的吸收谱带与 ASTER 波段的对应关系

Table 1 Correspondence relation between the absorption bands of the minerals and ASTER bands

Mineral	ASTER band										
	1	5	6	7	8	9	10	11	12	13	
Bi							Reflex	Absorption		Reflex	
Mus	Absorption	Reflex	Absorption	Reflex							
Am				Reflex	Absorption	Reflex					
Chl	Absorption	Reflex			Absorption	Reflex					
Gt									Absorption		Reflex
Act			Reflex		Absorption	Reflex					

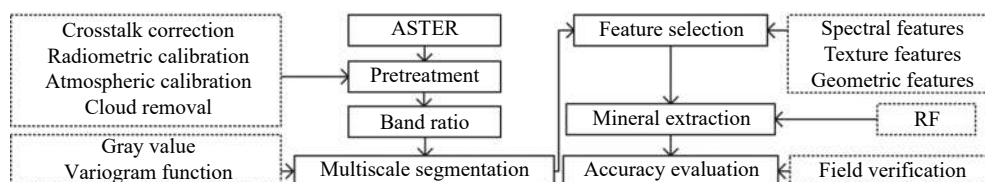


图 3 技术流程

Fig.3 Technical process

2.2.1 比值运算

黑云母在 8.6 μm 处(波段 11)强吸收, 8.0 μm (波段 10)和 9.0 μm 处(波段 12)强反射. 白云母在波段 1 和波段 6 强吸收, 波段 5 和波段 7 强反射.

角闪石在波段 7、波段 9 处强反射, 而在波段 8 强吸收. 绿泥石在波段 1 和波段 8 处有较强吸收谷, 在波段 5 和波段 9 有强反射. 石榴石在热红外 9.25 μm 处(波段 12)有强吸收, 10.21 μm 处(波段

13)有强反射. 阳起石在 2.3 μm 处(波段 8)有吸收谷, 2.2 μm (波段 6)、2.35 μm (波段 9)处有反射峰. 增强各变质矿物信息的波段比值公式见表 2, 其中 b_i 表示 ASTER 影像第 i 波段的光谱反射率.

表 2 各矿物比值公式

Table 2 Ratio formula of minerals

Bi	Mus	Am	Chl	Gt	Act
$(b_{12}+b_{10})/b_{11}$	$(b_5+b_7)/b_6$	$(b_6+b_9)/(b_8+b_7)$	$(b_1+b_9)/b_8$	b_{13}/b_{12}	$(b_6+b_9)/b_8$

2.2.2 多尺度分割

比值法增强的影像有很多伪信息, 需进一步处理. 因矿物信息多呈块状和条带状, 面向对象的分析法可利用高分辨率影像相邻像素之间的关系^[16]. 面向对象特征提取的基础是影像分割, 其目的是将影像分割成若干均匀的区域, 分割性能至关重要. 分割算法需指定尺度参数来控制对象的大小, 多尺度策略较单尺度参数更能满足分割的需要. 较为有效的方法是从多尺度分割结果中选择全局最优尺度, 然后改进欠分割和过分割区域, 最后结合最优分割结果与改进后的分割区域. 本文先给定一个尺度参数进行过分割, 然后使用区域合并细化分割结果, 合并的约束条件为变差函数. 比值影像为单波段影像. 进行初分割时区域合并的依据为光谱相似性, $D(X, Y)$ 为两个相邻区域 X 和 Y 的平均灰度的欧式距离式(1), n 、 m 分别为 X 、 Y 区域的像素总数.

$$D(X, Y) = \frac{1}{n} \sum_{i=1}^n x_i - \frac{1}{m} \sum_{j=1}^m y_j \quad (1)$$

计算每一个初分割对象块的变差函数, 得到变差函数矩阵. 变差函数可描述影像像素的空间相关性和变异性, 提取更详细纹理信息, 有 3 个参数: 方向 θ 、步长 d 、窗口大小 α ^[17], 可表示为:

$$\gamma(\theta, d) = \frac{1}{2N(\theta, d)} \sum_{i=1}^{N(\theta, d)} [F(y_i) - F(y_i + d)]^2 \quad (2)$$

式中, $\gamma(\theta, d)$ 为像元在方向 θ , 步长为 d 的变差函数; $N(\theta, d)$ 为影像在 θ 方向的步长为 d 的像素对的数量, θ 取 0° 、 45° 、 90° 、 135° 4 个方向; $F(y_i)$ 和 $F(y_i + d)$ 分别为 y_i 和 $y_i + d$ 的灰度值. 像素灰度值具有空间自相关性, 自相关性随距离增大而变小, 所以变差函数的步长 d 不应超过窗口大小 α 的一半. 本文各对象块为不规则形状, 不使用传统的 $\alpha \times \alpha$ 移动窗口, 而是以各对象块覆盖范围作为各自的计算窗口. 步长 d 的上限设为不规则对象块各方

