doi:10.6041/j.issn.1000-1298.2017.07.017

基于 MESMA 和 RF 的山丘区土地利用信息分类提取

陈元鹏[1,2]　郧文聚[2]　周　旭[2]　彭军还[1]　李少帅[1,2]　周　妍[2]
(1. 中国地质大学(北京)土地科学技术学院, 北京 100083; 2. 国土资源部土地整治中心, 北京 100035)

摘要: 探讨了基于多端元混合像元分解(Multiple endmember spectral mixture analysis,MESMA)和随机森林(Random forest,RF)相结合的土地利用信息分类提取方法。以 Landsat-8 OLI 卫星遥感影像为主要数据,基于植被-不透水面-裸土(Vegetation-impervious surface-soil,VIS)模型,利用 MESMA 将影像分解为植被、不透水面和裸土 3 类组分,将生成的 3 类组分变量和基于光谱、纹理信息计算选取的 20 个特征变量组合后开展 RF 分类实验,将分类结果与相同特征变量下的支持向量机(Support vector machine,SVM)、最大似然(Maximum likelihood classification,MLC)分类结果进行比较分析。结果表明:MESMA 可以获得较为精确的组分丰度信息;RF 分类结果优于相同特征变量下的 SVM 和 MLC 分类结果;在 MESMA 生成的组分信息变量参与分类后,3 种方法的分类精度均有所改善,分别达 90.50%、88.85%、86.35%,其中 RF 的分类精度改善最为显著;MESMA 与线性混合分解(Linear spectral mixture analysis,LSMA)生成的组分信息变量相比,前者对于改善分类精度效果更为明显。MESMA 对于提高影像分类精度起到一定积极作用,基于 MESMA 和 RF 的方法对中等空间分辨率影像山丘区土地利用信息分类提取精度较高,利用该方法开展遥感影像解译可为大尺度的土地利用监测和管理工作提供技术支持和理论参考。

关键词: 土地利用分类; 中等空间分辨率; 多端元混合像元分解; 随机森林; 山丘区
中图分类号: TP79; S127　　**文献标识码**: A　　**文章编号**: 1000-1298(2017)07-0136-09

Classification and Extraction of Land Use Information in Hilly Area Based on MESMA and RF Classifier

CHEN Yuanpeng[1,2]　YUN Wenju[2]　ZHOU Xu[2]　PENG Junhuan[1]　LI Shaoshuai[1,2]　ZHOU Yan[2]
(1. *School of Land Science and Technology*, *China University of Geosciences* (*Beijing*), *Beijing* 100083, *China*
2. *Land Consolidation and Rehabilitation Center*, *Ministry of Land and Resource*, *Beijing* 100035, *China*)

Abstract: Due to the factors such as sensor spatial resolution and heterogeneity of surface features, the mixed-pixels were commonly found in medium-spatial resolution remote sensing data, especially in hilly areas, strong topographic relief, diversity, breakage, mixed distribution and scattered layout of the surface features and other factors constituted the difficulties of remote-sensing image classification mapping. In order to improve the classification accuracy for land use in hilly areas and provide data support for land use monitoring, a combined approach of multiple endmember spectral mixture analysis (MESMA) and random forest (RF) was explored. Based on data source of Landsat-8 operational land imager (OLI) sensor data, the fractional abundance of vegetation, impervious surface and soil was firstly extracted through MESMA. Secondly, totally 20 feature variables were figured out and three combined models were constructed on the basis of data image spectrum, texture and fraction variables to carry out random forest classification experiment. Through comparing between the optimal result from the experiment and SVM and MLC classification results, including the same number of variables, the results indicated that MESMA can derive accurate fraction information. The inclusion of fraction information could help to improve the mapping accuracy of all classification methods (RF, SVM and MLC), which can be up to 90.50%, 88.85% and 86.35%, respectively, the gain of RF classification accuracy was most significant. Comparing with LSMA, the fraction variable generated by MESMA was more useful for improving the accuracy. The combined method of MESMA and RF can achieve the comparatively accurate classification map in the multi-feature variables. The accuracy was better than those of SVM and MLC

收稿日期: 2017-04-28　修回日期: 2017-05-18
基金项目: 国土资源部公益性行业科研专项(201511010-02)
作者简介: 陈元鹏(1988—),男,博士生,国土资源部土地整治中心工程师,主要从事遥感技术和土地复垦研究,E-mail: cyp520918@163.com

classification results with the same feature variables. Therefore, the proposed method can obtain high precision in land use classification in hilly area. Based on this method, remote sensing image interpretation of large scales can provide technical support and rational reference for land reclamation monitoring.

Key words: classification of land use; medium-spatial resolution; multiple endmember spectral mixture analysis; random forest; hilly area

引言

准确掌握国家土地资源利用信息是国家相关部门调整土地利用结构、合理开发土地资源、动态监测土地利用状况等工作的基础[1]。土地资源与遥感数据在时空特性方面具有高度一致性，对土地资源的研究成为遥感技术应用的主要领域之一。过去数十年，国内外开展了大量研究，促使了遥感技术在土地资源领域的长足发展。现今遥感技术已为土地利用的信息提取与分类识别提供了丰富的信息源和多元的实现手段[2]，为准确掌握国家土地利用信息贡献了巨大力量。

在大尺度的土地资源遥感研究应用中，兼顾经济、效率的同时，中等空间分辨率数据往往成为较佳选择，但在地物异质性和传感器空间分辨率影响下，中等空间分辨率数据普遍存在混合像元，尤其在我国山地丘陵区，地形起伏大、地物分布破碎，导致影像数据中混合像元占比更高，这一定程度上影响了地物识别和分类精度。为此，充分利用混合像元分解技术，并与目前应用范围较广的机器学习分类方法有效结合[3]，对基于中等空间分辨率影像的山丘区土地利用信息提取精度的提高很有帮助。

