

草场资源遥感调查专家系统模型

程 涛 李德仁 舒 宁

摘 要 本文根据专家系统及草场综合分类原理研制了一个草资源遥感调查专家系统原型(GRIES)。GRIES具有与一般专家系统相同的结构,并以模块化编程。系统中设计了一个非精确推理模型以对知识和数据的不确定性与不完全性进行表示和处理。实验结果证明 GRIES能充分利用各种辅助信息及专家的知识 and 经验进行草场资源调查,提高了影象的分类精度。通过对知识库及数据库的适当更改,GRIES还可用于植被、土地利用等遥感资源调查领域。

关键词 遥感;影象分类;专家系统;非精确推理;草场资源

1 引 言

随着信息技术和航天技术的发展,遥感技术在国土资源调查、环境监测和规划管理决策领域发挥了重要作用。用遥感进行草场资源调查,可对大面积的区域进行调查,且可省去许多野外调查的工作,缩短了调查周期。同时,用计算机可对各种资料进行综合处理,提高了数据处理的速度和精度。因此,利用遥感及计算机技术在某种程度上提高了常规草地调查的速度和精度,可以更好地为草场动态监测服务。但在目前已有的一些应用实例中,还普遍地存在着不完善的一些问题,主要有:

其一,目前对遥感影象分类常用的是基于地物光谱的统计分类方法,这是种用影象特征来进行分类的方法。因而,不同的人会对影象有不同的理解;选择的样区各异,会导致分类结果各异。因此,仅用影象特征来进行分类,效果是不甚理想的。其二,现在较好的分类方法是利用辅助信息对分类结果进行修正,这就要用到许多空间数据及相关的属性数据。如何对这些信息进行恰当的处理和应用是一个十分棘手的问题。目前一般仅能结合一种或两种辅助信息,这种结合是不完全的。其三,上述的辅助信息多为统计数据,是不精确和不完全的,难以用基于算法和精确模型的过程式语言对其进行恰当的描述和处理。

因此,现存的遥感计算机草资源调查系统,由于缺乏特定的专家知识的支持和恰当的知识及数据表示形式,对系统中大量辅助数据的使用是低效和不完全的,难以真正达到专家处理问题的水平。

专家系统是一种具有特定专业领域求解的专家知识和技能的计算机程序系统,它利用人工智能技术,用来源于一个或多个专家的知识进行推理,模拟专家在解决实际问题时处理问题

的过程,并以专家水平提供答案。专家系统以描述性语言实现,能以灵活的符号表示知识和数据,特别适合于任务复杂且任务需要符号操作和启发式知识的研究领域。目前专家系统已被广泛应用于语言理解、图象分析、医学诊断、林业规划、地质勘探等领域,并取得了不少成功的经验。

在测绘领域内,人工智能、专家系统的应用也十分活跃,在解析图象匹配、航空摄影相片识别、多光谱遥感影象分类、地图制图专家系统及智能地理信息系统等方面的应用研究正在逐步深入。本文旨在将专家系统技术与遥感技术结合起来进行草资源调查。设计了一个草资源遥感调查专家系统原型(GRIES),用逻辑编程的方法实现,并以实地数据进行了论证和分析。此项研究也属于多光谱遥感影象分类专家系统的范畴。因此,本文也将对专家系统用于遥感图象处理的问题作一些有益的探讨。

2 草场资源遥感调查专家系统模型(GRIES)

为了将草场资源调查原理用专家系统方法实现,本文用图1所示的模型来完成。GRIES的基本结构与一般的专家系统是相似的,亦使用模块化编程方法,包括四个大的模块:

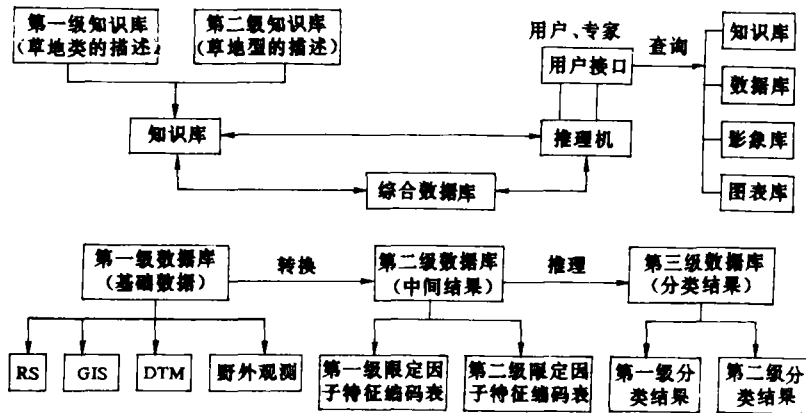


图1 GRIES的结构框图

1. 综合数据库模块:它是一个综合数据库,含有草地调查所需的各种图形、影象及相关的属性数据。此外,数据库中还存在了许多中间运算结果及最终的分类统计结果;

2. 知识库模块:知识库中的规则体现了草场综合顺序分类的原理(示意图见图2),反映了草地类型与各种因素(气候、地形、地貌及影象色调等)的关系,这是进行草场分类的主要依据;

3. 推理机模块:推理机引导系统匹配数据、事实和规则进行推理,得出分类结果。GRIES属于一个诊断系统(结论已知的问题),故选用逆向推理机制,以提高推理效率。此外,该模块中含有控制整个系统运行的机制。

4. 用户接口模块:它运用下拉弹出式菜单以实现人机对话,便于用户对系统操作。即通过用户界面,用户可以对知识库和数据库进行查询,并能了解推理路线。而且,系统也可以向用户询问推理所需的初始数据等信息。

GRIES 是基于谓词逻辑的专家系统,它的推理机用 Turbo-Prolog 语言实现,其它的数据处理功能由 Turbo-C 完成。整个系统可在 IBM-PC/AT,XT 及其兼容机上实现。

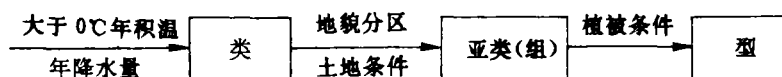


图2 草场综合分类顺序

3 GRIES 的组成部分及功能

本节按图 1 所示的四个模块来介绍 GRIES 的组成及各部分的功能。

3.1 草资源综合数据库

全库由三级组成。

1. 第一级数据库(基础数据库):是指关于草场可获得的最原始的资料数据。这些资料包括各种遥感数据、GIS 数据、地形数据及其它的野外测量数据。对各种专题图件及地形图,先进行数字化,然后灰度化,与影象数据一起以栅格形式存入第一级数据库中。

2. 第二级数据库(中间结果数据库):是指由基础数据经过一定的转换而得的,其数据均以编码形式表示,且接 Prolog 所要求的表结构存贮,以便与知识库中的知识表示相匹配,便于推理工作的进行。从第一级数据到第二级数据库主要有以下几种转换:

气候数据的影象化:气候因子决定了草地的类,将其影象化,以便与影象数据复合,以式(1)作为气候模型^[2]。

$$L_{ij} = a_0 + a_1(h_{ij} - h_0) + a_2 \cdot i + a_3 \cdot j \quad (1)$$

式中: i, j ——影象象素的行、列值; L_{ij} ——象素 (i, j) 上的气候观测值,如年积温、年降水量等;

h_{ij} ——象素 (i, j) 所对立的高程; h_0 ——起始点 $(i=j=0)$ 的高程; a_i ——待求系数 $(i \in [0, 3])$

由数字地面模型(DTM)生成坡度、坡向数据,以得到地形因子。

生物量指标的计算:生物量指标反映了植被的生长状况,按下式计算:

$$\text{生物量指标} = \frac{\text{波段 7} - \text{波段 5}}{\text{波段 7} + \text{波段 5}} \quad (2)$$

色度空间的变换:为了利用专家对遥感影象进行目视解译时所用的色调判读经验,必须选择一个恰当的色度空间,这个色度空间应和人眼观察色彩的习惯相一致,而且易于在计算机上实现。为此 GRIES 选用了 HSV 色彩模型^[3],将 MSS 的 7、5、4 三波段作为计算机可实现的色彩模型——RGB 电视颜色空间的红(R)、绿(G)、兰(B)三原色,实现 RGB 到 HSV 色度空间的转换,以模拟专家目视解译时所定义的色彩。

