结合小波包变换和随机森林的 ASTER 蚀变信息提取

唐淑兰^{1,2)},曹建农¹⁾,王国强³⁾,卜涛³⁾

1) 长安大学地球科学与资源学院,西安,710054; 2) 西安财经大学管理学院,西安,710100;
 3) 中国地质调查局西安地质调查中心,西安,710054

内容提要:为了更准确地提取蚀变信息,本文选择新疆、甘肃和内蒙古三省交界部位为研究区,结合小波包变换和随机森林提取 ASTER 蚀变信息。首先,选择主要蚀变类型的诊断性波段进行特征向量主成分分析,得到主分量影像;接着,对主分量影像进行小波包变换,使用代价函数选择最优小波包树,并提取高低频信息构造分类向量; 然后,经过特征筛选构造随机森林分类模型,并提取矿化蚀变信息;最后,通过野外采样、薄片鉴定对提取结果进行 精度评价。铁染、Al-OH 及 Mg-OH 蚀变信息的主成分分析波段组合分别选择 Band 1、2、3、4,Band 1、3、4、6 及 Band 1、5、8、9。结果表明,本文方法提取铁染、Al-OH 基团及 Mg-OH 基团蚀变信息的总体精度可达到 88.7443、 85.5469 及 91.7594,Kappa 分别为 0.7767、0.6732 及 0.8362,与成矿区带及已有的该区域的成矿特征相关性较 好。本研究采用的最优小波包树能充分利用矿物光谱的能量特征,随机森林可削弱矿物组分的噪声干扰,研究结 果可为遥感蚀变信息提取提供技术参考。

关键词: ASTER; 蚀变信息; 小波包变换; 随机森林; 主成分分析

围岩蚀变是遥感找矿的重要解译标志之一(Fu Shuixing et al.,2004; Huang Shuang et al.,2015; Li Zhizhong et al.,2015; Liu Dechang et al.,2017; Zhao Yingjun et al.,2015; Zhu Weiping et al., 2014)。热液成矿过程中生成的 Fe³⁺或 Fe²⁺离子、 Mg-OH 和 Al-OH 等基团在遥感影像上显示特征 性的光谱特征和颜色。根据不同蚀变矿物的光谱及 空间特征,结合遥感和图像处理技术可以定位蚀变 矿物分布区,能够免于自然条件约束,提高矿产资源 勘查效率。

有效的遥感蚀变信息提取方法有:比值法、主成 分分析(Principal Component Analysis,PCA)、光谱 角(Spectral Angle Mapping,SAM)、混合像元分解 等(Fallatah et al.,2020;Gong Hao et al.,2009; Kaur et al.,2019;Mahboob et al.,2019;Uddin et al.,2020;Vignesh et al.,2020;Wang Runsheng et al.,2011;Xu Kai et al.,2019)。ASTER 影像因其 高空间分辨率有较好的矿物和岩性绘图能力。 Zhang Ruisi et al. (2016)基于 ASTER 光谱特征利 用波段比值法和假彩色合成法提取了与 Fe³⁺、硅酸 盐、Al-OH 等成分相关的矿物。Xie Minghui et al. (2015)利用特征导向主成分分析法提取了辽宁兴城 地区的褐铁矿(Fe³⁺)、绿泥石(Mg-OH 基团)和高 岭石(Al-OH 基团)蚀变异常信息。Zhang Tingbin et al. (2016)将 ASTER 和 Landsat-8oli(AO)结合 起来探测了青藏高原多龙斑岩铜金矿。更多图像处 理技术被应用到 ASTER 蚀变信息提取中。Xu Ka et al. (2019)利用 ASTER 数据基于支持向量机和 主成分分析识别了古龙金矿地区的热液蚀变矿物。 Moore et al. (2008)应用主成分分析和匹配滤波根 据已知矿床的训练集识别未知目标。Hosseinjan et al. (2011)使用亚像素分解来确定每个像素内不 同矿物的相对比例。Abdi et al. (2012)还将光谱角 映射器应用于 ASTER 数据,以区分热液蚀变。

transform and random forest. Acta Geologica Sinica, 95(3): 924~933.

注:本文为中国地质调查局项目(编号 DD20190812)、国家自然科学基金项目(编号 41571346)、陕西省教育厅项目(编号 18JK0317)、陕西省自然科学基础研究计划项目(编号 2020JM-585)、西安财经大学项目(编号 16FCJH05)资助的成果。

收稿日期:2020-08-06;改回日期:2020-09-25;网络发表日期:2021-02-21;责任编委:范宏瑞;责任编辑:李曼。

作者简介:唐淑兰,女,1979年生。讲师,博士研究生,从事遥感影像模式识别研究。Email: mzwsbjh@126.com。通讯作者:曹建农,男, 1963年生。教授,博士,从事遥感影像模式识别研究。Email: zyjncao@chd.edu.cn。

引用本文:唐淑兰,曹建农,王国强,卜涛. 2021. 结合小波包变换和随机森林的 ASTER 蚀变信息提取. 地质学报, 95(3):924~933, doi: 10.19762/j. cnki. dizhixuebao. 2021208. Tang Shulan, Cao Jiannong, Wang Guoqiang, Bu Tao. 2021. Aster alteration information extraction based on wavelet packet

Abbaszadeh et al. (2013)在伊朗 Rabor 地区使用最 小二乘法比较图像光谱与参考光谱的拟合度来增强 热液蚀变。这些方法有效地提取了 ASTER 影像的 蚀变信息,但是,没有充分利用高分辨率遥感影像的 多尺度细节特征,导致提取结果噪声较多,精度 有限。

小波包分解是在小波分解的基础上发展的能对 高低频信号同时进行分解的方法,能保留影像全部 信息,有时频局部化及多尺度分析优势(Seo et al., 2016)。随机森林(Random Forest, RF)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,提取遥感信息的准确度 较高(He Xianjin et al.,2018)。本文利用小波包分 解与 RF 的优势,从各蚀变矿物的主分量影像上提 取蚀变信息,旨在利用矿物区域特征及高级分类器 为蚀变矿物的遥感解译提供技术参考。