多端元混合像元分解是在线性混合分解基础上发展而来的一种混合像元分解方法，MESMA 在引入地物端元波谱库的基础上，充分考虑了像元内地物的不同组合方式及光谱变化性，利用不同类型和数量的地物端元组合模拟每个像元，它针对每一类地物选取多个端元光谱参与混合像元分解，可以有效解决“同物异谱”问题[4-6]。随机森林分类方法目前在基于机器学习的遥感影像分类方法中精度相对较高，处理大数据集时运行速度更快，不容易过拟合，处理多维变量能力强，可生成变量重要性估计，所以该算法在多维数据分类回归中被广泛应用，并取得较好的效果[7]。但与其它的机器学习分类算法相比，国内对于随机森林算法在遥感影像分类中的研究和应用相对较少，已有的研究区域主要集中在平原区[8-10]，在山丘区的研究应用不多。

本文提出一种 MESMA 和 RF 相结合的遥感影像分类方法，利用该方法对四川省古蔺县域内的土地利用信息进行分类提取研究；以中等空间分辨率 Landsat－8 OLI 影像为唯一数据源，不使用辅助数据参与分类；利用 MESMA 方法解混生成植被、不透水面和裸土 3 种组分信息，将组分信息作为特征变量和基于光谱、纹理信息计算选取的 20 个特征变量进行组合，构建 3 种变量组合模型开展 RF 分类实验；对精度最高的模型 3 分类结果进行变量重要性估计，分析各变量的重要性成因；将精度最高的模型 3 分类结果与相同特征变量下的 SVM、MLC 分类结果，基于 LSMA 和 RF 相结合方法的分类结果进行比较分析。

1　研究区概况

研究区位于四川省泸州市古蔺县，地理坐标为 27°59′43″～28°7′24″N、105°55′41″～106°4′47″E，区域内海拔在 340～1 207 m 之间，中亚热带气候，年平均气温 17.1～18.5℃、平均降水量 748.4～1 184.2 mm[11-13]。参照 GB/T 21010—2007《土地利用现状分类》，研究区内主要土地利用类型为耕地、有林地、工矿仓储用地（简称“工矿用地”）、住宅用地、交通运输用地、水域及水利设施用地（简称“水域”）。本文在信息提取分类体系构建中，将住宅用地与交通运输用地归并为一类简称建设用地，为支撑工矿用地复垦利用管理工作，工矿用地作为一种单独地类进行提取实验，而未归并至建设用地类。研究区地理区位及遥感影像数据如图 1 所示。

2　数据选取及预处理

本研究中使用的主要数据为 Landsat－8 OLI 卫星遥感影像，参考数据包括：泸州市工矿废弃地现状数据、无人机航拍影像、地面实测数据、Google Earth 数据。其中，Landsat－8 OLI 卫星遥感影像用于土地利用信息分类提取；4 类参考数据用于样点采集参考和精度评价验证。

2.1　Landsat－8 OLI 卫星遥感影像

Landsat－8 是 Landsat 系列卫星中的一员，发射于 2013 年，它搭载了 2 个主要载荷：陆地成像仪（OLI）和热红外传感器（TIRS）。OLI 包含 9 个波段，其中多光谱波段的空间分辨率为 30 m，全色波段的分辨率为 15 m，成像宽幅为 185 km×185 km。本文选用的 Landsat－8 OLI 遥感数据为 30 m 分辨

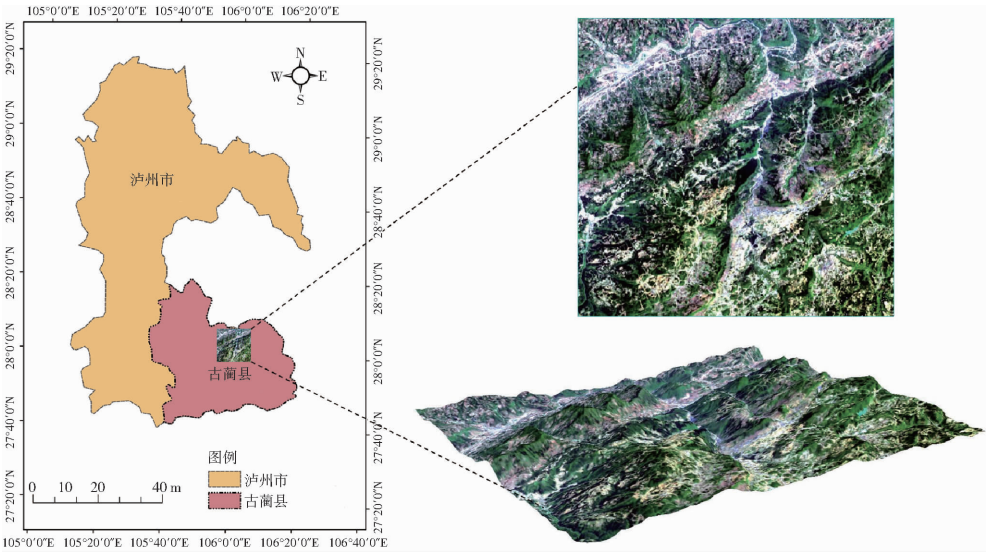


图 1 研究区地理位置和 3D 遥感影像

Fig. 1 Location of study area and 3D representation of remote sensing image

率多光谱波段数据 1 景,波段数为 7,分别为 *C*、*B*、*G*、*R*、NIR、SWIR1、SWIR2,获取日期为 2016 年 4 月 3 日(该时间窗口内植被与裸土可分性较好),研究区域内无云量。影像的预处理在 ENVI 5.3 软件平台中完成,对多光谱数据的预处理包括波段筛选(去除了 Coastal 波段)、辐射定标、FLAASH 模块大气校正、裁剪生成研究区影像数据。

2.2 参考数据

参考数据中,泸州市工矿废弃地现状数据为 1∶200 000 现状图,由泸州市政府于 2012 年编制完成;航拍影像由 UV-II 型无人机搭载的 Pentax-645D 型相机拍摄,拍摄时间为 2016 年 11 月,空间分辨率为 0.2 m;地面实测数据的采集在无人机航拍过程中同步开展,实测仪器采用天宝手持式 GPS,水平精度优于 1 m(Trimble geoexplorer 2008 Series GeoXH, trimble navigation limited, USA)。