向像素数(α)的一半, 即去掉小数部分保留整数部分(用 $|\alpha|$ 表示). 因不规则对象块的各方向变差函数向量维数不同, 取各方向不同步长变差函数的累计平均值:

$$\gamma(\theta) = \frac{1}{|\alpha|/2} \sum_{i=1}^{|\alpha|/2} \gamma(\theta, i) \quad (3)$$

将各方向变差函数累计平均值顺序连接得到最终纹理特征向量:

$$T = t(\gamma(0^\circ), \gamma(45^\circ), \gamma(90^\circ), \gamma(135^\circ)) \quad (4)$$

设两个相邻对象块纹理向量分别为 γ_a 和 γ_b , 通过卡方距离式(5)来评价两块的相似度, 当距离小于设定值时合并这两块. 其中, k 为变差函数向量的维度, 这里 $k=4$.

$$\chi^2(\gamma_a, \gamma_b) = \sum_{i=1}^k \frac{(\gamma_a^i - \gamma_b^i)^2}{(\gamma_a^i + \gamma_b^i)} \quad (5)$$

通过分割指数(Global score, GS)选择最优多尺度分割结果, GS 指数是衡量对象块内部同质性和块间异性的参数, 该值越小说明分割效果越好.

2.2.3 随机森林(RF)

RF 通过集成多个弱分类器(树), 采用平均或投票法得到最终分类结果, 精度和泛化能力较高, 擅于处理高维数据. 具体过程: 每颗树采用有放回的抽样(Bagging)随机选择原始样本集的子样本集, 利用分类和回归树(Classification and regression trees, CART)算法训练二叉决策树, 构建弱分类器, 对各分类器的分类结果采用多数投票法输出结果. 原始样本集的 2/3 用于分类, 1/3 用于验证. RF 有 2 个重要参数: 分类的特征数量和决策树的数量. 没有被 Bagging 采用的数据称为袋外(Out of bag, OOB)数据, 利用 OOB 预测结果平均错误率来表征不同特征的重要性. RF 特征数量一般选输入变量总数的平方根^[18]. 决策树数量的上限一般设置为 1000.

根据已有地质资料, 以多尺度分割之后的各对象块为单位进行特征选择, 共选择特征 3 类, 9 个维度:

(1)光谱特征: 各矿物在 ASTER 影像的光谱范围具有特征性光谱特征, 选择各对象块的平均灰度(Sp)及标准差(De)作为光谱特征.

(2)纹理特征: 选择各对象块的变差函数纹理 4 个(Va1、Va2、Va3、Va4 分别代表 0° 、 45° 、 90° 、 135° 变差函数纹理).

(3)几何特征: 选择面积特征(Ar)、形状指数(Sh)、走向特征(Tr). Ar 是指对象块的像素总数(T). 对象边缘的平滑性指标即 Sh: $s=l/4\sqrt{T}$ ^[19], 其

中, s 为形状指数, l 为对象块的边缘长度. 矿物分布具有丛集性, 走向特征能标志矿物的分布情况. 走向特征选多个特征向量中, 特征值较大的向量的方向.

利用 RF 筛选特征, 首先使用 OOB 检测某一棵决策树 (L) 的正确分类个数 $L(a)$, 然后对 OOB 中特征 a 进行扰动, 再次检测 L 的正确分类个数 $L(a^{\wedge})$, $|L(a^{\wedge})-L(a)|$ 为 a 对 L 的重要性, 最后计算 a 对所有决策树的重要性的平均值, 即为 a 对整个森林的重要性^[20]. 根据重要性排序, 降低特征空间维度.

2.2.4 精度评价

进行两次精度评价, 一次通过 OOB 计算精度, 评价算法稳定性、并筛选树数; 第二次通过野外实地调查、采样、薄片鉴定得到矿物鉴定结果, 用混淆矩阵评价矿物提取精度, 分别统计制图精度 (Production accuracy, PA)、用户精度 (User accuracy, UA)、总精度 (Overall accuracy, OA) 和 Kappa 系数.

3 结果与分析

3.1 图像增强

经过比值运算, 目标矿物增强的同时其他光谱相似矿物的信息也增强了 (图 4). $(b_{12}+b_{10})/b_{11}$ 同时增强了黑云母信息和中基性斜长石矿物信息; $(b_5+b_7)/b_6$ 增强了白云母及高岭土等黏土矿物信息, 对伊利石、蒙脱石等亦有增强; $(b_6+b_9)/(b_8+b_7)$ 增强了黑云母和角闪石信息, 含绿泥石的变质岩也呈现高值特征; $(b_1+b_9)/b_8$ 增强了绿泥石信息, 对碳酸盐岩矿物也有较好的区分作用; b_{13}/b_{12} 增强石榴石信息的同时, 突出了碱性长石族矿物; $(b_6+b_9)/b_8$ 可增强阳起石信息, 也能突出黑云母、角闪石信息. 另外, 由于不同岩性所含矿物成分百分比不同, 导致比值运算结果有较多不确定信息. 但是, 增强影像依然可为后续工作提供数据基础.