1 研究区地质概况

研究区(图1)位于新疆、甘肃和内蒙古三省交 界部位的甘肃省玉门市地区,大地构造隶属于北山 造山带,是古亚洲成矿域的重要组成部分。成矿带 上属于甘-蒙北山南部的磁海-公婆泉成矿带上(Li Junjian et al., 2016), 该成矿带是在前寒武纪基底 之上发育的早古生代裂谷带,经历了复杂的地质构 造演化和构造岩浆活动,是铁-铜-金-铅-锌-锰-钨-锡-铷-钒-铀-磷的有利成矿部位,成矿类型以斑岩型 和岩浆热液型为主,其次为沉积变质型、沉积改造型 和热液型,已发现的典型矿床有铜矿、金矿、铁矿、锰 矿和磷矿等。出露的地层有:新太古界、长城系、蓟 县系、石炭系、二叠系、侏罗系、白垩系和第四系,并 发育有石炭纪和二叠纪中酸性侵入岩。新太古界、 长城系和蓟县系中可见绢英岩化和绿泥石化,并含 大理岩或白云岩,大理岩与酸性侵入岩接触界线附 近常见砂卡岩化;蓟县系大理岩中偶见阳起石;石炭 系和二叠系碎屑岩中可见绿帘石化。

2 数据与方法

2.1 数据

(1) ASTER 数据: ASTER 数据共有 14 个波段:可见-近红外(VNIR)有 3 个空间分辨率为 15m的波段(0.52~0.86μm);短波红外(SWIR)有 6 个空间分辨率为 30m的波段(1.60~2.43μm);热红外(TIR)有 5 个空间分辨率为 90 m的波段(8.125~11.65μm)。研究区为甘肃中盐池地区的 ASTER



达幅地质简图(实测)

Fig. 1 Geological map of 1 : 50000 delkhada sheet in Mazongshan area, Beishan, Gansu Province (Survey)

数据,影像为 2003 年 8 月 15 日拍摄的 L1T 级数据,图像清晰,无积雪、植被覆盖,有云覆盖。

(2)主要蚀变类型及光谱特征:成矿地质条件有 利部位包括特殊的构造带、蚀变带和特殊岩层等,受 多期次区域构造和热液活动的影响。研究区内岩石 发生蚀变和变质作用,类型主要有绿帘石化、绿泥石 化、阳起石化、碳酸盐化、砂卡岩化、绢云母化、赤铁 矿化、硅化和角岩化等。蚀变岩类型取决于原岩的 成分,比如:砂卡岩集中产出于大理岩和侵入岩的接 触带上,大量互不相连的碳酸盐脉常出现在灰岩和 大理岩中;绢云母化表现为火山碎屑岩中的碎屑及 泥质物被绢云母所取代;绿泥石化在基性火山岩和辉 长岩中常见,部分酸性岩和泥质岩中可见。参考美国 地质调查局(USGS)矿物的波谱曲线,三类蚀变矿物 吸收谱带与 ASTER 波段的对应关系如图 2。

2.2 方法

本文工作流程(图 3):首先,对 ASTER 影像进 行预处理;其次,选择各蚀变矿物诊断性光谱特征进 行特征向量主成分分析(Feature Oriented Principal Components Selection, FPCS),并选择各蚀变矿物 的主分量影像;接着,对主分量影像进行小波包变 换,用代价函数选择最优小波包树,从变换结果的高 频和低频部分分别提取纹理和光谱特征,构造分类





Fig. 2 The corresponding relationship between the spectrum curve of altered minerals and aster band (a)—铁染;(b)—Al-OH;(c)—Mg-OH (a)—Iron-stain;(b) Al-OH;(c)—Mg-OH

特征向量;然后,采集各蚀变信息的样本集,构建 RF并提取蚀变矿物;最后通过检验样本集、野外勘 查及薄片鉴定结果进行精度评价,统计总体精度 (OA)和 Kappa 系数。ASTER 数据预处理包括串 扰校正、辐射定标、大气校正及去云处理。采用 FLAASH 模块进行大气校正,提取近红外波段异常 高值进行掩膜运算去除云覆盖。另外,进行几何精 校正,将 VNIR 15m 分辨率重采样为 30m 分辨率。 精度评价时将野外实证点投影到遥感影像上,以投 影点为中心随机生成面积不超过 80 的不规则多边 形,随机提取多边形内部 1/3 的点作为检验样本。



图 3 结合小波包变换和随机森林的 ASTER 蚀变信息提取流程

Fig. 3 Process of extracting aster alteration information by combining wavelet packet transform and random forest

2.2.1 主成分分析

主成分分析可去除 ASTER 影像各波段的相关 性并降维。选择目标蚀变矿物的 3~4 个诊断性光 谱波段进行主成分分析的方法称为 FPCS,4 个波段 FPCS 的 Crosta 技术能有效提取铁化、泥化蚀变信 息(Loughlin et al.,1991)。FPCS 之后可生成 4 个 主分量,与该矿物反射峰波段对应的特征向量载荷 因子符号为正、吸收谷波段对应符号为负的主分量, 即为蚀变异常主分量,这时蚀变位于该主分量亮值 区,符号相反时蚀变位于暗值区。蚀变信息主分量常 为 PC4 或 PC3。铁染蚀变信息的诊断性波段可选择 PCA(1,2,3,4)或 PCA(1,3,4,5);Al-OH 基团蚀变 信息的诊断性波段选择 PCA(1,3,4,6)、PCA(1,4, 5,7)或 PCA(1,4,6,7);Mg-OH 及 CO_3^{2-} 基团蚀变 信息的诊断性波段选择 PCA(1,5,8,9)、PCA(1,3, 4,8)、PCA(1,3,8,9)或 PCA(1,3,5,7)(Wu Zhichun et al.,2018;Liu Jianyu et al.,2019)。