2.3 样点选取

根据泸州市工矿废弃地现状数据、地面实测数据、航拍影像的目视解译分类结果以及 Google Earth 影像数据,在 Landsat-8 OLI 影像上选取了训练与验证样点,其中训练样点 706 个,占比 22%;验证样点 2 495 个,占比 78%。各类地物样点数量如表 1 所示。

表 1 各地物类型样点数量

Tab. 1 Numbers of samples for each class

类别名称	训练样点	验证样点	总样点数
有林地	118	1 284	1 402
耕地	252	639	891
工矿用地	134	272	406
建设用地	91	153	244
水域	111	147	258
总数	706	2 495	3 201

3 研究方法

3.1 技术方法

采用 MESMA 和 RF 相结合的方法进行土地利用信息提取分类,技术流程主要为:影像预处理;利用 MESMA 进行混合像元分解;利用 RF 方法进行影像土地利用信息提取分类;精度评价。如图 2 所示。

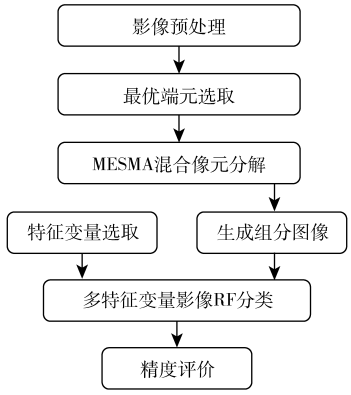


图 2 技术流程

Fig. 2 Flow chart of research technique

具体步骤为:对 Landsat-8 OLI 多光谱影像进行预处理,将航拍影像与卫星影像进行配准、裁剪;基于植被、不透水面、裸土模型[14],利用 MESMA 方法解混生成 3 种组分丰度数据(植被、不透水面、裸土);结合影像的光谱、纹理信息进行特征变量的计算、选取;将组分丰度数据与各类型特征变量组合后进行 RF 分类(利用网格搜索法进行参数寻优)[15-16];对分类结果进行精度评价,利用特征变量重要性估计评价各特征变量的重要程度并探讨原因;与 SVM 和 MLC 分类方法进行比较,评估随机森林分类方法的性能。

3.2　MESMA 方法的混合像元分解

MESMA 是在 LSMA 基础上发展而来的一种混合像元分解方法，该方法针对每一类地物选取多条光谱，并以此生成多个端元组合（每个端元组合由不同地物中的某一条光谱组成），并对每个像元寻找最小二乘法误差最小的端元组合，进而求出每个像元的端元比例，通常采用均方根误差（RMSE）指标来评价模型的拟合优度[17]，公式为

$$R=\sqrt{\frac{\sum_{\lambda=1}^{S}\varepsilon_{\lambda}^{2}}{S}} \tag{1}$$

式中　S——遥感影像波段数

λ——波段　　ε_{λ}——λ 的残差

3.2.1　光谱库建立和端元选取

端元光谱的选择对于生成高精度的组分丰度数据至关重要，本文基于无人机影像和地面实测数据，按照 VIS 建模原则，在影像上选取了植被、不透水面、裸土 3 种地物类型的端元，数量分别为植被端元 606 个、不透水面端元 209 个、裸土端元 271 个，总计 1 086 个，将这些端元的光谱收集并建立光谱库，利用“Viper Tools”计算影像光谱库每条光谱的端元平均均方根误差（Endmember average root mean square error，EAR）[18-19] 和最小平均光谱角（Minimum average spectral angle，MASA）[20]，选择 EAR、MASA 值较低的端元为最优端元。通过比选，最终从 1 086 条端元光谱中选取了 18 条作为最优端元光谱。光谱曲线如图 3 所示。

3.2.2　端元组合模型

文献[21]研究表明，2 端元模型适合自然地物景观的模拟分解，3 端元模型适合人为扰动下的地物景观模拟分解，4 端元模型适合城镇建设区的模拟分解。研究区域内地形起伏、土地利用覆被状况较为复杂、人为扰动因素大，所以本文主要测试了 3 端元、4 端元的组合模型。约束条件分别为：非阴影组分丰度范围在 -0.05 ~ 1.05 之间；最大允许阴影组分为 0.8；最大 RMSE 为 0.025。各端元模型的评价如表 2 所示。

通过表2的解混比例和RMSE的比较结果可

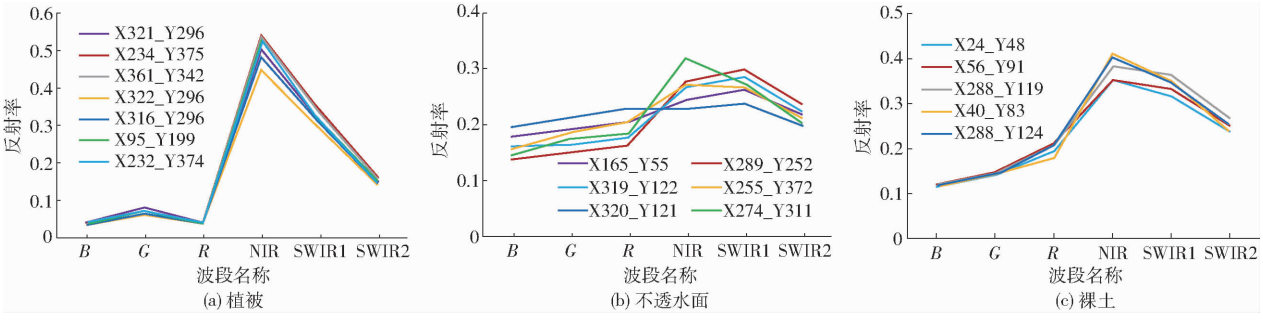


图 3　端元光谱曲线

Fig. 3　Spectral curves of endmember

表 2　端元组合模型评价

Tab. 2　Assessment of combinations of endmember model

类型	模型数量	端元模型组合	解混比例/%	均方根误差
3 端元模型	42	植被 + 不透水面 + 阴影	91.0	0.006 9
	35	植被 + 裸土 + 阴影	86.5	0.006 9
	30	不透水面 + 裸土 + 阴影	18.8	0.002 2
4 端元模型	210	植被 + 不透水面 + 裸土 + 阴影	91.0	0.005 6