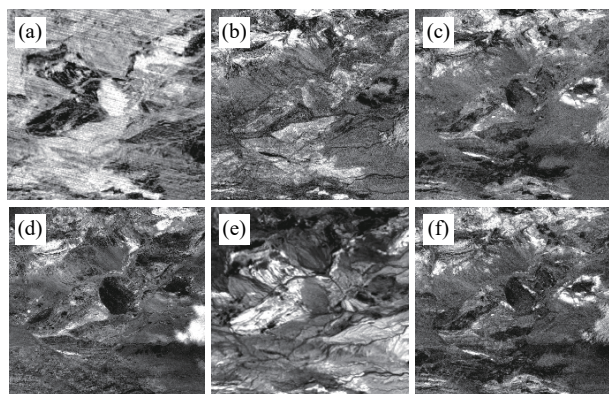


图 4 比值增强结果. (a)Bi; (b)Mus; (c)Am; (d)Chl; (e)Gt; (f)Act
Fig.4 Ratio enhancement results: (a) Bi; (b) Mus; (c) Am; (d) Chl; (e) Gt; (f) Act

3.2 多尺度分割

本文多尺度分割流程 (图 5) 如下:

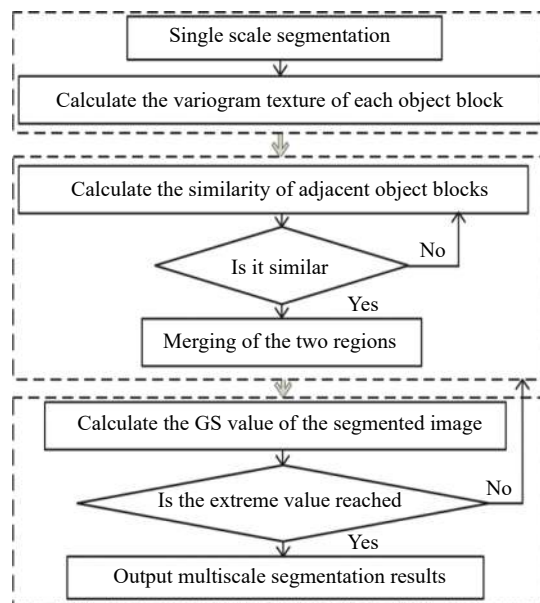


图 5 多尺度分割流程

Fig.5 Multiscale segmentation process

(1) 对目标矿物增强之后的影像进行单尺度分割. 从左上角到右下角遍历影像, 相邻像素灰度距离小于 0.002 时合并两区域, 循环遍历, 直到达到尺度限制. 分割尺度参数设为 0.2, 紧致度参数设为 0.5, 形状参数设为 0.1.

(2) 遍历初分割之后的影像, 计算各对象块的变差函数纹理, 构造整体影像的变差函数矩阵.

(3) 遍历各对象块, 计算当前对象块和相邻对象块的纹理相似性, 当两个纹理向量的卡方距离小于 0.0003 时, 合并两对象块, 修改矩阵标识. 循环执行, 直到选定最优分割结果.

(4) 计算每次合并之后的整体影像的 GS 值, 记录 GS 值的变化趋势, 输出 GS 值最小时的分割结果.

随着多尺度分割过程中合并次数增多, 各分割结果的 GS 值均呈现出最小值 (图 6). 黑云母、白云母、角闪石、绿泥石、石榴石、阳起石分别经过 1 次、5 次、6 次、4 次、1 次、2 次合并, GS 值最小. GS 值最小时的分割结果即为最优分割结果 (图 7). 增强黑云母和石榴石信息时采用了 ASTER 热红外波段影像, 热红外波段分辨率较低, 影像分割收敛更快^[21].

