

基于拓扑约束理论的森林小班自动聚类分析

薄志义^{1,2}, 鲁绍伟³, 罗旭³, 景海涛⁴, 曹志刚⁵

(1. 中国矿业大学 资源与安全工程学院, 北京 100083; 2. 北京工业职业技术学院建筑工程系, 北京 100042;
3. 北京林业大学 资源与环境学院, 北京 100083; 4. 徐州师范大学, 江苏 徐州 221116;
5. 河北科技师范学院, 河北 秦皇岛 066004)

摘要: 为提高林业调查规划专题图编制的工作效率, 讨论了基于拓扑约束理论的数字专题图制作方法和原则, 利用点位的拓扑关系建立小班地域的邻接矩阵, 和小班属性矩阵作为自动聚类的基础矩阵, 并通过 MapBasic 编程序对 MapInfo 进行二次开发, 结合 Excel 等软件, 实现了小班多边形自动聚合及分类, 为林业专题地理信息系统研究奠定基础。小班多边形聚类方法与人机交互操作相比, 具有速度快、精度高等特点, 克服手工制图工作量大、速度慢、成本高等缺陷。

关键词: 自动聚类分析; 小班; 约束理论; MapInfo
中图分类号: S711 文献标识码: A 文章编号: 1005-3409(2006)06-0013-03

The Auto-clustering Analysis of Subcompartment Based on the Constrained Theory

BO Zhi-yi^{1,2}, LU Shao-wei³, LUO Xu³, JING Hai-tao⁴, CAO Zhi-gang⁵

(1. School of Resource and Safety Engineering, China University of Mining and Technology, Beijing 100083, China;
2. Beijing Vocational & Technical Institute of Industry, Beijing 100042, China;
3. School of Resource and Environment, Beijing Forestry University, Beijing 100083, China;
4. Xuzhou Normal University, Xuzhou, Jiangsu 221116, China;
5. Hebei Normal University of Science & Technology, Qinhuangdao, Hebei 066004, China)

Abstract: To improve the efficiency of making the theme-map of forest investigation and planning, the authors discussed the methods and principles of making digital theme-map based on the Constrained recognition and clustering theory. The adjoining-matrix of points on the earth can be made based on Constrained connection, and while it is the basic matrix with attribute-matrix of auto-clustering. Based on MapBasic and Excel, a program is made to realize the second-develop of MapInfo, the auto-combination of sub-compartment polygons can be achieved. This research is groundwork of thematic-GIS. That it not only automated the combination of sub-compartment polygons, but also raised the data quality compared with the method alternated between man and computer, this method avoids some limitations of heavy-work, low-efficiency and high-cost.

Key words: auto-clustering analysis; subcompartment; Constrained theory; MapInfo

森林分布图编制中工作量最大、精度最难控制的是小班多边形聚类。近年来, 在森林资源调查、林业规划等图件编制工作中已广泛应用地理信息系统(GIS)技术, 已基本上实现计算机化^[1-3]。基础地理信息系统的功能面向的对象是多方面, 其主要操作是地图的空间数据采集、空间数据处理及空间数据分析, 但是针对具体专业的专题地理信息系统研究还不是很深入, 因此在编制森林分布图等林业调查规划设计专题图时, 一般采用人机交换的方法进行, 存在着工作量大、速度慢、成本高等问题^[4]。本文以森林小班分布图制作为例, 讨论了拓扑约束理论在林业调查与规划中的应用原则与方法, 外业调查及室内数据处理过程中不同属性的编码原则, 并基于 MapInfo 的二次开发编程实现^[5-6]。结果表明, 该小班多边形聚类方法与人机交互操作相比, 具有速度快、精度高等特点。

1 森林小班分类图制作的一般方法

小班分类图制作方法很多, 但多以外业调查建立空间数据库为基础, 通过专用软件对调查原图进行扫描矢量化或跟踪矢量化, 形成众多的数字多边形(polygon)小班分类结果^[7]。实际操作过程不可避免地要进行图形编辑、图形检查等大量的手工劳动, 由于详细的森林分布图, 通常包括数万个小班, 数据量巨大, 给室内工作带来诸多不便甚至出现错误。该种工作流程不便数据的快速显示及属性定义, 分类结果没有为后续进一步工作打下良好基础。