知，3 端元组合模型中虽然“植被 + 不透水面 + 阴影”组合与 4 端元模型组合的解混比例相同为 91.0%，但 RMSE 相对较高为 0.006 9，而 3 端元组合模型中“不透水面 + 裸土 + 阴影”的组合模型虽然 RMSE 最低为 0.002 2，但解混比例也最低，只有 18.8%，所以相比较下 4 端元组合模型的解混效果更优，因此本研究选取 4 端元组合模型进行 MESMA 混合像元的分解，通过阴影归一化后生成植被、不透水面、裸土 3 种组分丰度图像。

3.3　特征变量计算与选取

研究基于影像的光谱、纹理信息进行特征变量的计算、选取。光谱特征变量主要包括波段变量（B、G、R、NIR、SWIR1、SWIR2），光谱指数变量包括 NDVI（Normalized difference vegetation index）、BCI（Biophysical composition index）[22-23]、EVI（Enhanced vegetation index）、MNDWI（Modified normalized difference water index）、MNLI（Modified non-linear index）、SAVI（Soil adjusted vegetation index），共 12 个光谱特征变量。

纹理特征变量选取时，经多次试验对比分析，选用 3 × 3 大小的移动窗口，利用灰度共生矩阵（GLCM）分别计算影像 6 个光谱波段变量的 8 种纹

理特征[24]：均值(Mean)、方差(Variance)、同质性(Homogeneity)、对比度(Contrast)、差异性(Dissimilarity)、熵(Entropy)、二阶矩(Second moment)、相关性(Correlation)，计算后得到 48 个纹理特征变量，因 48 个变量之间相关性较高，利用主成分变换(PCA)对其进行降维，以标准差值 7 为阈值正向排序择优选取了前 8 个主成分变量(PC1 ~ PC8)参与影像分类，合计 8 个纹理特征变量。

结合前述的 MESMA 解混生成的植被、不透水面、裸土组分特征变量，总计 23 个特征变量参与 RF 影像分类，所有的特征变量如表 3 所示。

表 3 特征变量统计

Tab. 3 Statistic of feature variables

类别	特征变量	数量
光谱特征	*B*、*G*、*R*、NIR、SWIR1、SWIR2	6
光谱指数特征	NDVI、BCI、EVI、MNDWI、MNLI、SAVI	6
纹理特征	PC1 ~ PC8	8
组分特征	植被、不透水面、裸土	3
合计		23

3.4 随机森林分类算法

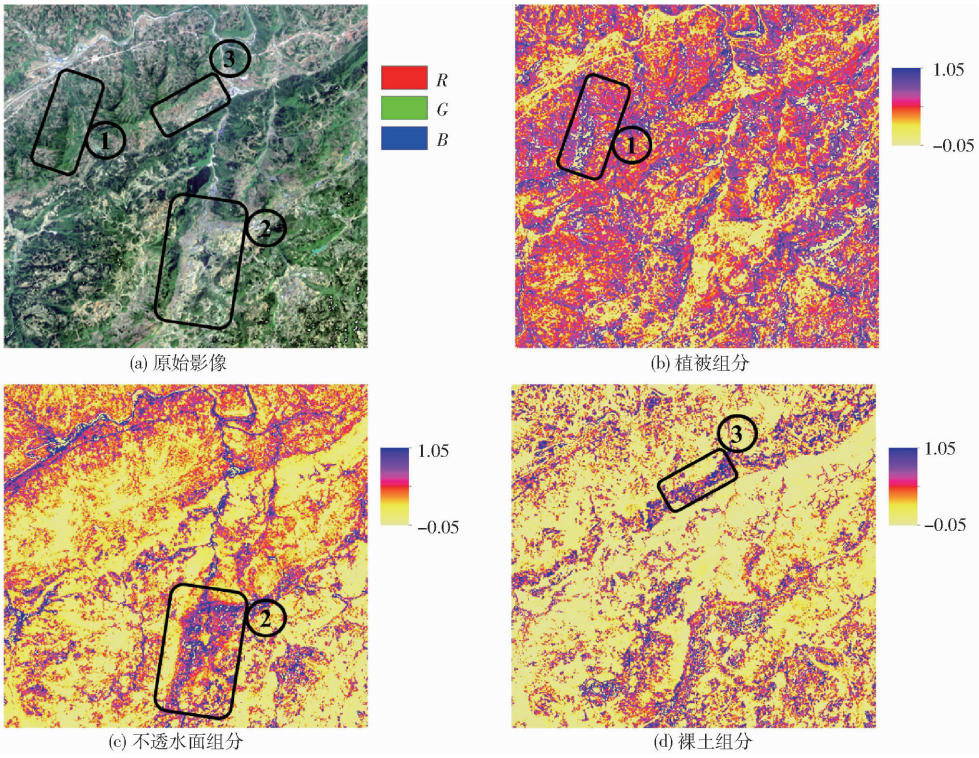
随机森林分类算法是由决策树组合构成的机器学习算法，该算法适用于处理高维数据且运行速度相对较快[25]，其执行过程如下：首先采用随机 Bootstrap 方法从原始数据中有放回的抽取 N 组训练集，每组训练集的大小约为训练数据集的 2/3，这一过程称为 Bagging；然后利用 N 组训练集构建 N 棵决策树，在每棵树生长过程中，从全部 M 个特征变量中随机抽选 m 个($m \leqslant M$)进行内部节点划分；最后，集合 N 棵决策树的预测结果，采用投票的方式决定新样本的类别。在训练数据抽取过程中，约 1/3 的数据未被抽中，这些数据称为袋外数据(OOB)，OOB 可用于类别错分的误差评估和变量重要性估计，在变量选取过程中采用了 Gini 系数测量变量的不纯度。参数设置中通常树的默认数量(ntree)为 100，变量默认数量(mtry)为影像波段总数的平方根[26-29]。