3.3 RF 提取矿物

3.3.1 样本选择

根据野外调查和目视解译结果, 选择目标矿物对象作为样本集. 6 种矿物样本集如表 3, 随机

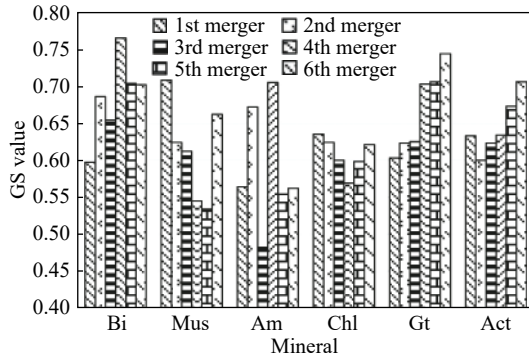


图 6 GS 值变化趋势

Fig.6 Change trend of GS values

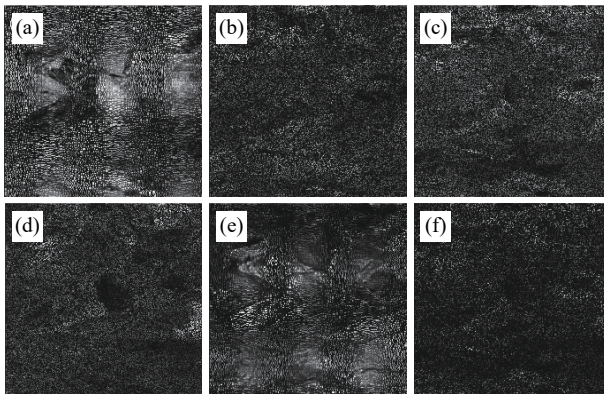


图 7 多尺度分割结果. (a)Bi; (b)Mus; (c)Am; (d)Chl; (e)Gt; (f)Act
Fig.7 Multiscale segmentation results: (a) Bi; (b) Mus; (c) Am; (d) Chl; (e) Gt; (f) Act

选择每一样本集的 2/3 样本数进行分类, 1/3 进行精度验证. 因角闪石分布较为分散, 覆盖面积大, 选择的样本最多; 而石榴石和阳起石只有少量出现, 选择的样本较少.

表 3 矿物样本数

Table 3 Mineral samples

Bi	Mus	Am	Chl	Gt	Act
1577	1348	1945	1403	912	832

3.3.2 特征选择

通过 OOB 检测误差率, 进行特征的重要程度归一化排名(图 8). 由于比值运算增强了各矿物的光谱特征, 光谱特征的重要程度最高; 变差函数统计了地质体空间相关性, 研究区在 45° 方向的区域化变量的相关性最强; 岩石由多个矿物组分, 岩性分布多为条带状或块状分布, 面积大小较大程度影响分类结果. 随着特征维数的增加, 计算效率降低. 树数不同时, 按照重要程度排名选择不同特征个数进行分类, 各结果表明随着特征个数增加提取精度先增大后变小, 综合看 3 个特征量时矿物提取精度最高, 如表 4 所示. 表 4 显示了树数为

100 时不同特征个数的提取精度, 相比于采用 9 个特征, 采用 3 个特征的各矿物提取精度分别提高了 4.3543%、2.4350%、4.6721%、3.8256%、4.9135% 及 1.6645%. 经综合判断, 最终选择 Sp、Va2、及 Ar 3 个特征量提取矿物.

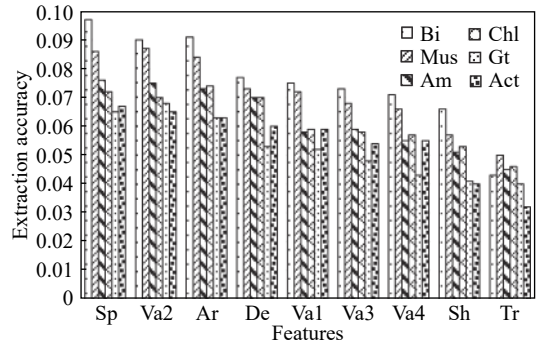


图 8 特征重要程度排名

Fig.8 Ranking of feature importance

表 4 不同特征数量矿物提取精度

Table 4 Extraction precision of minerals with different characteristic numbers

Features	Extraction precision/%					
	Bi	Mus	Am	Chl	Gt	Act
1	80.0041	75.1212	72.3755	66.0052	64.6564	60.0027
2	82.3421	76.2342	75.1137	67.2311	65.0033	60.0456
3	83.0302	76.7802	76.0232	67.6121	67.9111	60.0101
4	83.0012	76.8767	75.3545	67.5689	65.5889	60.01
5	82.4632	76.2069	75.1182	66.7865	65.3421	60.0043
6	81.8795	76.1951	74.1147	66.0011	65.001	59.9876
7	80.789	76.0022	73.2433	65.5045	64.7768	59.3425
8	79.6574	75.5467	72.1389	64.8675	64.3345	59.0001
9	78.6759	74.3452	71.3511	63.7865	62.9976	58.3456

3.3.3 决策树个数选择

为每个训练集构建验证树(从 50 到 500, 步长为 50). 通过 OOB 进行精度验证, 各矿物提取精度最高时, 对应的树数都不同(图 9). 黑云母、白云母、角闪石、绿泥石、石榴石、阳起石精度最高为 0.8404、0.7914、0.7634、0.6836、0.6791、0.6165, 精度最高时对应的数的棵树分别为 350、300、50、200、100、50. 树数对矿物提取精度的影响在可控范围之内, 从 50 棵到 500 棵引起的精度变化最大的是石榴石, 精度变化幅度为 0.0471, 最小的是阳起石, 变化幅度为 0.0177.

