

# 图斑数据自动概括中面向地理特征的层结构选取

卢林<sup>1</sup> 吴纪桃<sup>2</sup> 柳重堪<sup>1,2</sup>

(1 北京航空航天大学电子信息工程学院, 北京市海淀区学院路37号, 100083)

(2 北京航空航天大学理学院, 北京市海淀区学院路37号, 100083)

**摘要:**在信息定义和结构化选取的基础上, 根据图斑的空间分布与距离联系, 给出图斑数据地理要素层结构的定义和建立过程, 从而实现图斑数据制图概括中对象选取的模型化。

**关键词:**制图综合; 结构化; 聚类

**中图法分类号:** P283.1; P208

制图概括的自动化是近来 GIS 研究领域内的热点和前沿问题, 但是对制图概括问题的解决远没有达到人们所期望的程度, 主要原因在于制图概括问题本身的复杂性。传统的制图概括主要依靠制图员人工操作, 最终的制图结果掺杂了大量的主观因素, 因而不是不稳定的。

图斑数据作为一类特殊的地理要素数据, 记录着诸多类型的地理信息。本文通过对图斑数据自身的地理特征分析, 总结出图斑空间层的定义。在空间距离关系的基础上, 建立图斑的层结构关系; 在制图概括的内容选取过程中, 实现图斑数据地理特征及层次性表达。

## 1 图斑数据的制图概括

### 1.1 图斑数据的数量选取

Toeper<sup>[1]</sup> 在研究制图概括时发现, 地物要素的选取数量与地图比例尺之间有着密切的关系, 于是建立了开方根模型的基本公式:

$$N_b = N_a \cdot \sqrt{M_a/M_b}$$

其中,  $N_a$  为原图地物数量;  $N_b$  为新图上选取的地物数量;  $M_a$  为原图比例尺;  $M_b$  为新图比例尺。

诸多制图学者们对该公式的趋向性意见认为, 开方根规律对编图实践具有指导意义, 特别是对于离散分布的点状和面状要素的选取基本上是

正确的。后来, 开方根规律公式得到扩展, 增加了两个系数。在本文中, 原数据对符号基本没有依赖性, 故采用上述开方根公式。

### 1.2 图斑数据的内容对象选取

作为制图概括的主客观依据, 地理认知贯穿于制图概括的整个过程。一方面, 制图概括本身就是一个主观与客观相统一的过程; 另一方面, 制图工作者在地理认知中不仅全面、综合地分析和理解地理环境, 而且其认知中还包含了对一些客观的制图标准和规范的接收或者认同<sup>[3]</sup>。

随着比例尺的缩小, 地图的几何精度逐渐让位于地理形态与结构的表达, 这一点对于制图概括具有很重要的指导意义<sup>[4]</sup>。

在确定了选取的数量之后, 具体选择哪些对象就成为最关键的问题。问题的核心在于结构识别, 即数据对象的重要性排序, 不仅包括在全局中的重要性指标, 还包括数据相互之间的相对重要性指标。对于图斑数据而言, 衡量重要性的因素主要有位置、大小和形状, 有时候需要针对不同类型的图斑增加一些因素。在比例尺变换过程中, 形状的重要性随着比例尺缩小而降低, 所以本文只用图斑的位置和大小这两个因素来建立图斑数据有序的层结构, 从而获得数据对象选取的顺序。

给定图斑数据集  $S = \{ \{ K_i, i = 1, \dots, N \} \mid K_i = \{ (x_j, y_j), j = 1, \dots, M_i \} \}$ , 其中,  $N$  为数据集中图斑的个数,  $M_i$  为第  $i$  个图斑的轮廓的坐标点数

据个数。通过计算可以得到图斑的面积数据  $\{A_i, i = 1, \dots, N\}$ 。图斑单体的位置参数为:

$$C_i = \left( \sum_j x_j / M_i, \sum_j y_j / M_i \right), (x_j, y_j) \in K_i$$

图斑单体的大小参数为:

$$R_i = N \cdot A_i / \sum A_i$$

$R_i$  可以衡量每个图斑在整体中所处的等级或地位, 而  $C_i$  则可以反映整体图斑的分布情况。有了这两个参数, 可以联合组成一个三维矩阵(位置参数是二维的)。

在图斑数据的制图概括过程中, 当某个图斑的面积在新比例尺下面积很小 ( $< 0.2 \text{ mm}^2$ ) 时, 由于人的视觉精度有限, 故可删除。本文在图斑数据的层结构建立的过程中, 没有考虑删除的情况, 这种情形可在图形化简时加以处理。

有了图斑数据的两个数值参数  $C_i$  和  $R_i$ , 可以定义图斑之间的距离:

$$d(K_i, K_j) = \sqrt{(C_i - C_j)^2 + (R_i - R_j)^2}, \\ 1 \leq i, j \leq N$$

下文中简记为  $d_{ij}$ 。有了距离的定义, 可以计算图斑数据的距离矩阵  $D = \{d_{ij}, i, j = 1, \dots, N\}$ 。利用距离矩阵, 可以按照如下方式建立图斑数据的层结构关系。

1) 将所有图斑都定义为第一层(最低层)的数据;

2) 在  $D$  中选择最小的元素, 设为  $d_{ij}$ , 则  $d_{ij}$  对应的图斑  $K_i$  和  $K_j$  形成新层(图斑组)  $K_r$ ;

3) 重新计算其他图斑  $K_l$  与这个新图斑组  $K_r$  的距离 ( $l \neq r$ ), 得到新的距离矩阵  $D_r$ ; 并代替原来的距离矩阵  $D$ ;

4) 回到 2), 在  $D_r$  中重复 2)、3) 操作, 得到  $D$  和一个新层;