地貌因子的提取:地貌因子是确定草地亚类和型的重要因子。若草场区内有地貌图,则可直接将其数字化,反之,可根据其地形图进行勾绘。

3. 第三级数据库(结果数据库):两级草场分类结果(“类”、“型”)及相关的统计数据就构成了结果数据库。

综上所述,建立综合数据库的过程实质上就是对原始资料进行预处理的过程,只有完成这

一系列的工作,并将数据库以与知识库、推理机相匹配的格式存贮,才能便于推理工作的完成,得到最终的结果。

3.2 GRIES 的多级知识库

根据草场综合顺序分类的原理,将知识库分为两级。第一级中存贮确定草场类的规则,即由气候因子来确定草场类的知识;第二级中存贮确定地型的规则,即根据地形、地貌和影象色彩特征来确定草地型的知识。这两级规则具有相同的知识表示形式,均以 Turbo-Prolog 的谓词逻辑表示,这样就便于使用相同的推理机制和统一的数据表示形式。

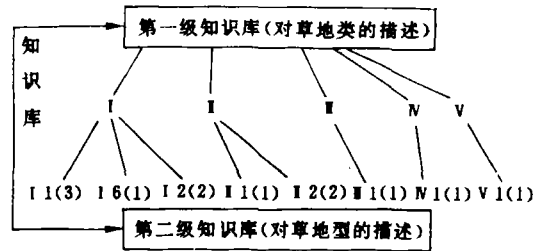


图3 GRIES 知识库的层次结构

对知识库进行分级的目的是为了便于对知识库管理,使草地类和型的确定在不同的层次上进行,这样就缩小了搜索空间,提高了推理效率。而且,第二级推理是在第一级推理完成后进行的,即先确定草地的类后再确定草地的型,将知识库分级后就可实现分级推理,使推理路线简洁清晰。图3为知识库分级层次示意图。

3.3 GRIES 的非精确推理模型

GRIES 中大多数的规则来源于《西藏那曲地区草资源调查报告》,限定因子特征区间的划分也是根据报告中给出的数据确定的。这些数据以及专家对草地判别经验的描述是不十分精确的和不完全的,有的甚至是模糊的。因此,在 GRIES 中引入不精确推理。目前,处理不确定性的方法有许多,有代表性的是确定性理论(MYCIN 中的)、主观贝叶斯方法(PROSPECTOR 中的)、D/S 证据理论及发生率计算等等。这些方法各有优缺点^[4],都不能完全适用于 GRIES,为此笔者自行设计了一个不精确推理模型。

3.3.1 限定因子的可信度

GRIES 中的知识和数据是不精确和不完全的,表现在对各类草地的描述中,所给的限定因子区间是不连续的。这样就有许多分类单元的特征值在两个区间之间,为了后续的推理,必须将这些特征值归并到区间之内。为了将这些强行归并的编码(不确定性的)与本来就落在区间内的编码(确定性的)区别开来,在 GRIES 中引入了限定因子编码可信度(CF)概念。

定义:某一限定因子有两个特征编码区间 $[a, b]$ 及 $[c, d]$,其特征编码分别为 e_1 及 e_2 。现有一分类单元,其特征值为 $x \in (b, c)$ 。则定义:

将 x 归并到区间 $[a, b]$ 的可信度为

$$CF_1 = \left| \frac{x - c}{c - b} \right| \quad (3)$$

将 x 归并到区间 $[c, d]$ 的可信度为

$$CF_2 = \left| \frac{x - b}{c - b} \right| \quad (4)$$

而 x 的最终可信度 CF 及其特征编码 e 按下式判断:

$$\begin{cases} CF = CF_1 & e = e_1 & \text{若 } CF_1 \geq CF_2 \\ CF = CF_2 & e = e_2 & \text{若 } CF_1 < CF_2 \end{cases} \quad (5)$$

对于特征值属于某个确定区间的,其可信度为 1。

3.3.2 GRIES 的不精确推理模型

不精确推理模型一般包括三个方面^[9],在 GRIES 中表现为:

(1)知识的不精确刻划:由于专家并不能明确回答每条规则的可信度,笔者姑且认为所有规则的可信度都是相同的,且将所有规则的可信度均设为 1,即 $CF(H, E)=1$;