2.2.2 小波包变换

小波包是小波的推广,能更大程度地控制时间-频率平面的分开度,对高频和低频部分都进行分解。 对于图像,变换时考虑二维、四子带滤波器组。尺度 和平移基函数为:

 $\varphi_{j,m,n}(x,y) = 2^{j/2} \varphi(2^{j}x - m, 2^{j}y - n)$ (1) $\psi^{i}_{j,m,n}(x,y) = 2^{j/2} \psi^{i}(2^{j}x - m, 2^{j}y - n), i = \{H, V, D\}$ (2)

其中,H、V、D 分别为表示水平、垂直和对角线方向。大小为 $M \times N$ 的图像 f(x,y)的离散小波变换是:

$$W_{\varphi}(j_{0},m,n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1N-1} \int_{y=0}^{N-1} f(x,y) \varphi_{j_{0},m,n}(x,y)$$
(3)

$$W^i_{\psi}(j,m,n) = \frac{1}{\sqrt{MN}}$$

 $\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \psi_{j,m,n}^{i}(x,y), i = \{H,V,D\}$ (4)

 j_0 是一个任意的开始尺度, $W_{\varphi}(j_0, m, n)$ 系数定义 f(x,y)在尺度 j_0 处的近似。 $W_{\varphi}^i(j, m, n)$ 系数对尺 度 $j \ge j_0$ 附加了水平、垂直和对角方向的细节。在 $W_{\varphi}(j,m,n) = f(m,n)$ 时,该过程可以对尺度 j = j $-1, j - 2, \dots, j - P$ 反复迭代生成 P 尺度变换。图 4 为一个三尺度、二维小波包树的一部分。



图 4 三尺度、完全小波包分解树

Fig. 4 Complete wavelet packet decomposition tree based on three scales

由于小波包树有多种分解选择,实际应用中无法 将所有的分解都列举出来,需要有效的算法来筛选节 点,用以熵为基础的代价函数选择最优小波包树:

$$E(f) = \sum |f(m,n)| \tag{5}$$

 $E(f) 度量 f 的能量。对于所有的 m 和 n,函数 f(m,n)=0的能量为零。计算节点的能量,用 <math>E_p$ (作为父节点的能量)表示,此节点的 4 个子节点的能量分别为 E_A , E_H , E_V 和 E_D 。对于二维小波包分解,父节点是近似系数或细节系数的一个二维阵列; 子节点分别是滤波后的近似、水平、垂直和对角线细节。若子节点的联合能量小于父节点的能量,即 E_A + E_H + E_V + E_D < E_P ,则保留这些子节点;否则,修 剪掉这些子节点,保留父节点。

统计小波包变换高频系数的值可以得到影像的 多尺度纹理特征,本文选择的统计量有:熵(En)、均 值(M)、方差(V)、能量(E),公式见 Sankar A S et al. (2015)。由于高频部分系数为水平、垂直和对角 三个方向,统计量取三个方向的均值。从小波包变 换低频子带中可提取不同尺度光谱特征(S)。构造 分类特征向量:

$$T = [E_n, M, V, E, S] \tag{6}$$

2.2.3 RF

RF 是以多棵决策树为基础的非参数分类器, 最终分类结果为多棵决策树平均或者投票的结果。 不同的训练样本子集可以提高分类模型的预测能力,分类结果精度高、性能稳定(Feng Wenqing et al.,2017),对于矿物组分复杂、噪声较多的岩性或 矿物识别效果较好。具体过程:每颗树采用 Bagging 随机选择原始样本集的子样本集,利用 CART(Classification and Regression Trees)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。没有被 Bagging 采用的数据称为 OOB(out of bag)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 有两个重要参数:决策树的数量和特征数量。本文决策树数量的上限设置为 500,特征数量选输入变量总数的平方根(Zhang Lin et al.,2020)。

RF 评估特征的重要程度并进行筛选的原理: 对某棵决策树 L 使用其对应的 OOB 数据进行性能 检测,记正确分类的个数为 FL;接着对 OOB 数据 所有样本的 F 特征进行随机干扰,并记正确分类个 数 FL',对于决策树 L 特征 F 的重要程度表示为: $D(F)_i = |FL_i - FL_i|(i=1,2,3,...,N);对于整个$ RF 的 F 特征的重要程度表示为:I(F) = [D(F)1+D(F)2+...+D(F)N]/N,其中,N 为决策树数量。特征筛选流程(图 5)中 m 为输入的特征数量,n 为 最终选定的特征的维数。



图 5 RF 提取蚀变信息时的特征筛选流程 Fig. 5 Feature selection process for RF extraction of alteration information

根据野外调查结果,分别选择蚀变信息样本点 各 300 个,以这 300 个样本点为中心随机产生面积不 超过 20 的不规则多边形作为原始训练样本集。原始 样本集的 2/3 用来构建 RF,1/3 评价提取精度。

3 结果与分析

3.1 主成分分析

赤铁矿 是含铁蚀变矿 物的代表,在 0.54 μm (Band 1)和 0.85 μm(Band 3)处吸收最强烈;褐铁 矿因含 Fe³⁺在 0.76~0.90 μm(Band 3)处有强吸收 谷,而在 Band 2 处有反射峰。选择 Band 1、2、3、4 进行 FPCS 提取铁染蚀变。特征向量矩阵(表 1)中 Band 2 和 Band 4 载荷因子符号为正而 Band 1 和 Band 3 载荷因子符号为负的主分量为 PC4,则 PC4 为主要贡献源。

与羟基有关的蚀变矿物有高岭石、白云母、绢云母、蒙脱石等。高岭石在 Band 4 有吸收谷;白云母

表 1 ASTER 波段 1、2、3、4 主成分分析特征向量矩阵 Table 1 Eigenvector matrix of Aster Band 1, 2, 3 and 4 obtained by PCA

