

# 基于知识库分类法在植被提取中应用研究

王旭红,肖 平

(陕西省基础地理信息中心,陕西 西安 710054)

## The Research on the Way of Knowledge-based Classification in Vegetation Extraction

WANG Xu-hong, XIAO Ping

**摘要:**目前,计算机运算速度和容量的不断增长、数据库技术及信息网络技术的发展和人工智能的引用为分析和处理大数据量的遥感和地理空间数据创造了更佳的条件。通过对地学知识表示、规则库、推理机和专家分类器的研究,提出一套基于知识的适用于太白红杉数据生产的实用方法。

**关键词:**专家分类器;规则库;推理机

### 一、引 言

随着空间技术和计算技术的发展,遥感信息的空间分辨率和光谱分辨率不断提高,扩展了其应用领域;计算机运算速度和容量的不断增长、数据库技术及信息网络技术的发展和人工智能的引用为分析和处理大数据量的遥感和地理空间数据创造了更佳的条件。这些都为遥感技术的实用化奠定了基础,因而如何应用计算机技术建立地学知识分析处理模型和遥感信息的分析处理方法,并在此基础上,如何通过扩展传统地学分析方法和数理统计分析方法,引入人工智能领域中新发展的高新技术、新方法,开拓智能化的遥感地学分析模型,改善对遥感信息分析、处理、识别的精度,提高智能化、自动化的水平显得尤为重要和迫切。本文研究基于知识库分类法,建立太白红杉的数据提取知识库。太白红杉是太白落叶松的一种,是秦岭的特有树种,分布区很小,在太白山下限海拔 2 800 m,上限海拔达 3 400 m,位于高山灌丛草甸之上,冷杉林带之下,仅在太白山有较大面积的纯林。综合运用计算机技术、人工智能、地学知识、地理信息和遥感技术是克服遥感信息进行空间决策分析中遇到的不确定性和多解性问题的一个重要途径,具有重大的现实意义。

### 二、专家分类系统

专家分类系统是一类计算机应用,它主要基于一系列某一领域专家所定义的规则、条件或假设解决某一问题或做出决策。这些规则集就是所谓的知

识库,也叫作决策树。它通过图解地组装成一个树形结构,其中,每一规则表达了一个问题,其答案被下送到由含有条件查询的新的分支上。专家分类系统是模拟人类组合各种带有因果关系的知识进行推理并得出结论。

### 三、几种分析方法的引入

#### 1. 引入高程数据

Franklin 和 Peddle (1989, 1990, 1992) 使用 SPOT HRV 遥感数据和 DEM 产生的地形类别数据,提出多种数据源的证据推理分类方法,结果优于传统的统计分类方法。

#### 2. 不确定推理论

根据 D-S 理论,知识库中证据或事实不仅可被作用于单一类别,而且可同时被作用于一组类别,比如对于某一影像中地物对象,如果没有足够的证据来表明其属于某一类别,则可以表示为更高的超类别进行描述。

#### 3. 对监督训练数据集频率分析

虽然 D-S 证据推理理论对多源数据的分析、集成、分类提供了适宜的框架,但是由于当前基于机械式学习所获得的知识往往主观性比较强而导致结果有异,D. R. Peddle(1995)提出了通过对监督训练数据集频率分布的分析、计算来产生证据的方法。

### 四、关键技术

1. 通过研究地物在航空影像上光谱信息、纹理特征以及地物分布的地带性特征,与地面坡度、起伏

度、地理位置和海拔高度的关系等地理知识,运用计算机技术、人工智能、遥感图像处理 and 地理信息技术,建立用于太白红杉地物提取的知识规则库,提出一套基于知识的适用于太白红杉地物提取数据生产的实用系统和方法;

2. 用于辅助判断的地理知识获取与表达;
3. 专家知识库的建立;
4. 基于统计分析、知识模型的半智能化逻辑判断规则建立;
5. 用于逻辑判断规则的专家分类器的建立。