手工制作小班分类图的一般工作流程:

- (1) 建立外业调查的小班空间数据库;
- (2) 对小班原图进行矢量化(建立新的图层);

* 收稿日期: 2006-04-21

基金项目: 国家 863 课题(2003AA245030); 国家自然科学基金项目(90302014); 北京市自然科学基金项目(4032015) 共同资助
作者简介: 薄志义(1964-), 男, 辽宁锦州人, 博士生, 副教授, 主要从事测绘科学、地理信息系统教学与研究。

- (3) 在新建图层上按一定原则进行邻接小班的聚合, 并对图斑填充颜色.;
- (4) 编制图例、图框、图头和图街落款;
- (5) 以适当尺寸的图纸打印输出。

2 拓扑约束理论

基于拓扑约束理论是多元分析的一个分支^[8,9]。本文提出利用点位的邻接拓扑关系构成约束识别思想来判断森林小班的分布。其作法是将研究范围概化成节点图, 采用矩形节点分布形式。首先对节点的小班属性、联接方式以及拓扑关系进行分析, 根据属性变化的一致性对节点进行分类, 同

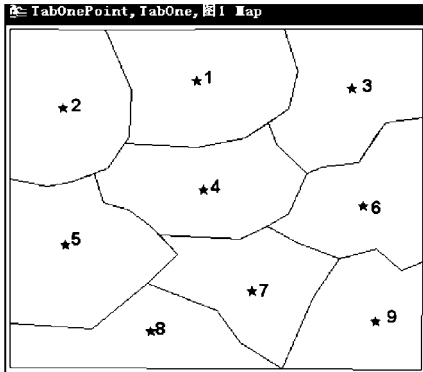


图 1 概化多边形节点分布

节点的拓扑约束矩阵 A 为:

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & \dots & 0 \\ 1 & 1 & 0 & 1 & 1 & \dots & 0 \\ 1 & 0 & 1 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 1 & 1 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 \end{bmatrix}_{9 \times 9}$$

2.2 邻接矩阵生成图的遍历序列

图的遍历是从图上某一顶点出发访问图中其余顶点(如图 2), 且使每一个顶点仅被访问一次, 遍历图的过程实质是通过边或弧寻找邻接点的过程。假设初始状态是图中所有顶点(n) 未曾被访问, 从图中某个顶点 v_1 出发, 访问此顶点, 然后依次从 v_1 的未被访问的邻接点出发优先遍历图, 直至图中所有和 v_1 有路径相通的顶点都被访问到; 若此时图中还有未被访问的顶点, 则从一个未被访问的顶点开始, 重复上述过程, 直至图中所有顶点都被访问为止。显然, 这是一个复杂的递归过程。基本步骤如下:

- 构造图的邻接矩阵;
- 将所有相关边或弧串连成一个或多个“非闭合”的边或弧线性表, 直至串连的边或弧数为 $n-1$;
- 出发顶点与右关联顶点顺序连接优先遍历序列 (dfs)。

根据图 2 地表节点及邻接矩阵, 生成的边线性表:
(1, 3) (3, 6) (6, 9) (9, 8) (8, 5) (5, 2) (2, 4)
(4, 7)

图 2 的 dfs 序列是: 136985247
由于在访问出发顶点后接下来访问顶点都是随机的, 因此边线性表不是惟一的。

3 小班属性定义及编码

3.1 小班属性定义

外业森林调查过程中就对众多森林小班进行属性定义

一类的点构成一个同属性区域。但是, 这可能会造成不连通的分类, 即相隔很远的一致性点被归入同一单元, 或者一个单元内部的某一点属于另一个单元。这种纯统计学上的识别结果可能造成物理上的矛盾。因此, 要达到合理的分类就必须兼顾点的相似性、几何上的连通性以及其它先验信息。