			•	
主分量	Band 1	Band 2	Band 3	Band 4
PC1	0.9783	0.1335	0.1013	0.1215
PC2	0.151	-0.98	-0.1242	-0.0359
PC3	0.0541	0.1218	-0.9631	0.2338
PC4	-0.1308	0.0828	-0.2162	0.964

(或绢云母)在 2.3 μ m 处有弱吸收谷(Band 8);高岭 石和绢云母在 2.2 μ m(Band 6)处均有 Al-OH 引起 的强吸收特征。选择 Band 1、3、4、6 进行 FPCS,特 征向量矩阵(表 2)中 PC4 分量 Band 6 载荷因子符 号为负,且绝对值较大,Band 4 载荷因子符号为正, 可选为主要贡献源。

绿泥石与绿帘石波谱特征相似,在 2.335 μm 处有强吸收(Band 8),在 Band 5 和 Band 9 有强反 射。方解石在 Band 1 有弱吸收谷,Band 8 有强吸 收谷,Band 9 处为反射峰,因此选择 Band 1、5、8、9

表 2 ASTER 波段 1、3、4、6 主成分分析特征向量矩阵 Table 2 Eigenvector matrix of Aster Band 1, 3, 4 and 6 obtained by PCA

	-			
主分量	Band 1	Band 3	Band 4	Band 6
PC1	0.9842	0.1003	0.0498	-0.1369
PC2	0.1237	-0.976	0.0114	0.1786
PC3	-0.0667	-0.0218	0.9882	-0.1358
PC4	-0.1074	-0.1919	0.1441	-0.9648

Fig. 6

进行 FPCS,提取碳酸盐岩化蚀变。特征向量矩阵 中(表 3),Band 5 和 Band 9 载荷因子符号为正, Band 1 和 Band 8 载荷因子符号为负,对应的 PC4 为主要贡献源。

表 3 ASTER 波段 1、5、8、9 主成分分析特征向量矩阵 Table 3 Eigenvector matrix of Aster Band 1, 5, 8

and 9 obtained by PCA

主分量	Band 1	Band 5	Band 8	Band 9
PC1	-0.9867	-0.151	-0.0587	-0.0104
PC2	-0.1496	0.9884	-0.0266	-0.0066
PC3	0.0624	-0.0177	-0.9973	-0.035
PC4	-0.0091	0.0043	-0.0357	0.9993

3.2 小波包变换

对各蚀变主分量影像利用小波函数 db1 进行 4 层小波包分解,其中低频部分提取影像的光谱信息, 高频部分提取影像的纹理信息。通过代价函数选择 最优小波包树(图 6),其中,铁染、Al-OH 及 Mg-OH 主分量影像的最优小波包树的结点数分别为 113、89 及 121。树中被合并的结点索引序号如表 4。

各蚀变信息样本的小波包统计量的均值如表 5。计算每一层分解结果的分类向量,利用 RF 进行 不同尺度的蚀变信息提取,根据提取结果的精度择 优。构造分类向量时,要将各统计量归一化: $y^i = x^i/x^i_{max}$,其中, x^i 为分类向量的第i个分量, y^i 为 第i个分量归一化结果, x^i_{max} 为第i个分量最大值。



图 6 通过代价函数选择的各蚀变信息的最优小波包树 The optimal wavelet packet tree of each alteration information selected by cost function

(a)—铁染;(b)—Al-OH;(c)—Mg-OH

(a)-Iron-stain;(b)-Al-OH;(c)-Mg-OH

表 4 被合并的结点索引序号

Table 4 Number of merged nodes

铁染	80	79	78	77	76	75	73	72	71	69	59	57	56	55	53	48	47	
	46	45	44	43	40	39	28	27	24	23	20	16	15	12	8	7		
AL-OH	72	71	70	69	64	63	62	61	60	59	56	55	53	48	47	46	45	40
	39	36	35	34	33	32	31	30	29	28	27	24	23	16	12	10	4	
Mg-OH	63	62	61	60	55	31	24	23	16	4	2							

		Table 5 Wreat value of v	wavelet packet statistics of	cach alteration sample	
蚀变类型	统计量	1 层分解	2 层分解	3 层分解	4 层分解
	En	-4.9577×10^{11}	-3.1859×10^{11}	-1.0711×10^{11}	-7.3787×10^{10}
	М	-0.0347	0.0611	-0.3527	-1.9510
铁染	V	180.2453	585.6054	2708.2	12205.2
	Е	9.7641	17.4124	36.6449	77.3027
	S	265.0264	529.9511	$\widehat{\mathbf{W}}$ $3 \not\in \Delta \widehat{\mathbf{W}}$ $4 \not\in \Delta$ $\widehat{\mathbf{W}}$ $3 \not\in \Delta \widehat{\mathbf{W}}$ $4 \not\in \Delta$ $\times 10^{11}$ -1.0711×10^{11} -7.378 1 -0.3527 -1.9 54 2708.2 1220 24 36.6449 77.3 11 1059.5 2117 $\times 10^{11}$ -1.4282×10^{11} -1.163 59 -0.2969 -1.7 72 2354.02 1094 04 34.8553 72.4 13 1197.7 2394 $\times 10^{11}$ -1.1019×10^{11} -5.090 2 -0.1137 2.16 16 3183.52 9471 56 42.9211 70.4 93 910.6320 1820	2117.4
	En	-6.1314×10^{11}	-3.0182×10^{11}	-1.4282×10^{11}	-1.1631×10^{11}
	М	0.0553	-0.0359	-0.2969	-1.7780
Al-OH	V	36.9241	527.8372	2354.02	10945.8
	Е	3.9870	16.4594	34.8553	72.4237
	S	299. 3707	598.7713	1197.7	2394.5
	En	-3.6840×10^{11}	-2.3624×10^{11}	-1.1019×10^{11}	-5.0907×10^{10}
	М	0.0029	0.1452	-0.1137	2.1672
Mg-OH	V	1.0845	879.1616	3183.52	9471.38
	Е	0.4976	21.9756	42.9211	70.4202
	S	227.6074	455.2493	910.6320	1820.7