(2)初始证据的不精确刻划:即前提因子(限定因子)的可信度 $CF(E)$,由式(5)给出;

(3)不精确更新算法:

若有规则

IF $\Delta E_1 \Delta E_2 \Delta \dots \Delta E_n$. THEN $H(CF(H, E))$

则结论的可信度为:

$$CF(H) = CF(H, E) \max\{0, CF(E)\}$$

$$\therefore CF(H, E) = 1 \quad CF(E) > 0$$

$$\therefore CF(H) = CF(E) = CF(E, \Delta E_1 \Delta \dots \Delta E_n)$$

$$= CF(E_1)K_1 + CF(E_2) \cdot K_2$$

$$+ \dots + CF(E_n) \cdot K_n$$

$$= \sum_{i=1}^n CF(E_i) \cdot K_i$$

$$(K_i = 0 \text{ 或 } 1) \quad (6)$$

其中 $CF(E_i)$ 为第 i 个限定因子编码 E_i 的可信度, n 为规则前提中限定因子的数目。

笔者认为选择的限定因子数目愈多,则分类结果愈准确。但从实际的角度考虑,某个分类单元可能不满足规则中所规定的全部条件,即其特征编码仅与规则中的条件编码部分元素相匹配,它仍可能是此规则中所描述的草场类型。因此,用 K_i 来表示两者的元素是否匹配。若匹配则 $K=1$,否则, $K=0$ 。这样又会导致一个分类单元与多个规则匹配,解决冲突的办法就是将所有规则结论的可信度作比较,取其值最大的结论作为此分类单元的分类结果。

整个不精确推理过程可用图 4 示意。

3.4 用户界面

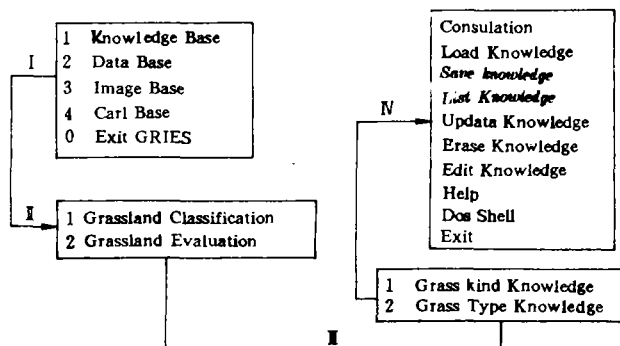


图 5 用户界面分层菜单示意图

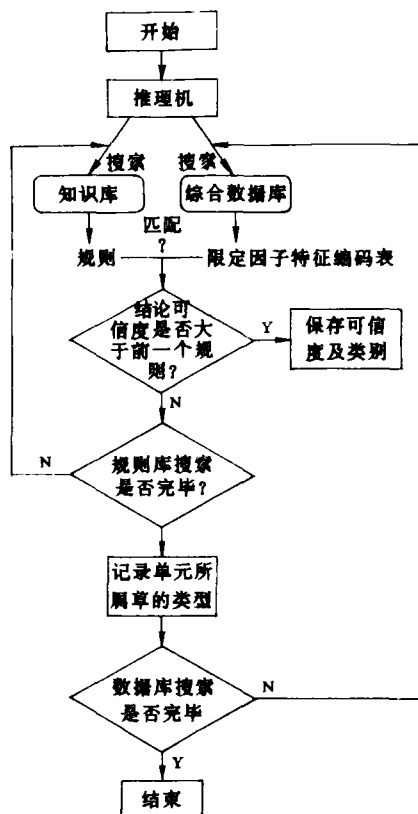


图 4 不精确推理模型

目前,用户界面一般有菜单、文本及图形三种方式。GRIES 选用下拉式菜单,通过对不同层次菜单的选择,用户可对系统进行不同的操作。GRIES 的菜单分为四层,可用图 5 作简单的示意。

本节所述的四个模块是相互联系的,GRIES 已解决了四个模块的内接联接,但需要进一步完善它们的外部接口。

4 实例与结果

本试验以西藏自治区那曲地区班戈县为试验对象,试验中用到了试验区以下资料(班戈县的):