2.1 地域邻接矩阵表达

为得到具有连通性的聚类结果, 定义一个描述节点邻接拓扑关系的拓扑约束矩阵 A :

$$A = [\alpha_{ij}]_{n \times n}$$
$$\alpha_{ij} = 0 \quad i, j \text{ 不相邻}$$
$$\alpha_{ij} = 1 \quad i, j \text{ 相邻或 } i = j$$

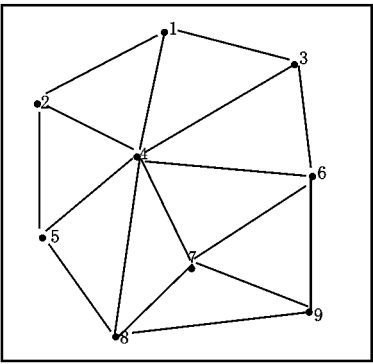


图 2 节点连接图

(编码), 不同的多边形 ID 号其属性可能相同, 也可能不同, 编码表示所代表的小班类型等特征, 如图 1 的模拟外业属性编码如下:

TabOnePoint Browser				
	ID	Area	Sort	Code
<input type="checkbox"/>	1	1,235		A
<input type="checkbox"/>	2	342		A
<input type="checkbox"/>	3	157		B
<input type="checkbox"/>	4	432		C
<input type="checkbox"/>	5	821		A
<input type="checkbox"/>	6	257		B
<input type="checkbox"/>	7	193		B
<input type="checkbox"/>	8	164		A
<input type="checkbox"/>	9	246		A

图 3 小班调查数据库

图 3 中 Code 字段表示小班的不同属性, 是小班自动分类聚合的基础, 相同的属性如果多边形是相邻的则自动聚类在一起, 虽然属性相同但物理上不连通也不能聚类一起。

多边形定位可采用 GPS 采集多边形内的坐标, 输入 GIS 进行配准。

3.2 建立编码矩阵

根据调查编码建立森林小班的编码矩阵如下:

$$B_{n \times n} = \begin{bmatrix} A & 0 & 0 & 0 & \dots & 0 \\ 0 & A & 0 & 0 & \dots & 0 \\ 0 & 0 & B & 0 & \dots & 0 \\ & & \dots & \dots & \dots & \\ 0 & 0 & 0 & 0 & \dots & A \end{bmatrix}$$

4 森林小班的自动聚类

4.1 小班多边形聚合原则

小班多边形邻接, 主要属性相同或相近, 根据需要及外

业调查属性,可依据树种、坡度等在输出图纸上,保留的最小图班尺寸为:规整多边形 $\geq 2\text{ mm} \times 2\text{ mm}$,带状多边形一般宽度 $\geq 2\text{ mm}$ 。同时还要考虑土地种类优先级及优势树种(组)优先级^[2]

4.2 搜索约束聚类对

聚类原则是属性变化一致而又邻接的点作为一类,因此可根据矩阵 A 、 B 进行搜索。通过搜索,若 L 、 m 两点即满足属性变化一致性又同时满足邻接性,则 L 、 m 两点可归入同一个小班单元。

改造高差矩阵与拓扑约束矩阵:

l 、 m 两点聚到一起后,可将它作为一点对待,用 (l, m) 表示。

这样,原 n 阶值矩阵变为 $n-1$ 阶的矩阵,约束矩阵 B 也应降一阶。

根据新的邻接矩阵与拓扑约束矩阵,重复(1)和(2)步骤,直到将所有同属性点聚为一类。

同一类属性的小班标注相同颜色。

5 MapBasic 编程关键及实例

5.1 MapBasic 的合并对象并改变标注等属性

合并对象是指将多个单独的地图对象合并成一个对象,根据约束矩阵聚类结果,应用 Combine() 函数实现多个邻接属性相同的多个对象合并。

合并后的数据属性表可通过两种方法自动建立,其一是通过 MapBasic 的 Open Table() 函数打开其他软件形成的数据表(如 Excel)实现数据连结;其二是直接应用 col_name=Sum(col_name) 等函数编程实现。