表 5 各蚀变样本的小波包统计量的均值

Table 5	Mean value of	wavelet p	backet statistics	of each	alteration	sample

计算各层分解之后蚀变信息样本的均值与背景均值的欧式距离,用 M-L 表示,其中 M 表示各背景均值,L 表示各层级分解目标矿物样本均值, L=1,2,3,4(表 6)。结果表明,随着分解层级增 多,各目标矿物与背景的距离先增大后减小,进行 3 层级分解时距离最大,所以初步判定 3 层级小波 包分解结果分类效果最佳,后面根据分类精度进 一步判定。

表 6 小波包分解各层统计量均值与背景均值的欧式距离

 Table 6
 Euclidean distance between the mean value of each level statistics and the background mean

value decomposed by wavelet packet

蚀变矿物	M-1	M-2	M-3	M-4
铁染	0.3243	0.5478	1.2902	0.8027
Al-OH	0.2971	0.4879	1.2848	0.8599
Mg-OH	0.4002	0.4432	1.2905	0.8752

3.3 RF 提取

3.3.1 特征筛选

通过 OOB 检测误差率,进行特征的重要程度 归一化排名(图 7)。主成分分析突出了目标矿物的 光谱特征,光谱特征的重要程度最高;能量可以表征 矿物的时空分布可能的变化程度;岩性的组成成分 较为复杂,目标矿物在区域内的分布规律性差,信息 熵可以衡量某区域目标信息量的复杂程度及大小。 根据重要程度排名,最终选择光谱、能量、熵特征进 行矿物提取。

3.3.2 树数选择

用选定的特征子集构建验证树(从 50 到 500, 步长为 50)(图 8)。通过 OOB 验证精度,各矿物提



图 7 RF 分类时所有输入特征的重要程度排名 Fig. 7 Importance ranking of all input features in RF classification

取精度最高时对应的树数即为最佳树数。各矿物提取精度最高时,对应的树数都不同,铁染、Al-OH及Mg-OH精度最高为 0.8843、0.8544及 0.89,对应的最佳树数分别为 200、200及 300。树的棵树引起的精度变化幅度分别为 0.0441、0.0393及 0.0463。



3.4 结果及分析

3.4.1 提取结果

提取结果(图 9)显示,铁染异常主要出现在太 古宙和长城纪地层中,少量出现在南部石炭纪花岗 岩中,零星出现在二叠纪、侏罗纪和白垩纪地层中, 整体呈零散的点状或团块状分布,未见明显的环状



图 9 蚀变信息提取结果示意图 Fig. 9 Alteration information extraction results (a)—铁染蚀变提取结果;(b)—Al-OH 蚀变提取结果;(c)—Mg-OH 蚀变提取结果

(a)-Extraction results of iron stain alteration; (b)-extraction results of Al-OH alteration; (c)-extraction results of Mg-OH alteration

或带状分布特征,与区域构造线不协调。Al-OH 异 常主要分布在北部的前寒武纪地层区和侵入岩周 围,零星出现在南部,呈条带状或似环状。前寒武纪 特别是太古代和长城纪地层中含较多的绢英岩化, 与区域变质作用关系密切,呈条带状分布,与区域构 造线方向一致。碳酸盐岩与酸性侵入岩接触界线附 近常见矽卡岩化,该类蚀变呈似环状或线状,与区域 构造线不协调。Mg-OH 和 CO²⁻异常在研究区北 部和中部广泛分布,侵入岩和南部太古代地层中有 零星出露,多呈条带状,局部呈零星团块状。长城纪 和蓟县纪大理岩、白云岩及太古代角闪岩出露区的 Mg-OH 异常信息与区域构造线方向一致。绿帘石 化热液蚀变规律不明显,形成了零散的点状或团块 状区域。总之,提取的蚀变信息与区域地质背景相 吻合。本文选取了 42 个野外样品点实地验证(图 10),验证结果表明,该区铁染异常与碱性热液蚀变 有关(图 11a),在酸性石英脉中偶见。Al-OH 异常 常见有绢英岩化和砂卡岩化,绢英岩化出露在前寒 武系特别是长城系和太古宇中(图 11b、c), 砂卡岩 化出露在蓟县纪大理岩与二叠纪侵入岩接触界线附 近(图 11d)。Mg-OH 和 CO²⁻主要异常类型有碳酸 盐化、绿泥石化和绿帘石化(图 11e、f)、角岩化(图 11g)以及阳起石化(图 11h)等。

3.4.2 精度评价

对各小波包分解层级使用 RF 提取各蚀变信息,随着分解层级增加提取精度先升高后降低,在3 层级分解时的精度达到最高,与前面选定结果一致(表7)。铁染、Al-OH 及 Mg-OH 在3 层级分解的 提取 精度分别为 88.7443、85.5469 及 91.7594, Kappa 分别为 0.7767、0.6732 及 0.8362。可见,分 解层数要根据目标矿物来确定。



图 10 矿化蚀变信息提取结果野外验证点 Fig. 10 Field verification points of mineralization alteration information extraction results

3.4.3 影响本文方法的因素分析

影响本文蚀变信息提取准确性的因素有:FPCS时 诊断性波段的选择;先验样本集的选择;噪声干扰。

最终进行蚀变信息提取的影像为 FPCS 之后选择的主分量影像。FPCS 时诊断性波段的选择会影响提取结果。如铁染蚀变矿物有赤铁矿、褐铁矿、针铁矿、黄钾铁矾、磁铁矿、角闪石等,波谱特征会因铁离子的价态及矿物质的透明度和含水性而不同。如赤铁矿在 0.66 μ m 和 0.85 μ m 处有强吸收谷,对应ASTER Band 3,而角闪石因含有大量 Fe²⁺而在 2.3 μ m 波长处吸收谷较强,对应 ASTER Band 8,提取这两种铁染蚀变信息时需选择不同的波段组合。实际应用中应对各蚀变类型进行精细划分选择波段,也可进行二次主成分分析来突出目标矿物。