- 遥感资料:1986 年夏季的 MSS 的 7、5、4 三个波段影像,由它们可计算生物量指标及进行色度空间变换。(选取 200×200 个象元大小的区域为试验区);

- 高程资料:西藏自治区 1:10 万的地形图,由此生成 DTM、坡度、坡向文件及进行地貌勾绘;

- 图形资料:包括草地等级图(1:20 万)、草地类型图(1:20 万)及土壤类型图(1:50 万);

- 气候资料:那曲地区五个气象台站从 1971~1980 年的观测记录,由此生成气候模型;

根据《西藏那曲地区草资源调查报告》及兰州草原生态研究所的草地专家经验建立了知识库,两级知识库中有近六十条规则。

主要完成以下实验内容:

首先,用最大似然法对影像进行预处理,实验结果表明分类结果破碎,且与样区的大小及数目的多少有较大关系,即受主观因素的影响,且拒分类较多。其次,不进行影像分割,直接以单个象元为分类单元,按精确推理和非精确推理两种模型进行分类。精确推理模型所得的结果中有许多错分和拒分的情况。而在非精确推理中则不存在拒分类。此外,为了提高系统的运行速度,采用以区域生成法对生物量指标影像进行分割后的斑块为分类单元进行分类。三种方法所得结果的比较见表 1。

表 1 三种分类方案的比较

比较项目 方法名称	相差象元个数	正确率 (%)	运行时间			
			预处理	分类	后处理	总计
最大似然法	10885	72.79	30min	2h	1h	3h30min
以象元为分类单元	4292	89.27	45min	3h	1min	3h46min
以斑块为分类单元	6932	82.67	5min	4min	1min	10min

5 结束语

从以上的试验及在构建 GRIES 的过程中,笔者有下列体会:

- 构建专家系统,应用到混合编程技术,即用过程性语言进行数据处理,而用描述性语言设计推理机,以提高整个系统的运行效率。

· 对知识的描述应用选择充分而无冗余的限定因子。且对知识库和推理机应根据情况实行分级设计,以便对知识库进行管理和提高推理效率。

· 对不精确知识和数据采用不精确推理是一种较好的方法。它不仅可对知识和数据的不确定性进行恰当的描述,而且能用恰当的算法进行处理,在某种程度上提高了知识和数据的利用率,也提高了分类精度。且不精确推理模型可根据实际情况自行设计。

· 用专家系统方法实现草资源遥感调查,能够利用多种辅助信息并结合专家经验,比用传统的监督分类法精度高、速度快且自动化程度好。GRIES作为草资源遥感调查专家系统的一个原型,对其知识库、数据库作适当的修改、完善和扩充,它还可用于森林、植被、土地利用及农作物生长预测等领域。

遥感、地理信息系统与专家系统的结合正在方兴未艾,这种结合必将推动地理信息系统及遥感技术更加深入广泛的应用。

参 考 文 献

- 1 舒宁.草场分类原理的计算机实现.武汉测绘科技大学学报,1989(2):25~32
- 2 王野乔等.地学编码数据与遥感信息进行长白山气候研究.遥感信息,1990(2):8~10
- 3 吴健康.数字图象分析.人民邮电出版社,1989,49
- 4 卢有杰,吴炜煜,王家厥编译.专家系统技术指南.世界图书出版社,1989,210~220

An Expert System Model of Grass Resource Investigation by Using Remote Sensing Data

Cheng Tao Li Deren Shu Ning

Abstract According to the theory of expert system and the principle of complex classification of grass land, a prototype of grass resource investigation expert system by remote sensing is developed. The system——GRIES(Grass Resource Investigation Expert System) has the same structure as a usual expert system. The authors design an model of uncertain inference to solve the uncertainty and uncomplety of the knowledge and the data. With the experiments of some examples, a combination of the technique of expert system and remote sensing can make full use of various auxiliary-information to overcome the disadvantages existing in maximum-likelihood classification. At the meanwhile, GRIES can take the advantage of the knowledge and experiences of the experts to improve the accuracy of classification. After some changes of the knowledge-base and the data-base, GRIES can also be used in the areas such as vegetation field analysis, landuse investigation, predication of growth of crops and other kinds of resources investigation by remote sensing.

Key words remote sensing; image classification; expert system; uncertain inference; grass resource