

一种遥感混合分类方法及其在山区土地覆盖分类中的应用

范 莉¹, 刘洪斌^{1,2}, 武 伟¹

(1. 西南农业大学资源环境学院, 重庆 北碚 400716; 2. 中科院地球化学研究所, 贵阳 550053)

摘 要: 混合分类既可以改进纯粹的监督分类或非监督分类过程的效率, 又可以提高纯粹的监督分类或非监督分类过程的精度。介绍一种新的遥感混合分类方法——控制聚类法(guided clustering)。首先对该方法的原理进行阐述, 然后利用该方法对重庆市具有代表性的山区酉阳县的遥感影像进行分类实验。实验结果表明, 这种方法能有效的提高分类精度, 特别适合我国南方山区土地覆盖信息的提取。

关键词: 遥感; 控制聚类法; 山区

中图分类号: F301.24; TP79

文献标识码: A

文章编号: 1005-3409(2006)03-0235-03

Application of Hybrid Classification on the Land Cover Classification in Mountain Areas

FAN Li¹, LIU Hong-bin^{1,2}, WU Wei¹

(1. Southwest Agriculture University, Beibei, Chongqing 400716, China;

2. Research Institute of Geochemistry of Chinese Academy of Sciences, Guiyang 550053, China)

Abstract: The hybrid classification can improve the efficiency of process on supervised classification or unsupervised classification. First the theory of guided clustering was described. Then this method was developed to make a classification experiment using TM imagine in the representative mountain area—Youyang. The results showed that the guided clustering can increase the classification accuracy, especially suitable to apply in this south mountain area.

Key words: remote sensing; guided clustering; mountain areas

土地覆盖和土地利用与全球环境变化和人类的生产生活密切相关, 近年来成为地球系统科学研究领域中的一个重要分支^[1-3]。山区由于地形复杂, 且常年多云雾, 在此种数据条件下很难采用常规方法进行覆盖分类^[4]。遥感技术具有快速及时地提供地表信息和方便的图像处理能力, 遥感数据覆盖范围广, 重复周期高, 能对地物进行多光谱、多时相和多角度的观测, 因此遥感成为山区获取地面信息的主要手段。

传统上, 利用遥感技术进行土地覆盖分类多采用非监督分类和监督分类的方法。非监督分类方法虽然能够自动地将影像划分为若干光谱特征相对一致的若干类, 但需要人工判别分出类型所对应的实际类别, 该方法带有一定的“盲目”性, 实际上分出的类别是“谱类”而不是“地类”^[5]。监督分类比非监督分类更多地要用户来控制, 常用于对研究区比较了解的情况, 在对影像分类前需要借助其它信息来定义训练样本, 然后基于该样本使计算机系统自动识别具有相同特征的像元, 由此得到分类结果。不管是非监督分类还是监督分类, 主观性都很强, 即使对同一幅影像用同一种分类方法, 不同的人也会得到不同的分类结果。传统的分类方法结果不尽如人意, 既为了简化并改进纯粹的监督分类和非监督分类的过程效率, 又为了提高纯粹的监督分类和非监督分类过程的精度, 各种形式的混合分类已经被提出^[6]。控制聚类法(guided clustering)是近几年提出的一种混合分类算法。这种分类方法将非监督分类和监督分类方法结合起来对遥感影像进行分类, 减少人为主观性的干扰, 在遥感分类问题上表现出很大的优势。而且“控制聚类法”(Bauer et al., 1994)

特别适于山区的土地覆盖划分^[7](例如: 森林、农田、草地、水域、未利用地和灌木林)。

1 研究区域概况

酉阳县位于重庆直辖市东南边缘(108°18′25″~109°19′02″E, 28°19′28″~29°24′18″N), 境内群山环抱, 沟谷纵横, 地势起伏大, 地形变化复杂, 位于渝、黔、湘、鄂、黔沿线褶皱带中, 属巫山大娄山中山区, 地势中部高, 海拔 600~1 800 m, 多为船形山、柱状山; 东西两侧低, 海拔 400~600 m, 多为中山、低丘、槽槽、平谷、洼地。故形成以毛坝盖为分水岭的沅江、乌江两大水系。酉阳县境南北长 119.85 km, 东西宽 98.9 km。幅员面积 5 173 km², 总面积居重庆直辖市各县(区)第一位。气候温润, 属亚热带季风气候。年平均气温 8.1~17.1℃, 全年日照时数为 1 131 h, 酉阳境内全年降雨量 1 150~1 660 mm。