3.3.4 矿物提取结果

变质矿物提取结果(图 10)主要分布在晚太古代变质地层中, 与实际地质情况吻合. 受多期次区

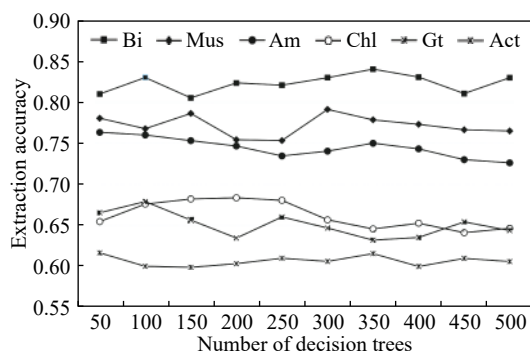


图 9 决策树个数选择

Fig.9 Number selection of decision trees

域构造和热液活动的改造,该区域是变质矿物集中分布区,也是矿化蚀变集中分布区域,发育绿帘石化、绿泥石化、阳起石、碳酸盐化、矽卡岩化及绢云母化。结果显示,黑云母主要分布在西北部和西南部的晚太古代变质地层中,相关岩性有:黑云斜长片麻岩、二云斜长片麻岩、二云二长片麻岩、斜长角闪岩、二云石英片岩和黑云母大理岩等。白云母主要分布在晚太古代变质岩区,少量出现在二叠纪花岗岩区,相关岩性有:石榴石白云母斜长片麻岩、二云斜长片麻岩、二云石英片岩、二云石英片岩和含白云母石英岩等。角闪石主要分布在晚太古代斜长角闪岩和斜长角闪片麻岩区域,少量在蓟县纪大理岩区。绿泥石矿物信息主要在晚太古代钠长绿帘绿泥片岩和黑云斜长片麻岩中,少量出现在长城纪绿泥石千枚岩和二叠纪的钠长绿泥石英千枚岩中。石榴石矿物只出现在西北部的晚太古代地层中,主要赋存于石榴石白云母斜长片麻岩和石榴黑云斜长片麻岩中。阳起石仅在蓟县纪蛇纹透辉白云质大理岩中有零星分布。

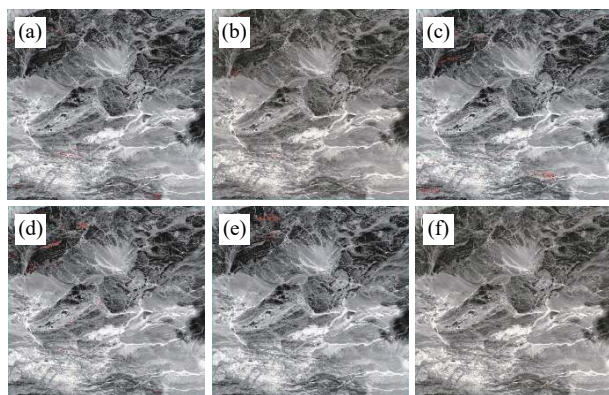


图 10 矿物提取结果。(a)Bi; (b)Mus; (c)Am; (d)Chl; (e)Gt; (f)Act
Fig.10 Mineral extraction results: (a) Bi; (b) Mus; (c) Am; (d) Chl; (e) Gt; and (f) Act

3.4 野外调查精度评价

通过 42 个点的野外调查、采样和实验室薄片

鉴定等步骤,对本文的矿物信息提取结果进行精度评价。由于野外勘查条件限制,实际采样点未能均匀分布在矿物显现区,根据岩性分布情况进行了补充采样。部分样品薄片鉴定结果如表 5 和图 11。

通过野外实测生成测试集,评价模型的泛化能力。精度评价结果(表 6)显示,提取精度较高的为黑云母、白云母、角闪石,较低的为绿泥石、石榴石、阳起石。黑云母在黑云母大理岩及黑云斜长片麻岩中质量分数最高可达 20%,且分布较广,实际采样点多覆盖在这些岩性上,总体精度为 85.4088%、Kappa 为 0.7779。白云母在石榴石白云母斜长片麻岩及二云石英片岩中质量分数最高可达 30%,二云斜长片麻岩中质量分数也达到了 10%,提取精度为 84.7640、Kappa 为 0.7833。角闪石在斜长角闪岩中质量分数最高可达到 75%,斜长角闪片麻岩中质量分数最高可达到 55%,矿物特征较为鲜明,提取精度为 85.7308、Kappa 为 0.7748。绿泥石在绿泥石英千枚岩、含绢云母石英岩、斜长角闪岩和斜长角闪片麻岩等岩性中均有体现,但质量分数一般不超过 15%(仅在极少量的钠长绿泥石英千枚岩中质量分数达到了 48%),因此,容易被其他高含量矿物干扰,提取精度为 70.6933、Kappa 为 0.5938。石榴石在石榴石黑云斜长片麻岩中质量分数为 5%、在石榴石白云母斜长片麻岩中质量分数为 3%,在其他岩性中均未出现,所以提取时样本训练程度不足导致它的总体精度仅为 65.5992、Kappa 为 0.5462。阳起石仅在蛇纹透辉白云质大理岩中出现了 5%,提取精度为 66.7509、Kappa 为 0.5560。