### 五、技术流程框图

太白红杉分布图的制作流程如图 1 所示。

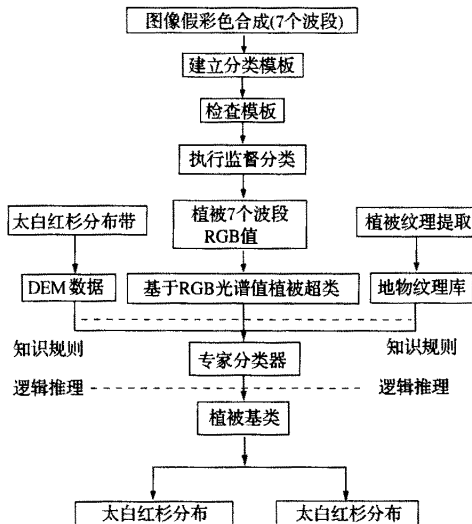


图 1 太白红杉分布图制作的流程

### 六、知识库的建立

#### 1. 知识的表示

地理知识的表示,要结合具体地理问题的实际特点。规则是地理知识的主要表示方法,能反映地理客观事物的基本规律性。但由于在实际问题中,所掌握的规律大多是经验性、常识性的知识,其含义是不精确的,所反映的规则也是不确定的。因此,采用放置假设信息,即设置规则的不确定可用可信度的方法来设置可信因子(Certainty Factor, CF)。我们采用产生式的模糊规则形式来表达知识,其基本形式为

if(条件), then(结论), CF(可信度因子)

其中,CF的取值为(0, 1),反映结论的可能性程度,当 CF = 0 时,则完全排除结论的可能性;当 CF = 1,表示结论完全可靠。

#### 2. 知识的获取

1. 植被的 RGB 值,监督分类建立植被光谱分波段的 RGB 值;
2. 纹理特征,制作纹理分类影像;
3. 高程数据,由数字高程模型数据取得。

#### 3. 规则库

这些规则集就是所谓的知识库,也叫作决策树。它通过图解地组装成一个树形结构,其中,每一规则表达了一个问题,其答案被下送到由含有条件查询的新的分支上,如图 2 所示。

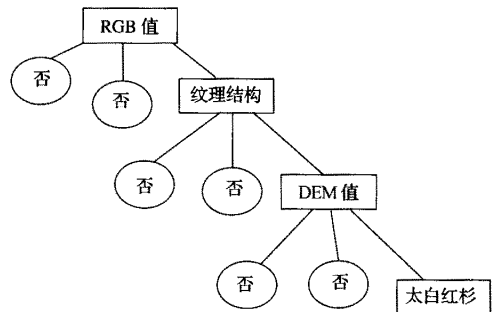


图 2

判别树归纳算法为:

算法 Generate\_decision\_tree 由给定的训练数据产生一棵判别树。

输入 训练样本 samples,由离散值属性表示;候选属性的集合 attribute\_list。

输出 一棵判别树。

方法为:

1. 创建节点 N
2. if samples 都在同一个类 C Then
3. 返回 N 作为叶节点,以类 C 标记
4. if attribute\_list 为空 Then
5. 返回 N 作为叶节点,标记为 samples 中最普通的类
6. 选择 attribute\_list 中具有最高信息增益的属性 test\_attribute
7. 标记节点 N 为 test\_attribute
8. for each test\_attribute 中的已知值  $a_i$ //划分 samples
9. 由节点 N 长出一个条件为 test\_attribute =  $a_i$  的分枝
10. 设  $s_i$  是 samples 中 test\_attributes =  $a_i$  的样本的集合;//一个划分
11. if  $s_i$  为空 Then
12. 加上一个树叶,标记为 samples 中最普通的

类

13. else 加上一个由 Generate\_decision\_tree (s<sub>i</sub>, attribute\_list\_test\_attribute)返回的节点