2 研究区数据源

以 TM 遥感影像为主要数据源, 同时, 收集研究区自然环境和社会经济方面的专题图件和数据资源。本文采用 1996 年 5 月 29 日 TM 图像, 酉阳县行政区划图, 酉阳县 1:25 万地形图, 重庆市 1:30 万土地利用现状图。

3 聚类控制法分类

3.1 控制聚类法原理

应用控制聚类法首先要求分析者为图像中每种要分类

的覆盖类型勾画了许多“类似监督的”训练样区,以帮助分析者识别需要定义的光谱类型。所选择的非监督训练样区是图像的子集,它在选择的目的在于监督分类训练样区完全不同^[6]。为了产生好几种光谱信号,在非监督分类集群程序中使用已给定信息类型的所有训练集的数据。这些信号被分析者检验,其中一些不纯的像元可能被丢弃或合并,剩下的就是描述所得到的信息类型的光谱子集。一旦获得这种光谱子类型足够大的数字,就可以对采样光谱子类型的全部子集采用最大似然法分类。具体分类过程如下:

3.2 针对信息类型勾画训练样区

分类前的遥感数据预处理及增强处理这里不再赘述。但需要提出的是,针对山区地势起伏大、地形复杂、常年多云雾、分类难度大的特点,着重加强对影像的地形校正和去云处理,以此来提高图像分类精度。

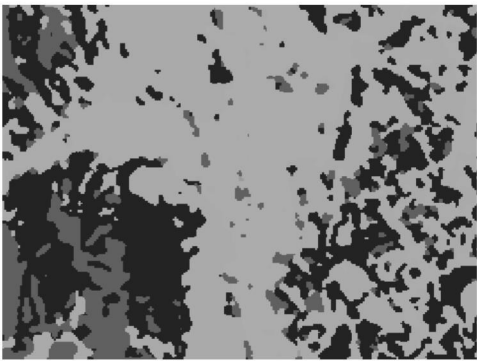
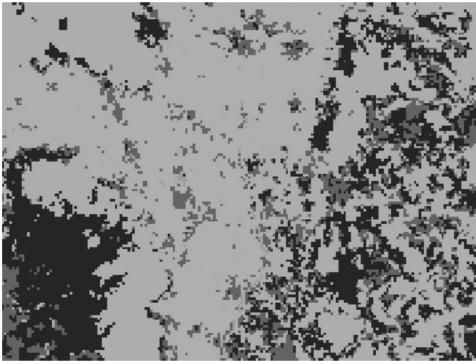


图 1 分类后平滑图示

3.3 对原始图像进行非监督分类

确定训练样区之后,在控制聚类法中,首先利用它的非监督分类法(ISODATA)对原始影像进行非监督分类,在传统的非监督分类中,计算机是按照一定规则,自动根据像元光谱或空间等特征组成集群组的,最终必须由分析者将每个集群组划分到某一类别中去,而控制聚类法中的非监督分类是根据训练本来决定每个光谱特征所组成的集群组是属于哪个信息类型的。非监督训练分类区在图像上的各种位置都包含了多种覆盖类型,这样就确保了图像中所有的光谱类型在各种子区域的某处被表示。根据训练样区将山区的土地覆盖类型分为耕地、林地、园地、草地、桑园、居民及建设用地、水域、火烧迹地、未利用地 9 类。

3.4 检验各信息类型的光谱

检验每个类型的光谱,及时合并或删除光谱类型。根据下列因素对集群的最后数目进行合并和删除: (1) 在原始图像上显示给定类型; (2) 对每种集群进行多维直方图分析; (3) 多元距离测量。最后将所有的类型的信息组合成一个单独的训练模板。

检验每个信息类型光谱后,一些信号被丢弃或合并,剩下的土地覆盖类型为耕地、林地、园地、草地、居民及建设用地、水域、未利用地。

3.5 监督分类

采用检验后的光谱文件训练监督分类器,对原始影像进行监督分类,这里采用最大似然法对其进行分类,从而得到最终的分类结果影像。

3.6 分类后处理

分类后滤波采用 ERDAS“聚类、过滤、填充”的运算法则来进行^[7]。同一土地覆盖类型中,滤除少于 4 个相邻像元的地类,取而代之的是相邻地类中没有产生方向偏离的像元。对山区进行滤波处理之后,面积小于 4 个像元,但被其他地类包围着的地类不删除。城镇的像元划分不进行滤波处理,开放的水体像元也不进行该处理,因为 T M 影像上水体的划