测试样本的拟合决定系数 R^2 被用来评价模型的预测准确性^[22]。各采样点实测值和预测值对比, Bi、Mus、Am 及 Chl 的 R^2 分别为 0.971、0.962、0.966 及 0.965, 预测值与实测值基本吻合; Gt 及 Act 的 R^2 分别为 0.935 及 0.932, 拟合效果不佳,但是由于在研究区石榴石和阳起石分布非常少,所以提取结果依然具备实用价值。

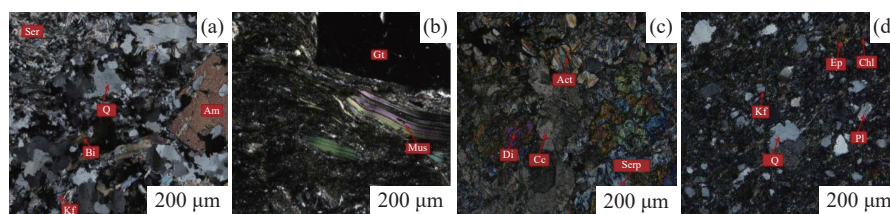
3.5 与其他方法提取精度对比

将本文方法与比值+阈值分割法、比值+SVM 法提取的总体精度进行对比(表 7)。各方法采用的比值运算公式一致, Bi、Mus、Am、Chl、Gt 及 Act 的分割阈值分别为: 1.4、3.1773、1.1406、3.01、1.405 及 2.25。采用 SVM 分类时的样本及特征与本文方法一致,核函数选择径向基函数(Radial basis function, RBF),参数 σ 分别取测试错误率最

表5 部分样品薄片鉴定结果

Table 5 Identification results of some sample slices

Sample number	Sampling location		Lithology	Mineral composition
	Longitude	Latitude		
D0158	96°32'13.1"	41°59'41.4"	Albite epidote chlorite schist	Quartz 48%, Feldspar 25%, Chlorite 12%, Epidote 8%, Biotite 3%, Sphene 3%
D0126	96°32'53.1"	41°59'20.4"	Plagioclase amphibolite	Amphibole 75%, Plagioclase 15%, Quartz 5%, Carbonate minerals 4%, Opaque minerals 1%
D0123	96°32'56.6"	41°59'19.5"	Garnet muscovite plagioclase gneiss	Muscovite 30%, Feldspar 62%, Quartz 5%, Garnet 3%
D0121	96°32'57.6"	41°59'19.3"	Biotite plagioclase gneiss	Quartz 45%, Plagioclase 33%, Biotite 20%, Amphibole 2%
D0117	96°33'8.2"	41°59'17.8"	Biotite plagioclase gneiss	Quartz 28%, Plagioclase 25%, Biotite 25%, Chlorite 15%, K-feldspar 5%, Opaque minerals 2%
D0107	96°33'16.8"	41°59'7.5"	Tonalite	Quartz 42%, Feldspar 34%, Biotite 12%, Amphibole 8%, Carbonate minerals 2%, Sericite 2%
D0105	96°33'16.6"	41°59'4"	Garnet biotite plagioclase gneiss	Quartz 50%, Feldspar 31%, Biotite 10%, Garnet 5%, Opaque minerals 3%, Apatite 1%
D0103	96°33'16.3"	41°59'3"	Garnet biotite plagioclase gneiss	Quartz 40%, Plagioclase 30%, K-feldspar 7%, Biotite 12%, Garnet 3%
D1125	96°31'36.8"	41°57'24.3"	Serpentine diopside dolomitic marble	Calcite 49%, Diopside 32%, Serpentine 14%, Actinolite 5%
D0901	96°40'47.7"	41°55'39.2"	Albite chlorite quartz phyllite	Quartz 32%, Plagioclase 16%, Chlorite 48%, Epidote 4%, Sphene
D0420	96°34'35.6"	41°52'36.5"	Plagioclase amphibolite	Plagioclase 55%, Amphibole 30%, Quartz 10%, Biotite 3%, Epidote 2%
D0426	96°34'34.5"	41°52'29"	Mica quartz schist	Quartz 50%, Muscovite 30%, Biotite 20%
D0450	96°34'14.9"	41°51'38.6"	Plagioclase hornblende gneiss	Plagioclase 35%, Amphibole 55%, Epidote 5%, Quartz 5%
D1007	96°30'44.9"	41°50'45.1"	Mica monzonite gneiss	K-feldspar 30%, Plagioclase 25%, Quartz 27%, Biotite 14%, Muscovite 4%
D0055	96°33'49"	41°50'6.3"	Biotite plagioclase gneiss	Plagioclase 38%, Quartz 35%, Biotite 20%, Muscovite 3%, Chlorite 4%
D0113	96°34'19.3"	41°51'46"	Plagioclase hornblende gneiss	Amphibole 45%, Albite 37%, Microcline 10%, Quartz 5%, Pyroxene 2%, Opaque minerals 1%



Q—Quartz; Pl—Plagioclase; Kf—K-feldspar; Ser—Sericite; Am—Amphibole; Bi—Biotite; Di—Diopside; Act—Actinolite; Serp—Serpentine, Cc—Carbonate minerals; Mus—Muscovite; Gt—Garnet; Chl—Chlorite; Ep—Epidote