本文通过特征变量选取构建了 3 种不同的特征变量组合模型进行随机森林分类实验，模型 1：多光谱特征模型(记为 MS)；模型 2：多光谱、纹理特征组合模型(记为 MS + TXT)；模型 3：多光谱、纹理、组分特征组合模型(记为 MS + TXT + EM)，测评各类型特征变量对影像分类的贡献度，选出最佳的分类方案。随机森林 ntree 和 mtry 的参数寻优采用了基于 OOB 误差值的网格搜索法，随机森林分类算法与网格搜索算法均在 IDL8.5 语言平台实现。

4 结果

4.1 多端元混合像元分解结果

由 MESMA 方法生成的植被、不透水面、裸土组分丰度图如图 4 所示，组分丰度范围在 -0.05 ~ 1.05 之间，丰度越高越接近该类地物。可以看出，与真彩色影像(图 4a)相比较，植被(图 4b)、不透水



(a) 原始影像 (b) 植被组分

(c) 不透水面组分 (d) 裸土组分

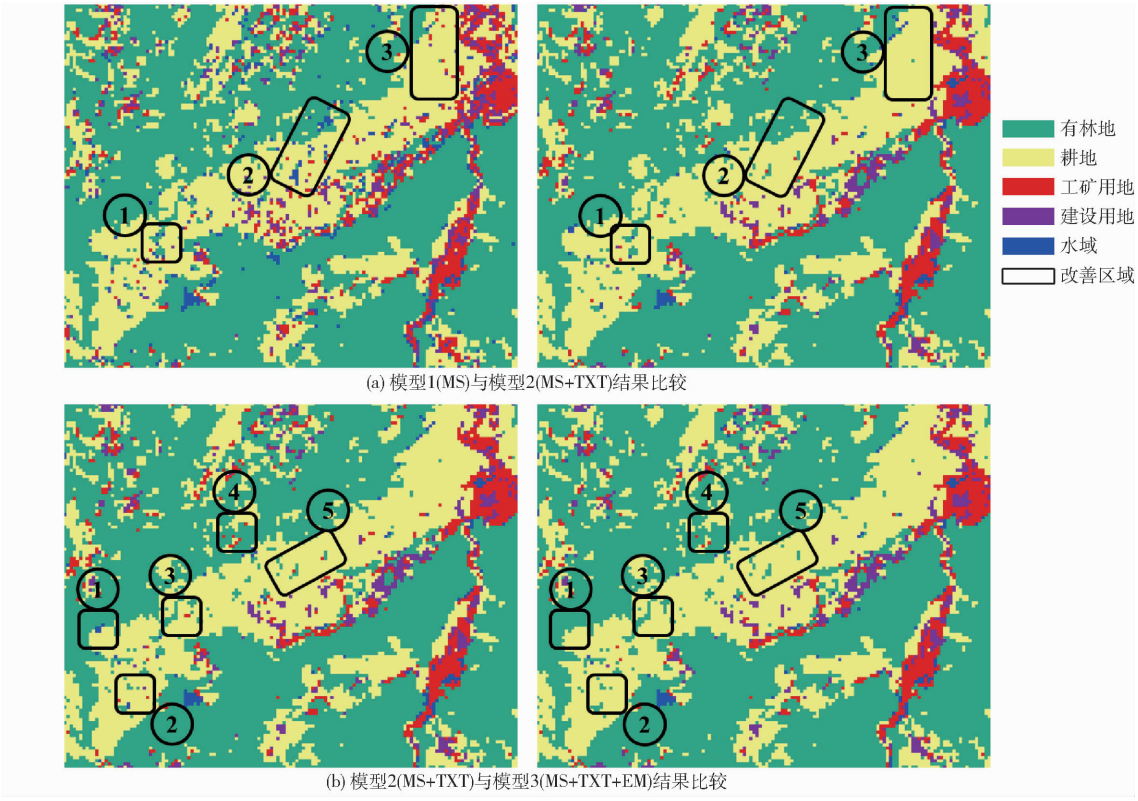
图 4 MESMA 组分丰度图

Fig. 4 MESMA fraction maps

面(图 4c)和裸土(图 4d)的组分丰度分布边界清晰,与研究区实际情况基本一致,如图 4 位置①中的植被、位置②中的不透水面和位置③中的裸土等。

4.2　随机森林分类结果与精度评价

利用基于 OOB 误差值的网格搜索法在 ntree 范围 1 ~600 内对 3 个模型进行参数寻优,3 个模型的 mtry 和 ntree 最优参数分别为(3,102)、(4,188)和(5,126)[30-31],利用以上参数分别执行随机森林分类算法得到分类结果,3 个模型分类总体精度分别为 85.36%、89.48%和 90.50%。由变化幅度看,多光谱特征模型在增加了纹理变量后,精度提升幅度最大为 4.12%,而在加入组分特征变量后分类精度得以进一步提高。各模型分类图的局部效果如图 5 所示。



(a) 模型1(MS)与模型2(MS+TXT)结果比较

(b) 模型2(MS+TXT)与模型3(MS+TXT+EM)结果比较

图 5　分类结果局部效果对比

Fig. 5　Comparisons of local classification result

从图 5 视觉效果看,图 5a 中模型 2 的分类结果明显优于模型 1,模型 1 位置①、③中错分为工矿用地的像素、位置②中错分为水域的像素,相应的在模型 2 中都得到了较好改善;模型 1 结果椒盐效果较为明显,而在模型 2 中增加了纹理特征变量后,椒盐效果消除显著,各地类的分布更加均匀连片。模型 3 与模型 2 结果相比较,分类精度进一步提高,如图 5b 所示,模型 2 中标记位置的错分像素点相应在模型 3 结果中有所改善,椒盐效果得以进一步消除,图面更加整洁美观。

3 个模型中,模型 3 分类精度最高达 90.50%,混淆矩阵见表 4,其中精度在 80% 以下的为工矿用地和建设用地,分别为 79.41% 和 67.97%,与林地、耕地、水域相比精度相对较低,说明实验方法在区分工矿和建设用地中还存在一定局限性,但较之模型 1 的 51.84% 和 51.63%,分别提高 27.57% 和 16.34%,可知在随机森林分类方法框架下,增加了纹理变量和组分变量后,影像的分类精度提高显著。