对 T M 影像中除第 6 波段外的其他 6 个波段进行分析,为获取进行土地覆盖分类的训练样本,主要参照重庆市 1996 年土地利用现状图,在影像上用 AOI 矢量工具勾画训练样区,所选择的非训练样区是图像的子集,它在选择目的上与监督训练样区完全不同^[6]。控制聚类法获取的每个非监督训练区所包括的每个像元的光谱类型不一定一致,但其信息类型必须一致(如水域,包含水体深浅不一的光谱类型);监督训练区位于含有均一覆盖类型的区域中,但是非监督训练区在影像的各个位置上包含了多种覆盖类型。这不但省去了涉及监督分类中选取训练样区时必须做到每个地类类型的训练样区中每个像元的光谱特征与其他地类类型的训练样区像元的光谱特征不一样,而且确保了影像中所有光谱类型在各种子区域的某处被表示,还在很大程度地减少了选择训练样区的主观性。

分通常接近 100% 的准确性^[8]。

例如在一片完整的旱地里有几个散布的像元点被分成了林地,反之亦然。滤波处理后的结果见图 2(b)

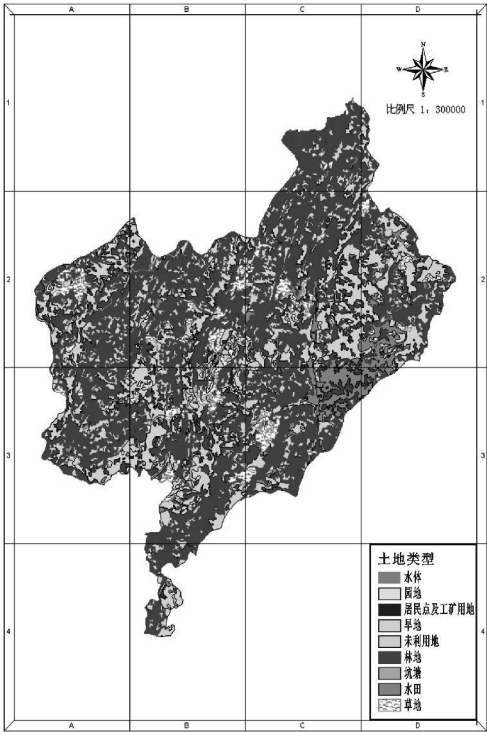


图 2 1996 年酉阳土地覆盖图

4 分类结果的验证及输出

酉阳县土地覆盖类型划分为耕地、林地、园地、草地、居民及建设用地、水域、未利用地 7 个地类 181 个训练样本,这

里以这些训练样本来测试分类结果的精度,其结果见表 1。

表 1 训练样本误差矩阵及精度测试结果 像元

地类	耕地	林地	园地	水域	居民点	草地	未利用地	行汇总
耕地	283	0	22	8	0	36	14	353
林地	10	290	51	0	0	42	6	360
园地	8	34	1495	0	0	30	5	1531
水域	0	0	0	255	0	0	0	255
居民点	0	0	0	0	109	0	37	130
草地	11	12	20	1	0	275	0	309
未利用地	0	0	0	0	5	0	214	214
列汇总	312	336	1578	264	114	383	276	3263
精度/%	90.71	86.31	94.74	96.59	95.61	71.80	77.54	

据表 1,分类总精度达 89.52%,分类精度较高的是水域、园地、居民点,达到 94% 以上,次之是耕地、林地,达到 85%~91%,分类精度较差的是未利用地、草地,精度为 70%~80%,这是因为所用卫星资料的获取时间是 5 月份,此时农作物主要是油菜,地物光谱和园地的光谱相似,旱地参考文献:

[1] 杨立民,朱智良.全球及区域尺度土地覆盖土地利用遥感研究的现状和展望[J].自然资源学报,1999,14(4):340—344.
[2] 赵庚星,李玉环,徐春达.遥感和 GIS 支持下的土地利用动态监测研究[J].应用生态学报,2000,11(4):573—576.
[3] 摆万奇,赵士洞.土地覆盖变化研究模型综述[J].自然资源学报,1997,12(2):169—175.
[4] 张洪亮,促绍祥,张军.国外遥感图像的地形归一化方法研究进展[J].遥感信息,2001,
[5] 戴昌达,姜小光,唐伶俐,等.遥感图像应用处理与分析[M].北京:清华大学出版社,2003.
[6] Thomas M Lillesand, Ralph W. Kiefer Remote Sensing and Image Interpretation[M]. Beijing: Publishing House of Electronics Industry, 2003.
[7] Bauer, Me, et al. Satellite inventory of Minnesota forest resources[J]. Photogrammetric Engineering and Remote Sensing, 1994, 60(3): 287—298.
[8] Heather M Reese, Thomas M Lillesand, David E Nagel et al. Statewide land cover derived from multiseasonal Landsat TM data— A retrospective of the WISCLAND project[J]. Remote Sensing of Environment, 2002, (82): 224—237.

(上接第 234 页)

的体系,其溶蚀过程不同于封闭体系中的溶蚀过程,溶蚀作用开始以后,在水动力条件不断变化的动态水介质中,溶解的 CaCO₃ 和残留的不溶堆积物不断被带走,因此溶蚀作用的推进更快。尤其是有侵蚀性 CO₂ 存在时,这一作用更为显著,岩石溶蚀作用产生初期所表现出的启动速度上的差异对溶蚀作用的进行有较大的影响,在动态水介质中,启动溶蚀速率快的其溶蚀作用推进也较快,从而使得不同岩性的岩石之间呈现出明显的差异溶蚀。灰岩类的岩石一般比白云岩类的岩石具有较高的启动溶蚀速率和较慢的溶蚀衰减速率,因此溶蚀作用的推进快,岩溶也相对发育。对于岩性较为接近的岩石而言,溶蚀也总是在启动速度较快的一方最先发育。但是水岩反应过程中,水与岩石的有效接触面积极大地影响着溶蚀的发育速度,对于致密块状或厚层状灰岩来讲,尽管石灰岩本身有较高的启动溶蚀速率和较慢的衰减速率,但其裂隙不发育,水岩反应过程中的有效接触面积就大大减小,因此其总体的溶蚀速率也相应降低;对于薄层状白云岩、泥质白云岩来说,薄层状的岩体结构造成了裂隙网络的发育,从而增大了水岩反应的有效接触面,溶蚀速率就大大增加,因此其溶蚀特征反而较石灰岩要明显。野外所观察到的白云岩类岩石的溶蚀特征较石灰岩类岩石更明显的参考文献:

[1] 彭汉兴,宋汉周.清江石门地区第四纪冰川沉积与岩溶发育期的探讨[J].中国岩溶,1992,11(2):173—179.
[2] 谢宇平.第四纪地质学及地貌学[M].北京:地质出版社,1994.65—71.
[3] 张倬元,王士天,王兰生.工程地质分析原理[M].北京:地质出版社,1994.470—488.
[4] 刘世凯,陆永清,欧湘萍.公路工程地质与勘察[M].北京:人民交通出版社,1999.112—116.
[5] 刘再华,WDreybrod,等.CaCO₃—CO₂—H₂O 岩溶系统的平衡化学及其分析[J].中国岩溶,2005,24(1):1—14.

和未利用地的光谱相似,导致分类精度较差。草地、耕地、未利用地(以裸露山石为主),由于光谱特征有一定程度的类似,加之山区地块零碎,这三种地类之间存在混分的现象。分类精度合格,可以进行最后的结果输出。

5 结 语

- (1) 针对山区地势起伏大、地形复杂的特点,采用控制聚类法进行土地覆盖分类,经研究证明,是一种切实有效的方法。不仅增加了用传统监督分类法分类的精度,还增加了用非监督分类法分类的精度,而且它也提高了整个分类过程的效率。
(2) 控制聚类法有助于分析者通过集群表示信息“自动”分类来识别各种光谱子类,这就增加了分类的客观性。
(3) 控制聚类法由于每次都标注一个信息类型,而且由于单一训练区包含复杂的覆盖信息类型条件,所以假的集群就能够被容易的识别(如显示另外的密闭有林地含有下层矮生植物,或者显示农作物耕地中裸地的情况)。同时,控制聚类法有助于识别那些由于被忽略而包含在训练区边缘附近的混合像元的情况。