图 11 部分样品显微照片。(a)云母石英片岩;(b)石榴石白云母斜长片麻岩;(c)蛇纹石透辉石白云质大理岩;(d)钠长石绿泥石石英千枚岩

Fig.11 Micrograph of some samples: (a) mica quartz schist; (b) garnet muscovite plagioclase gneiss; (c) serpentine diopside dolomitic marble; (d) albite chlorite quartz phyllite

表6 精度评价

Table 6 Accuracy evaluation

Mineral	PA/%	UA/%	OA/%	Kappa
Bi	80.95	78.64	85.4088	0.7779
Mus	83.35	75.60	84.7640	0.7833
Am	79.74	77.55	85.7308	0.7748
Chl	68.74	63.76	70.6933	0.5938
Gt	58.42	58.43	65.5992	0.5462
Act	59.61	64.41	66.7509	0.5560

小的 14、11、13、19、8 及 12, C 统一取值为 9。黑云母、角闪石及绿泥石出露较多且较为集中, 经多尺度分割之后精度提高幅度较大, 本文方法比比值+阈值分割法分别提高了 8.0542%、8.7624% 及 6.0339%; 比比值+SVM 法分别提高了 4.2747%、4.6306% 及 4.2401%。而白云母、石榴石、阳起石出露较少, 精度提高幅度不大, 本文方法比比值+阈值分割法分别提高了 3.481%、3.757% 及 3.9662%; 比比值+SVM 法分别提高了 1.777%、2.4106% 及 2.5158%。

表 7 本文方法与其他方法提取精度对比

Table 7 Comparison of extraction accuracy between the present method and other methods

Mineral	Extraction accuracy/%		
	Ratio + threshold segmentation	Ratio +SVM	Method of this paper
Bi	77.3546	81.1341	85.4088
Mus	81.2830	82.9870	84.7640
Am	76.9684	81.1002	85.7308
Chl	64.6594	66.4532	70.6933
Gt	61.8422	63.1886	65.5992
Act	62.7847	64.2351	66.7509

4 结论

根据变质矿物的特征性光谱特征进行比值运算增强 ASTER 影像, 并基于光谱特征和变差函数纹理进行多尺度分割, 然后, 通过 RF 提取变质矿物信息, 最后经过野外验证进行精度评价。结果表明, 黑云母、白云母、角闪石等在 ASTER 影像上具有鉴定性特征, 提取精度可分别达到 85.4088%、84.7640%、85.7308%; 而绿泥石、石榴石、阳起石作为次要矿物, 鉴定时被主要造岩矿物干扰, 提取精度达到 60% 以上。

基于野外调查和 ASTER 影像处理进行变质矿物的提取可有效提高变质岩野外调查效率和精度。本文方法可较为准确的鉴定 ASTER 的变质矿物; 可为其他遥感影像提取矿物提供借鉴; 也可作为辅助地质调查的有效手段。与其他类似研究相比, 基于变差函数的多尺度分割能增强形态特征对矿物信息的区分能力; RF 对训练数据的统计假设少、对矿物混合导致的噪声不敏感、分类变异性低, 对多种矿物组成的岩性具有较强的分析能力。

下一步研究方向: ①比值运算增强目标矿物的同时也增强了其他具有相似光谱特征的矿物, 导致误分类, 可通过原始岩性进行增强约束; ②多尺度分割对对象内部的同质性要求较高, 然而实际像元为多种矿物的混合, 过分割现象较多, 可通过解混像元突出主要信息; ③RF 的分类精度取决于树的棵树, 可通过寻优算法代替人工判定树的棵树。

参 考 文 献

- [1] Diener J F A, White R W, Link K, et al. Clockwise, low-P metamorphism of the Aus granulite terrain, southern Namibia, during the Mesoproterozoic Namaqua Orogeny. *Precambrian Res*, 2013, 224: 629