表 4　模型 3 精度评价

Tab. 4　Accuracy assessment of model 3　%

土地类型	有林地	耕地	工矿用地	建设用地	水域
未分类	0	0	0	0	0
有林地	98.29	3.60	0.37	0.65	0
耕地	0.78	85.92	8.82	10.46	0
工矿用地	0.16	6.89	79.41	16.99	12.93
建设用地	0	2.19	8.82	67.97	0.68
水域	0.77	1.40	2.57	3.92	86.39
总和	100	100	100	100	100

4.3　特征变量重要性估计

如前所述,随机森林分类方法的优势之一是可以评估特征变量的重要性和贡献度。图 6 表示根据 OOB 误差得出的模型 3 的特征变量重要性排序,由图可知排序靠前的特征变量主要为原始光谱特征变量和光谱指数变量,其中 R、SWIR2、MNDWI、B 4 个特征变量重要性值均大于 0.8,排序最为靠前,原因是研究区域内林地植被面积所占比重最高达

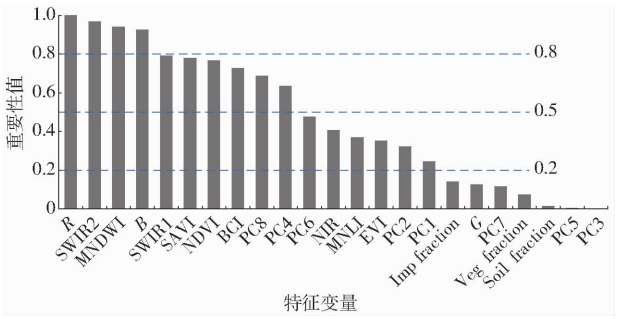


图 6 模型 3 特征变量重要性

Fig. 6 Feature variables importance value of model 3

55.58%,而植被在红光波段具有强吸收的特性,所以 *R* 变量在研究区域内对于大范围的植被信息提取贡献度最高;SWIR2 重要性高是因为土壤在短波红外区有强反射的特性,而研究区内耕地土壤面积占比 24.01%,仅次于植被林地,所以在分类过程中 SWIR2 对土壤信息的提取起到了重要作用;MNDWI 重要性较高的原因可能是因为研究区域内包含了赤水河的一部分和若干水域,虽然面积占比不高,但分布较为集中易于识别,所以使得 MNDWI 重要性较高;而 Landsat-8 OLI 的蓝光波段在水体制图、土壤和植被之间的区分时作用较大,所以 *B* 变量的重要性排序也较为靠前。重要性值在 0.5~0.8 之间的特征变量为 SWIR1、SAVI、NDVI、BCI、PC8、PC4,其中 3 个光谱指数分别适用于提取土壤(SAVI)、植被(NDVI)和不透水面(BCI),由此看出这 3 个光谱指数可有效增强影像类间的可分性。组分特征变量的重要性值在 0.2 以下,原因是组分变量是基于原始影像数据线性变换生成,与原始数据的相关性相对较高,且信息量的丰富程度上也不及原始数据,所以在排序中相对靠后,而各变量中纹理特征变量中的 PC3、PC5 重要性最低。以上结果说明虽然组分、纹理特征变量的融合可以提高随机森林分类方法的精度,但在变量的重要程度上,还是原始光谱特征变量与其相应光谱指数变量对分类的贡献度更高,这也与一些学者的研究结论一致。

4.4 不同方法分类结果比较

为评估本文方法在众多分类方法中的表现,将模型 3(记为 RF_MESMA)、模型 2(记为 RF)的分类结果分别与基于 LSMA 和 RF 方法的分类结果、目前应用范围较广的 SVM 和 MLC 方法分类结果进行了比较。其中,LSMA 与 MESMA 采用相同的端元进行混合像元分解,生成植被、土壤、不透水面 3 种组分数据,与 RF 方法相结合进行影像分类(记为 RF_LSMA);SVM 和 MLC 分成 2 种情况,即组分特征变量参与分类和未参与分类,以相同的训练和验证样点,构建了 4 种模型进行实验,模型分别为:组分特征变量参与分类的 SVM 模型(记为 SVM_MESMA),组分特征变量未参与分类的 SVM 模型(记为 SVM);组分特征变量参与分类的 MLC 模型(记为 MLC_MESMA),组分特征变量未参与分类的 MLC 模型(记为 MLC)。通过比较各方法的特征变量数、分类总体精度和 Kappa 系数,分析各组合、分类方法对山丘区土地利用信息分类提取的适用性,比较结果如表 5 所示。

表 5 不同方法精度评价

Tab. 5 Accuracy assessment of classification result based on different methods

分类方法模型	RF	SVM	MLC	RF_MESMA	RF_LSMA	SVM_MESMA	MLC_MESMA
特征变量数	20	20	20	23	23	23	23
总体精度/%	89.48	88.79	85.72	90.50	89.96	88.85	86.35
Kappa 系数	0.839 0	0.820 1	0.785 1	0.854 4	0.844 8	0.828 7	0.794 7

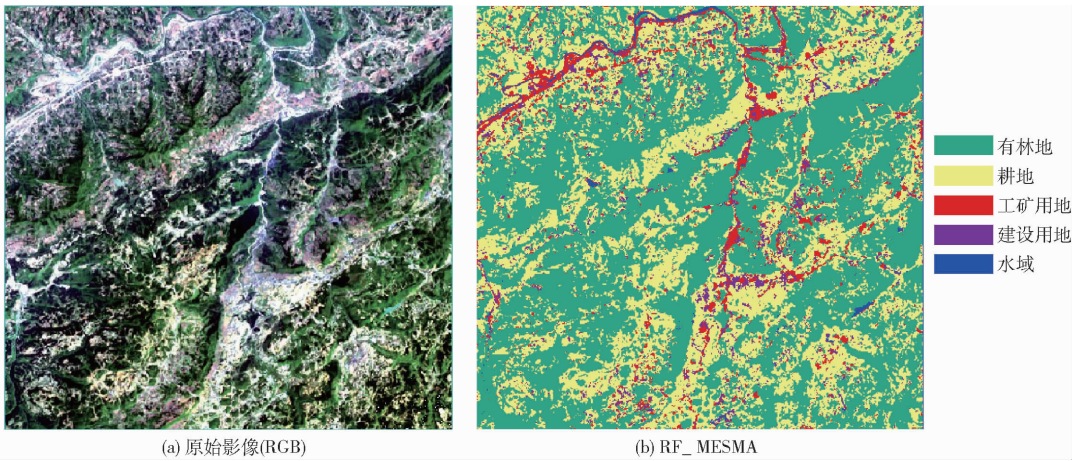
对照结果可知,RF、SVM、MLC 3 种方法相比较,RF 结果精度最优达 89.48%。各分类方法在加入 MESMA 生成的组分特征变量后,分类精度都有不同程度的提高,分别提高 1.02 个百分点、0.06 个百分点、0.63 个百分点,其中 RF_MESMA 方法模型的分类结果精度最高为 90.50%,与 SVM_MESMA、MLC_MESMA 相比精度分别高 1.65 个百分点和 4.15 个百分点,说明 MESMA 方法生成的组分信息数据作为特征变量参与分类后可在一定程度上提高影像的分类精度。