图 11 部分矿物照片及薄片鉴定结果

Fig. 11 Photos and thin section identification results of some minerals

(a)一长石石英岩中的铁染异常;(b)一绢英岩化长英质片岩;(c)一绢英岩化长英质片岩中的绢云母显微照片;(d)一大理岩中的砂卡岩; (e)一绿泥石化斜长角闪岩;(f)一钠长绿泥石英千枚岩中的绿泥石和绿帘石显微照片;(g)一长英质角岩;(h)一蛇纹石化白云质大理岩中的 阳起石显微照片

(a)—Iron staining anomaly in feldspar quartzite; (b)—sericitized feldspar quartz schist; (c)—micrograph of Sericite in sericitized feldspar quartz schist; (d)—skarn in marble; (e)—chloritized plagioclase amphibolite; (f)—micrographs of chlorite and epidote in albite chlorite quartz phyllite; (g)—feldspar quartz hornfels; (h)—micrograph of actinolite in serpentinized dolomitic marble

表 7 小波包分解不同层级的提取精度

Table 7 Extraction accuracy of different levels of waveletpacket decomposition

油亦米刑	1 层分解		2 层分解		3 层分角	解	4 层分解		
因文矢室	OA(%)	Kappa	OA(%)	Kappa	OA(%)	Kappa	OA(%)	Kappa	
铁染	85.5213	0.7427	86.9313	0.7556	88.7443	0.7767	87.8423	0.7447	
Al-OH	82.3239	0.7192	84.2339	0.7212	85.5469	0.7632	84.4849	0.6811	
Mg-OH	87.5364	0.7522	89.1464	0.8101	91.7594	0.8362	90.8574	0.8172	

表 8 各蚀变信息取不同样本时的分类精度

Table 8 The classification accuracy of different samples for each alteration information

蚀变类型	样本 1(%)	样本 2(%)	样本 3(%)	样本 4(%)	样本 5(%)	样本 6(%)	样本 7(%)	样本 8(%)	样本 9(%)
铁染	88.1365	85.6876	88.7443	86.8413	88.0371	85.6690	86.8212	86.6597	88.2136
Al-OH	84.8844	85.3519	85.5469	85.2244	84.0730	83.9970	84.4730	84.2949	84.5349
Mg-OH	90.2337	89.1221	91.7594	89.1169	89.9494	88.3340	90.9979	91.3992	88.3919

监督分类法基于先验样本集,本文采用 RF 分 类时的样本集是基于人工样本点随机产生的不规则 多边形样本区域,样本点的变化会引起分类结果的 动态变异。可通过样本可分离度(TD)(当 TD 值都 大于 1.9 时,样本可分离性较好)评估样本的可靠 性。表 8 为 3 层分解时不同样本引起的提取精度的 动态变化。铁染、Al-OH 及 Mg-OH 样本引起的提 取精度波动范围分别为 3.0753、1.5499 及 3.4254, 在实际应用中可采用结果投票法或者平均法来消除 样本的影响。

本文方法是基于像素的,岩石矿物组分的复杂 性会导致提取结果噪声较多。由于矿物分布具有丛 集特征,实际应用中应该考虑矿物分布的区域特征, 采用基于多尺度分割的面向对象的方法进行提取。

4 结论

(1)本文提出了结合小波包分解和 RF 的应用 于 ASTER 数据的矿化蚀变信息提取方法。

(2)实验结果表明,在该研究区 ASTER 数据中分别选取 Band 1、2、3、4,Band 1、3、4、6 及 Band 1、5、8、9 进行 FPCS,可以最大程度增强铁染、Al-OH 基团及 Mg-OH 和 CO₃²⁻ 基团的蚀变信息。

(3)小波包分解能充分利用高分辨率影像中岩石 的多尺度细节信息,最优小波包树剔除了冗余节点。

(4) RF 对矿物组分引起的噪声不敏感,且可筛选特征,构建最优分类向量。

(5)ASTER 数据蚀变信息提取结果与 42 个点 验证结果一致,说明蚀变信息提取方法可靠,获得的

多个蚀变信息集中区,为矿化线索的发现指明了方向,对矿产勘探具有很高的参考价值。

致谢:野外工作得到了中国地质调查局西安地 质调查中心孟勇等人的支持和帮助;在论文修改过 程中,评审专家和编辑老师提出了宝贵意见,在此一 并表示诚挚的感谢!