### 七、推理机

推理机是一组用来控制、协调整个专家系统的系统。它根据数据库当前输入的数据,利用知识库中的知识,按一定的推理策略,去求解当前的问题、解释外部输入的事实和数据,推导出结论并向用户提示等。本次实验推理策略采用正向推理,从已知事实出发,正向使用推理规则,其过程的算法描述如下。

1. 把用户提供的初始证据放入综合数据库。
2. 检查综合数据库中是否包含了问题的解,若已包含,则求解结束,并成功退出;否则执行下一步。
3. 检查知识库中是否有可用知识,若有,形成当前可用知识集,执行下一步;否则转 5。
4. 按照某种冲突解的策略,从当前可用知识集中选出一条知识进行推理,并将推出新的事实加入综合数据库中,然后转 2。
5. 询问用户可以进一步补充新的事实,若可补充,则将新的事实加入综合数据库中,然后转 3;否则表示无解,失败退出。

### 八、专家分类器

提供了一种基于规则的方法,为拥有第一手数据和知识的专家提供一个用户界面,让专家应用于确定变量、规则和感兴趣的输出类型(假设),生成层次决策树,建立预测类标记未知的对象类的,描述并区分数据类或概念的模式,最终建立知识库,如图 3 所示。

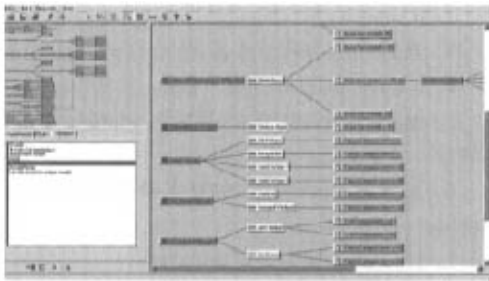


图 3

### 九、试验结果

根据上述方法建立分类模型,制作出图 4。

太白红杉分布图

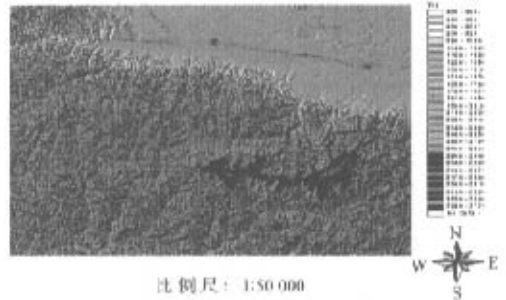


图 4

### 参考文献:

- [1] 朱志诚. 太白山顶植被的起源和发展[J]. 西北大学学报, 1998, 19(1): 104.
  - [2] 朱志诚. 秦岭落叶松林的主要类型及其演变的初步分析[J]. 西北大学学报. 1998, 19(3): 156.
  - [3] JORDAN L E. 更易于从新的地理数据中提取信息的专家分类系统[J]. 遥感信息, 2001, (1).
  - [4] 周成虎, 等. 遥感影像地学理解与分析[M]. 北京: 科学出版社, 1999.
  - [5] 史培军, 等. 土地利用/覆盖变化研究的方法与实践[M]. 北京: 科学出版社, 2000.
  - [6] 王旭红. 基于卫星影像制作土地利用和土地覆盖图的方法试验[J]. 测绘通报, 2000, (增刊).
  - [7] HAN J, KAMBER M. 数据挖掘概念与技术[M]. 北京: 机械工业出版社, 2001.
  - [8] 王万森. 人工智能原理及其应用[M]. 北京: 电子工业出版社, 2000.
- (上接第 44 页)
- 参考文献:**
- [1] 苗东升. 系统科学精要[M]. 北京: 中国人民大学出版社, 1998.
  - [2] 万建成. 软件体系结构的原理、组成与应用[M]. 北京: 科学出版社, 2002.
  - [3] BROOKS F P. The Mythical Man-Mouth[M]. [s.l.]: [s.n.], 2002.