RF_MESMA 与 RF_LSMA 方法相比较,前者精度为 90.50%,后者精度为 89.96%,说明 MESMA 方法对于提高信息分类提取精度作用更为明显。

模型 3(RF_MESMA)分类结果如图 7b 所示,与原始影像图 7a 相比可看出,影像中主要地类信息都得到了有效的提取。

5 结束语

MESMA 可以获得较为精确的组分丰度信息;RF 分类结果优于相同特征变量下的 SVM 和 MLC 分类结果;在 MESMA 生成的组分信息变量参与分类后,3 种方法的分类精度均有所改善,分别达到 90.50%、88.85%、86.35%,其中 RF 的分类精度改善最为显著;MESMA 与 LSMA 生成的组分信息变量相比,前者对于改善分类精度效果更为明显。MESMA 对于提高影像分类精度起到一定积极作



(a) 原始影像(RGB)　　(b) RF_ MESMA

图 7　影像分类图

Fig. 7　Classification results of image

用,基于 MESMA 和 RF 的方法对中等空间分辨率影像山丘区土地利用信息分类提取的精度较高,利用该方法开展遥感影像解译可为大尺度的土地利用、复垦监测和管理工作提供技术支持和理论参考。

参　考　文　献

1　林楠,姜琦刚,杨佳佳,等. 基于资源一号 02C 高分辨率数据的农业区土地利用分类[J/OL]. 农业机械学报,2015,46(1):278 – 284. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag = 1&file_no = 20150139&journal_id = jcsam. DOI:10.6041/j.issn.1000-1298.2015.01.039.
LIN Nan, JIANG Qigang, YANG Jiajia, et al. Classifications of agricultural land use based on high-spatial ZY1 – 02C remote sensing images [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(1): 278 – 284. (in Chinese)

2　张增祥,汪潇,温庆可,等. 土地资源遥感应用研究进展[J]. 遥感学报,2016,20(5):1243 – 1258.
ZHANG Zengxiang, WANG Xiao, WEN Qingke, et al. Research progress of remote sensing application in land resources[J]. Journal of Remote Sensing, 2016, 20(5): 1243 – 1258. (in Chinese)

3　FENG Quanlong, GONG Jianhua, LIU Jiantao, et al. Flood mapping based on multiple endmember spectral mixture analysis and random forest classifier—the case of Yuyao, China[J]. Remote Sensing, 2015, 7(9): 12539 – 12562.

4　马孟莉,朱艳,李文龙,等. 基于分层多端元混合像元分解的水稻面积信息提取[J]. 农业工程学报,2012,28(2):154 – 159.
MA Mengli, ZHU Yan, LI Wenlong, et al. Extracting area information of paddy rice based on stratified multiple endmember spectral mixture analysis[J]. Transactions of the CSAE, 2012, 28(2): 154 – 159. (in Chinese)

5　王浩,吴炳方,李晓松,等. 流域尺度的不透水面遥感提取[J]. 遥感学报,2011,15(2):388 – 400.
WANG Hao, WU Bingfang, LI Xiaosong, et al. Extraction of impervious surface in Hai Basin using remote sensing[J]. Journal of Remote Sensing, 2011, 15(2): 388 – 400. (in Chinese)

6　缪春华,张显峰,刘羽. 基于多端元光谱分解的干旱区植被覆盖度遥感反演[J]. 应用生态学报,2012,23(12):3243 – 3249.
MIU Chunhua, ZHANG Xianfeng, LIU Yu. Remote sensing retrieval of vegetation coverage in arid areas based on multiple endmember spectral unmixing[J]. Chinese Journal of Applied Ecology, 2012, 23(12): 3243 – 3249. (in Chinese)

7　马玥,姜琦刚,孟治国,等. 基于随机森林算法的农耕区土地利用分类研究[J/OL]. 农业机械学报,2016,47(1):297 – 303. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag = 1&file_no = 20160140&journal_id = jcsam. DOI:10.6041/j.issn.1000-1298.2016.01.040.
MA Yue, JIANG Qigang, MENG Zhiguo, et al. Classification of land use in farming area based on random forest algorithm[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(1): 297 – 303. (in Chinese)

8　郭玉宝,池天河,彭玲,等. 利用随机森林的高分一号遥感数据进行城市用地分类[J]. 测绘通报,2016(5):73 – 76.