AutoLabel() 函数实现合并后的图层进行重新标注,并在装饰图层中保存标注内容。

MapBasic 提供的 Shade 函数可以创建多种类型的专题图。

5.2 实例

通过 MapBasic 编制程序对原图图 1 进行自动分类,图

1 中根据外业编码可知,其分布形式为三类:

其中:

第一类多边形为(1, 2, 5, 8, 9),其聚合结果应为(11, 2, 5, 8, 9);

第二类多边形为(3, 6, 7),其聚类结果为(3, 6, 7);

第三类多边形为(4),其聚类结果为(4)。

MapInfo 分类结果图如图 4:

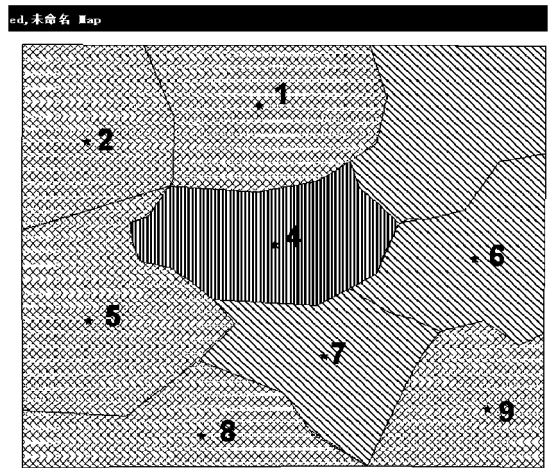


图 4 MapInfo 聚类结果

6 结论与讨论

(1) 小班多边形聚合是林业调查规划专题图编制中必不可少的工作。基于约束理论及地理信息系统采用自动聚合程序,不但大大减轻工作量,加快工作进度,而且避免了人机交换工作带来的误差,效果良好。

(2) 自动聚类结果通过地理信息系统实现与自然景观的热链接,使森林分布可视化。

(3) 基于约束理论的小班自动聚类易于程序实现,可基于 VB、VC 编程实现与 GIS 通讯,也可应用 GIS 的二次开发工具进行编程实现。

参考文献:

- [1] 吕洪利,武刚,卢泽洋. 基于 MAPINFO 的森林资源空间信息管理系统开发与应用[J]. 北京林业大学学报, 2001, 23(3): 87- 89.
- [2] 李春干,杨崇贵,景海涛. 森林小班多边形的自动聚合[K]. 北京林业大学学报, 2005, 27(1): 55- 58.
- [3] 王元胜,甘长青,周肖红. 香山公园古树名木地理信息系统的开发技术研究[J]. 北京林业大学学报, 2003, 25(2): 53- 57.
- [4] 朗璞玫,王相国,徐建民,等. 在 Autocad 平台上建立具有 GIS 特征的森林图形系统[J]. 北京林业大学学报, 2001, 23(3): 90- 93.
- [5] 吴立新. 林业计算机辅助制图技术初探[J]. 云南林业科技, 2002, (4): 71- 74.
- [6] Harris J R, Wilkinson L, Heather K, et al. Application of GIS Processing techniques for producing mineral prospectivity maps- a case study: mesothermal au in the Swayze Greenstone Belt, Ontario, Canada[J]. Natural Resources Research, 2001, 10(2): 91- 124.
- [7] M EL- Raey, Y Fouda, Gal P. GIS for environmental assessment of the impacts of urban encroachment on rosetta region, Egypt[J]. Environmental Monitoring and Assessment, 2000, 60: 217- 233.
- [8] 陆守一,唐小明,王国胜. 地理信息系统[M]. 北京: 中国林业出版社, 1998. 105- 108.
- [9] 蒋征,张正禄. 变形模式的拓扑约束识别[J]. 测绘学报, 1999, 28(4): 330- 334.
- [10] 蒋征,张正禄. 滑坡变形的模式识别[J]. 武汉大学学报, 2002, 27(2): 127- 132.
- [11] 刘永杰,冯仲科. 营林 GIS 数据采集过程中质量控制的研究[J]. 北京林业大学学报, 2003, 25(2): 48- 52.