References

- Abbaszadeh M, Hezarkhani A. 2013. Enhancement of hydrothermal alteration zones using the spectral feature fitting method in Rabor area, Kerman. Iran. Arab. J. Geosci, 6: 1957~1964.
- Abdi M,Karimpour M. 2012. Application of spectral angle mapper classification to discriminate hydrothermal alteration in southwest Birjand, Iran, using advanced spaceborne thermal emission and reflection radiometer image processing. Acta Geol. Sin, 86: 1289~1296.
- Fallatah A, Jones S, Mitchell D. 2020. Object-based random forest classification for informal settlements identification in the Middle East: Jeddah a case study. International Journal of Remote Sensing, 41(11): 4421~4445.
- Feng Wenqing, Sui Haigang, Tu Jihui, Sun Kaimin, Huang Weiming. 2017. Change detection method for high resolution remote sensing images using random forest. Acta Geodaetica et Cartographica Sinica, 46(11): 1880~1890. (in Chinese)
- Fu Shuixing, Zhang Shoulin, Li Chunxia, Feng Jianzhong, Fang Tonghui. 2004. Remote sensing technology for identification of alteration information of gold deposits in the eastern Tianshan area, Xinjiang. Acta Geologica Sinica, 78(2): 417~422.
- Gong Hao, Zhang Jingxiong, Sheng Shaohong. 2009. Object-based correspondence analysis for improves accuracy in remote sensing change detection. Geomatics and Information Science of Wuhan University, 34(5): 544~551. (in Chinese)
- He Xianjin, Zhang Xinchang, Xin Qinchuan. 2018. Recognition of building group patterns in topographic maps based on graph partitioning and random forest. ISPRS Journal ofPhotogrammetry and Remote Sensing, 136:26~40.
- Honarmand M, Ranjbar H, Shahabpour J. 2012. Application of principal component analysis and spectral angle mapper in the mapping of hydrothermal alteration in the Jebal Barez area, southeastern Iran. Resource Geology, 62(2):119~139.
- Hosseinjani M, Tangestani M. 2011. Mapping alteration minerals using sub-pixel unmixing of ASTER data in the Sarduiyeh area, SE Kerman, Iran. Int. J. Dig. Earth, 4: 487~504.
- Huang Shuang, Chen Shengbo, Zhou Chao, Zha Fengli, Lu Peng, Huang Weishi. 2015. The Application of spectral characteristic space method to the alteration information extraction of lake Superior-type iron deposit. Earth Science-Journal of China Universit y of Geosciences, 40(8): 1325~1329. (in Chinese)
- Kaur S, Bansal R K, Mittal M, Oyal L M, Kaur I, Verma A, Son L H. 2019. Mixed pixel decomposition based on extended fuzzy clustering for single spectral value remote sensing images. Journal of The Indian Society of Remote Sensing, 47(3):427~437.
- Li Junjian, Tang Wenlong, Fu Chao, Chen Zheng, Orolmaa Demberel, Oyuntuya Namsraijavyn, Delgersaikhan Adiya, Enkhbat Tserendash, Dang Zhicai, Zhao Zelin, Zhang Feng, Ren Junping, Zhao Lijun. 2016. The division of metallogenic belts in Sino-Mongolian border area. Geological Bulletin of China, 35(4):463~487. (in Chinese)
- Li Zhizhong, Wang Daming, Liu Dechang, Liu Yinnian, Zhao Huijie, Dang Fuxing. 2015. Hyperspectral remote sensing technology and its progress in resources exploration. Earth Science—Journal of China University of Geosciences, 40(8): 1287~1294. (in Chinese)
- Liu Dechang, Tian Feng, Qiu Junting, Ye Fawang, Yan Bokun, Sun Yu, Wang Zitao. 2017. Application of hyperspectral

remote sensing in solid ore exploration in the Liuyuan-Fangshankou area. Acta Geologica Sinica, 91 (12): 2781 \sim 2795. (in Chinese)

- Liu Jianyu, Chen Ling, Li Wei, Wang Genhou, Wang Bo. 2019. An improved method for extracting alteration related to the ductile shear zone type gold deposits using ASTER data. Remote Sensing for Land and Resources, 31(1): $229 \sim 236$. (in Chinese)
- Loughlin W P. 1991. Principal component analysis for alteration mapping. Photogrammetric Engineering & Remote Sensing, 57 (9):1163~1169.
- Mahboob M A, Genc B, Celik T, Ali S, Atif I. 2019. Mapping hydrothermal minerals using remotely sensed reflectance spectroscopy data from Landsat. Journal of The Southern African Institute of Mining And Metallurgy, 119(3): 279~289.
- Moore F, Rastmanesh F, Asadi H, Modabberi S. 2008. Mapping mineralogical alteration using principal-component analysis and matched filter processing in the Takab area, north-west Iran, from ASTER data. Int. J. Remote Sens, 29: 2851~2867.
- Sankar A S, Nair S S, Dharan V S, Sankaran P. 2015. Wavelet sub band entropy based feature extraction method for BCI. Procedia Computer Science, 46: 1476~1482.
- Seo Y, Kim S. 2016. River stage forecasting using wavelet packet decomposition and data -driven models. Procedia Engineering, 154,1225~1230.
- Uddin M P, Al Mamun M A, Hossain M A. 2020. PCA-based feature reduction for hyperspectral remote sensing image classification. Iete Technical Review, $(5): 1 \sim 21$.
- Vignesh K M, Kiran Y. 2020. Comparative analysis of mineral mapping for hyperspectral and multispectral imagery. Arabian Journal of Geosciences, 13(4):160.
- Wang Runsheng, Xiong Shengqing, Nie Hongfeng, Liang Shuneng, Qi Zerong, Yang Jinzhong, Yan Bokun, Zhao Fuyue, Fan Jinghui, Tong Liqiang, Lin Jian, Gan Fuping, Chen Wei, Yang Suming, Zhang Ruijiang, Ge Daqing, Zhang Xiaokun, Zhang Zhenhua, Wang Pinqing, Guo Xiaofang, Li Li. 2011. Remote sensing technology and its application in geological exploration. Acta Geologica Sinica, 85(11): 1700~1742. (in Chinese)
- Wu Zhichun, Ye Fawang, Guo Fusheng, Liu Wenheng, Li Hualiang, Yang Yi. 2018. A review on application of techniques of principle component analysis on extracting alteration information of remote sensing. Journal of Geoinformation Science, 20(11): 1644~1656. (in Chinese)
- Xie Minghui, Zhang Qi, Chen Shengbo, Li Yanqiu, Zha Fengli. 2015. Extraction of alterationanomaly information by featurebased principal component analysis from ASTER data. Earth Science—Journal of China University of Geosciences, 40(8): 1381~1385. (in Chinese)
- Xu Kai, Wang Xiaofeng, Kong Chunfang, Feng Ruyi, Liu Gang, Wu Chonglong. 2019. Identification of hydrothermal alteration minerals for exploring gold deposits based on SVM and PCA using ASTER data: A case study of Gulong. Remote Sensing, 11(24): 3003.
- Zhang Lin, Liu Zhe, Ren Tianwei, Liu Diyou, Ma Zhe, Tong Liang, Zhang Chao, Zhou Tianying, Zhang Xiaodong, Li Shaoming. 2020. Identification of seed maize fields with high spatial resolution and multiple spectral remote sensin using Random Forest Classifier. Remote Sensing, 12(3):362.
- Zhang Ruisi, CaoHui,Zeng Min,Zhang Dexian, Li Guangxu. 2016. Scientific belt lithologic mapping based on ASTER spectral analysis: A case study from the Tashkorgan area (Xinjiang, China). Acta Petrologica Sinica, 32(12), 3835 ~ 3846. (in Chinese)
- Zhang Tingbin, Yi Guihua, Li Hongmei, Wang Ziyi, Tang Juxing, Zhong Kuanghui, Li Yubin, Wang Qing, Bie Xiaojuan. 2016. Integrating data of ASTER and Landsat-8 OLI (AO) for hydrothermal alteration mineral mapping in Duolong Porphyry Cu-Au deposit, Tibetan Plateau, China. Remote Sensing, 8 (12): 890.