9　杨珺雯,张锦水,朱秀芳,等. 随机森林在高光谱遥感数据中降维与分类的应用[J]. 北京师范大学学报:自然科学版,2015,51(1):82 – 88.
YANG Junwen, ZHANG Jinshui, ZHU Xiufang, et al. Random forest applied for dimension reduction and classification in hyperspectral data[J]. Journal of Beijing Normal University: Natural Science, 2015, 51(1): 82 – 88. (in Chinese)

10　王书玉,张羽威,于振华. 基于随机森林的洪河湿地遥感影像分类研究[J]. 测绘与空间地理信息,2014,37(4):83 – 85.
WANG Shuyu, ZHANG Yuwei, YU Zhenhua. Classification of Honghe Wetland remote sensing image based on random forests[J]. Geomatics & Spatial Information Technology, 2014, 37(4): 83 – 85. (in Chinese)

11　张玉芳,庞艳梅,刘琰琰,等. 近 50 年四川省水稻生产潜力变化特征分析[J]. 中国生态农业学报,2014,22(7):813 – 820.
ZHANG Yufang, PANG Yanmei, LIU Yanyan, et al. Potential productivity of rice in Sichuan Province in recent five decades[J]. Chinese Journal of Eco-Agriculture, 2014, 22(7): 813 – 820. (in Chinese)

12 孙园园,徐富贤,孙永健,等. 四川稻作区优质稻生产气候生态条件适宜性评价及空间分布[J]. 中国生态农业学报,2015,23(4):506 - 513.
SUN Yuanyuan, XU Fuxian, SUN Yongjian, et al. Suitability evaluation of eco-climatic conditions for high quality rice production in Sichuan Province[J]. Chinese Journal of Eco-Agriculture, 2015, 23(4): 506 - 513. (in Chinese)
13 肖科. 泸州市土地整理效益分析与评价[D]. 雅安:四川农业大学,2009.
14 刘正春,曾永年,何丽丽,等. 基于光谱归一化的变组分光谱混合分析(NMESAM)方法及其应用[J]. 遥感技术与应用,2012,27(2):159 - 167.
LIU Zhengchun, ZENG Yongnian, HE Lili, et al. Method of normalized multiple endmember spectral mixture analysis and its application[J]. Remote Sensing Technology and Application, 2012, 27(2): 159 - 167. (in Chinese)
15 MAXWELL A E, WARNER T A. Differentiating mine-reclaimed grasslands from spectrally similar land cover using terrain variables and object-based machine learning classification[J]. International Journal of Remote Sensing, 2015, 36(17): 4384 - 4410.
16 WU Chaofan, DENG Jinsong, WANG Ke, et al. Object-based classification approach for greenhouse mapping using Landsat-8 imagery[J]. International Journal of Agricultural & Biological Engineering, 2016, 9(1): 79 - 88.
17 丁建丽,姚远. 干旱区绿洲典型地物 MESMA 模拟分解与验证[J]. 地球信息科学学报,2013,15(3):452 - 460.
DING Jianli, YAO Yuan. Research on pixel un-mixing of typical surface features in oasis based on the MESMA model[J]. Journal of Geo-information Science, 2013, 15(3): 452 - 460. (in Chinese)
18 POWELL R L, ROBERTS D A, DENNISON P E, et al. Sub-pixel mapping of urban land cover using multiple endmember spectral mixture analysis: Manaus, Brazil[J]. Remote Sensing of Environment, 2007, 106(2): 253 - 267.
19 DENNISON P E, ROBERTS D A. Endmember selection for multiple endmember spectral mixture analysis using endmember average RMSE[J]. Remote Sensing of Environment, 2003, 87(2): 123 - 135.
20 DENNISON P E, HALLIGAN K Q, ROBERTS D A. A comparison of error metrics and constraints for multiple endmember spectral mixture analysis and spectral angle mapper[J]. Remote Sensing of Environment, 2004, 93(3): 359 - 367.
21 LIU Jiantao, FENG Quanlong, GONG Jianhua, et al. Land-cover classification of the Yellow River Delta wetland based on multiple endmember spectral mixture analysis and a random forest classifier[J]. International Journal of Remote Sensing, 2016, 37(8): 1845 - 1867.
22 徐涵秋,王美雅. 地表不透水面信息遥感的主要方法分析[J]. 遥感学报,2016,20(5):1270 - 1289.
XU Hanqiu, WANG Meiya. Remote sensing-based retrieval of ground impervious surfaces[J]. Journal of Remote Sensing, 2016, 20(5): 1270 - 1289. (in Chinese)
23 DENG Chengbin, WU Changshan. BCI: a biophysical composition index for remote sensing of urban environments[J]. Remote Sensing of Environment, 2012, 127(5): 247 - 259.
24 王文静,张霞,赵银娣,等. 综合多特征的 Landsat 8 时序遥感图像棉花分类方法[J]. 遥感学报,2017,21(1):115 - 124.
WANG Wenjing, ZHANG Xia, ZHAO Yindi, et al. Cotton extraction method of integrated multi-features based on multi-temporal Landsat 8 images[J]. Journal of Remote Sensing, 2017, 21(1): 115 - 124. (in Chinese)
25 BREIMAN L. Random forest[J]. Machine Learning, 2001, 45(1): 5 - 32.
26 赵玉,王红,张珍珍. 基于遥感光谱和空间变量随机森林的黄河三角洲刺槐林健康等级分类[J]. 遥感技术与应用,2016,31(2):359 - 367.
ZHAO Yu, WANG Hong, ZHANG Zhenzhen. Forest healthy classification of *Robinia pseudoacacia* in the Yellow River Delta, China based on spectral and spatial remote sensing variables using random forest[J]. Remote Sensing Technology and Application, 2016, 31(2): 359 - 367. (in Chinese)
27 EISAVI V, HOMAYOUNI S, YAZDI A M, et al. Land cover mapping based on random forest classification of multitemporal spectral and thermal images[J]. Environmental Monitoring and Assessment, 2015, 187(5): 291.
28 GHOSH A, SHARMA R, JOSHI P K. Random forest classification of urban landscape using Landsat archive and ancillary data: combining seasonal maps with decision level fusion[J]. Applied Geography, 2014, 48(2): 31 - 41.
29 雷震. 随机森林及其在遥感影像处理中应用研究[D]. 上海:上海交通大学,2012.
30 王兴玲,李占斌. 基于网格搜索的支持向量机核函数参数的确定[J]. 中国海洋大学学报,2005,35(5):859 - 862.
WANG Xingling, LI Zhanbin. Identifying the parameters of the kernel function in support vector machines based on the grid-search method[J]. Periodical of Ocean University of China, 2005, 35(5): 859 - 862. (in Chinese)
31 刘颖. 基于机器学习的遥感影像分类方法研究[M]. 北京:清华大学出版社,2014.