- [2] Xie M H, Zhang Q, Chen S B, et al. Extraction of alteration anomaly information by feature-based principal component analysis from ASTER data. *Editorial Committee Earth Sci J China Univ Geosci*, 2015, 40(8): 1381
(谢明辉, 张奇, 陈圣波, 等. 基于特征导向主成分分析遥感蚀变异常提取方法. *地球科学—中国地质大学学报*, 2015, 40(8): 1381)
- [3] Zadeh M H, Tangestani M H, Roldan F V, et al. Sub-pixel mineral mapping of a porphyry copper belt using EO-1 Hyperion data. *Adv Space Res*, 2014, 53(3): 440
- [4] Wu Z C, Ye F W, Guo F S, et al. A review on application of techniques of principle component analysis on extracting alteration information of remote sensing. *J Geo-Inf Sci*, 2018, 20(11): 1644
(吴志春, 叶发旺, 郭福生, 等. 主成分分析技术在遥感蚀变信息提取中的应用研究综述. *地球信息科学学报*, 2018, 20(11): 1644)
- [5] Liu Y Z, Lai H R, Zhang D W, et al. Change detection of high resolution remote sensing image alteration based on multi-feature mixed kernel SVM model. *Remote Sens Land Resour*, 2019, 31(1): 16
(刘立志, 赖华荣, 张丁旺, 等. 多特征混合核 SVM 模型的遥感影像变化检测. *国土资源遥感*, 2019, 31(1): 16)
- [6] He Z H, He B B. Weight spectral angle mapper (WSAM) method for hyperspectral mineral mapping. *Spectrosc Spectr Anal*, 2011, 31(8): 2200
(何中海, 何彬彬. 基于权重光谱角制图的高光谱矿物填图方法. *光谱学与光谱分析*, 2011, 31(8): 2200)
- [7] Kaur S, Bansal R K, Mittal M, et al. Mixed pixel decomposition based on extended fuzzy clustering for single spectral value remote sensing images. *J Indian Soc Remote Sens*, 2019, 47(3): 427
- [8] Feng W Q, Sui H G, Tu J H, et al. Change detection method for high resolution remote sensing images using random forest. *Acta Geodaet Cartograph Sin*, 2017, 46(11): 1880
(冯文卿, 眭海刚, 涂继辉, 等. 高分辨率遥感影像的随机森林变化检测方法. *测绘学报*, 2017, 46(11): 1880)
- [9] Booyesen R, Zimmermann R, Lorenz S, et al. Towards multiscale and multisource remote sensing mineral exploration using RPAS: a case study in the Lofdal carbonatite-hosted REE deposit, Namibia. *Remote Sens*, 2019, 11(21): 2500
- [10] Cid Y D, Muller H, Platon A, et al. 3D solid texture classification using locally-oriented wavelet transforms. *IEEE Trans Image Process*, 2017, 26(4): 1899
- [11] Wang M, Zhang X C, Wang J Y, et al. Forest resource classification based on random forest and object oriented method. *Acta Geodaet Cartograph Sin*, 2020, 49(2): 235
(王猛, 张新长, 王家耀, 等. 结合随机森林面向对象的森林资源分类. *测绘学报*, 2020, 49(2): 235)
- [12] You Y F, Wang S Y, Wang B, et al. Study on hierarchical building extraction from high resolution remote sensing imagery. *J Remote*

- Sens*, 2019, 23(1): 125
(游永发, 王思远, 王斌, 等. 高分辨率遥感影像建筑物分级提取. *遥感学报*, 2019, 23(1): 125)
- [13] Li J Y, Zhao Y K, Xue Z E, et al. A survey of model compression for deep neural networks. *Chin J Eng*, 2019, 41(10): 1229
(李江昀, 赵义凯, 薛卓尔, 等. 深度神经网络模型压缩综述. *工程科学学报*, 2019, 41(10): 1229)
- [14] Cracknell M J, Reading A M. Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. *Comput Geosci*, 2014, 63: 22
- [15] Harris J. R, He J X, Rainbird R, et al. A comparison of different remotely sensed data for classifying bedrock types in Canada's arctic: application of the robust classification method and random forests. *Geosci Can*, 2014, 41(4): 557
- [16] Hossain M D, Chen D M. Segmentation for object-based image analysis (OBIA): A review of algorithms and challenges from remote sensing perspective. *ISPRS J Photogramm Remote Sens*, 2019, 150: 115
- [17] Diaz G F, Ortiz J M, Silva J F, et al. Variogram-based descriptors for comparison and classification of rock texture images. *Math Geosci*, 2020, 52(4): 451
- [18] Zhang L, Liu Z, Ren T W, et al. Identification of seed maize fields with high spatial resolution and multiple spectral remote sensing using random forest classifier. *Remote Sens*, 2020, 12(3): 362
- [19] Zhu J J, Fan X T, Du X P. Geometric feature representation and building extraction based on geometric features. *J Appl Sci*, 2015, 33(1): 9
(朱俊杰, 范湘涛, 杜小平. 几何特征表达及基于几何特征的建筑物提取. *应用科学学报*, 2015, 33(1): 9)
- [20] Masoumi F, Eslamkish T, Abkar A A, et al. Integration of spectral, thermal, and textural features of ASTER data using random forests classification for lithological mapping. *J Afric Earth Sci*, 2017, 129: 445
- [21] Pourmamdari M, Hashim M, Pour A B. Spectral transformation of ASTER and Landsat TM bands for lithological mapping of Soghan ophiolite complex, South Iran. *Adv Space Res*, 2014, 54(4): 694
- [22] Zhang B, He B B. Multi-scale segmentation of high-resolution remote sensing image based on improved watershed transformation. *J Geo-Inf Sci*, 2014, 16(1): 142
(张博, 何彬彬. 改进的分水岭变换算法在高分辨率遥感影像多尺度分割中的应用. *地球信息科学学报*, 2014, 16(1): 142)