- Zhao Yingjun, Qin Kai, Sun Yu, Liu Dechang, Tian Feng, Pei Chengkai, Yang Yanjie, Yang Guofang, Zhou Jiajing. 2015. Progress of Geological survey using airborne hyperspectral remote sensing data in the Gansu and Qinghai regions. Acta Geologica Sinica (English Edition), 89(5): 1783~1840.
- Zhu Weiping, Zhang Yongmei. 2014. Application of geophyusical and remote sensing methods to predict for potash resource. Acta Geologica Sinica (English Edition), 88(1): 289~290.

参考文献

- 冯文卿,眭海刚,涂继辉,孙开敏,黄伟明. 2017. 高分辨率遥感影像 的随机森林变化检测方法. 测绘学报,46(11):1880~1890.
- 龚浩,张景雄,申邵洪. 2009. 基于对象的对应分析在高分辨率遥感 影像变化检测中的应用.武汉大学学报(信息科学版),34(5): 544~551.
- 黄爽,陈圣波,周超,查逢丽,路鹏,黄唯实. 2015. 基于光谱特征空间的苏必利尔湖型铁矿高光谱遥感蚀变信息提取方法. 地球科学——中国地质大学学报,40(8):1325~1329.
- 李俊建,唐文龙,付超,陈正,Orolmaa Demberel,Oyuntuya Namsraijavyn,Delgersaikhan Adiya,Enkhbat Tserendash,党智 财,赵泽霖,张锋,任军平,赵丽君. 2016. 中蒙边界地区成矿区 带划分. 地质通报,35(4):463~487.

- 李志忠,汪大明,刘德长,刘银年,赵慧洁,党福星. 2015. 高光谱遥 感技术及资源勘查应用进展. 地球科学——中国地质大学学 报,40(8): 1287~1294.
- 刘德长,田丰,丘骏挺,叶发旺,闫珀琨,孙雨,王子涛. 2017. 柳园一 方山口地区航空高光谱遥感固体矿产探测及找矿效果. 地质学 报,91(12):2781~2795.
- 刘建宇,陈玲,李伟,王根厚,王博. 2019. 基于 ASTER 数据韧性剪 切带型金矿蚀变信息提取方法优化. 国土资源遥感,31(1): 229~236.
- 王润生,熊盛青,聂洪峰,梁树能,齐泽荣,杨金中,闫柏琨,赵福岳, 范景辉,童立强,林键,甘甫平,陈微,杨苏明,张瑞江,葛大庆, 张晓坤,张振华,王品清,郭小方,李丽. 2011. 遥感地质勘查技 术与应用研究. 地质学报,85(11):1700~1742.
- 吴志春,叶发旺,郭福生,刘文恒,李华亮,杨羿. 2018. 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,20(11):1644~1656.
- 谢明辉,张奇,陈圣波,李艳秋,查逢丽.2015. 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,40(8):1381~1385.
- 张瑞丝,曹汇,曾敏,张德贤,李广旭. 2016. 基于 ASTER 光谱特征 的科技廊带岩性填图:以新疆塔什库尔干地区为例.岩石学报, 32(12): 3835~3846.

Aster alteration information extraction based on wavelet packet transform and random forest

TANG Shulan^{1,2)} , CAO Jiannong $^{\ast\,1)}$, WANG Guoqiang $^{3)}$, BU Tao $^{3)}$

1) School of Earth Science and Resources, Chang'an University, Xian 710054;

2) School of Management, Xian University of Finance and Economics, Xian 710100;

3) Xian Center of Geological Survey, CGS, Xian 710054 * Corresponding author:zyjncao@chd.edu.cn

Abstract

In order to extract alteration information more accurately, this paper selects Xinjiang, Gansu and Inner Mongolia as the research area, and extracts aster alteration information by combining wavelet packet transform and random forest. Firstly, the diagnostic wavebands of main alteration types are selected for feature vector principal component analysis to obtain principal component images. Then, the principal component image is transformed by wavelet packet, and the optimal wavelet packet tree is selected by using the cost function, and the high and low frequency information is extracted to construct the classification vector. Then, the random forest classification model is constructed through feature screening, and the mineralization and alteration information is extracted. Finally, the accuracy of the extraction results is evaluated by field sampling and thin section identification. The principal component analysis band combinations of iron stain, Al-OH and Mg-OH alteration information are Band 1, 2, 3, 4, Band 1, 3, 4, 6 and Band 1, 5, 8 and 9, respectively. The results show that the overall accuracy of the method is 88, 7443, 85, 5469 and 91, 7594, and kappa is 0, 7767, 0, 6732 and 0, 8362, respectively. It has a good correlation with the metallogenic characteristics of the metallogenic belt and the existing area. The optimal wavelet packet tree used in this study can make full use of the energy characteristics of mineral spectrum, and random forest can weaken the noise interference of mineral components. The research results can provide technical reference for remote sensing alteration information extraction.

Key words: aster; alteration information; wavelet packet transform; random forest; principal component analysis