5) 如果距离矩阵  $D^n$  (重复了  $n$  次) 中只有一个数值, 则停止层的建立和距离矩阵的计算, 否则一直重复上述 2)、3)、4) 操作。

经过上面描述的层建立过程, 对于任何图斑数据, 都可以得到类似于图 1 中的层结构表达。图 1 中, 横轴对应每个图斑个体, 纵轴是得到新图斑层时的距离数值。

图 1 中最右端的两个数据集(点), 它们在整体分布中所处的层次是最高的, 和其他所有的数据点具有同等的重要性。这样, 就确定了在制图概括过程中内容(对象)选取的先后顺序。当给出了选取的数量以后, 具体图斑的选择顺序就可以

通过上述的层结构的高低得到。通过对图斑数据的层结构顺序的描述, 确定出对于整体地形分布最具有代表性的数据集合。最后通过图形的化简, 就能够解决制图概括的问题。

### 1.3 图斑数据的形状化简

图斑数据经过内容选取后, 需要进行图斑形态特征的概括。首先考虑当比例尺缩小后, 由于面积太小而不能在图中显示的图斑将被删除。另一方面, 当选取的对象属于同一个图斑组, 图斑尺寸大小接近, 并且图斑之间相隔的距离小于比例尺显示能力的时候, 这些图斑就需要被连接成为一个大的图斑。

对于如何生成新图斑, 关键在于搜索两个原始图斑的外围轮廓线, 即新图斑的轮廓数据点。已知两个图斑数据集合  $\{K_i(\{x_n, y_n\}, n = 1, \dots, M_i), K_j(\{x_l, y_l\}, l = 1, \dots, M_j)\}$ , 确定  $K_i$  中的点  $P_0(x_0, y_0)$ , 计算  $P_0$  与  $K_j$  中所有点的距离, 选择最接近比例尺显示能力的距离对应的  $K_j$  中的点  $P_1(x_1, y_1)$ , 当  $P_0$  遍历整个  $K_i$  时, 就可以找到  $K_i$  与  $K_j$  相连的两个端点, 如图 2 所示。这样类似于直接对图斑取舍后, 再对保留的图斑进行夸大处理。

得到了图斑的新轮廓, 并没有完成形状化的化简工作。在比例尺缩小以后, 原图斑中很多细节将不能在数据中表达出来了, 所以需要去除这些信息。小波的多分辨分析框架表明, 对于给定信号  $f(t) \in L^2(R)$ , 可分解为无穷个小波分量的直和:

$$f^j(t) = \sum_k c_k^j \phi_{j,k}(t), f^j(t) \in V_j$$

$$w^j(t) = \sum_k d_k^j \psi_{j,k}(t), w^j(t) \in W_j$$

其中,  $\phi(t)$  为尺度函数;  $\psi(t)$  为小波函数;  $c_k^j$  为近似分量系数;  $d_k^j$  为细节分量系数, 分别代表原信号在低频和高频下的分解系数。如果在信号重构时只保留近似分量或只保留细节分量, 则可以得到原信号的近似信号或细节信号, 而由近似信号和细节信号同样可以重构出原始信号。

近似分量表示的是原信号的粗略形状, 而细节分量则对应于信号中的细节弯曲。当舍弃细节分量, 只在近似分量上重构原信号时, 就可以得到平滑后的数据信息了。对于不同的比例尺, 小波的平滑能力会自动地变化, 非常适合图斑数据最后的形状化简过程。

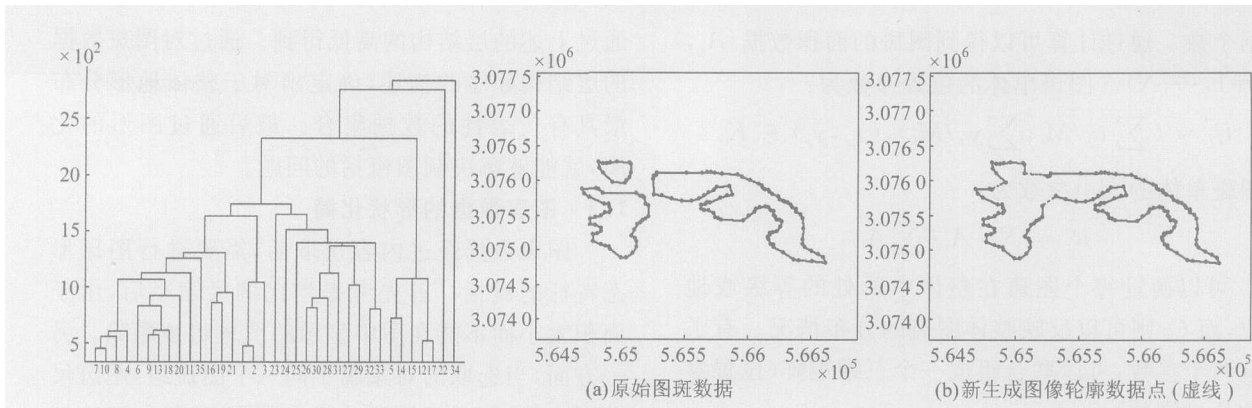


图1 图斑数据的层结构表达

Fig. 1 Layer Structure of Speckle Data

图2 新图斑轮廓的生成

Fig. 2 New Speckle Counter Data

### 2 实验

实验图形(图3)是某海域的岛屿数据图形,经1:20万的比例尺数字化得到。

在开方根规律下,岛屿数量的变化关系为:在比例尺为1:20万、1:40万、1:60万和1:80万的情况下,对应的岛屿数量分别为35、25、20、18。根据图3可建立岛屿的层结构。最后的岛屿概括结果如图4。

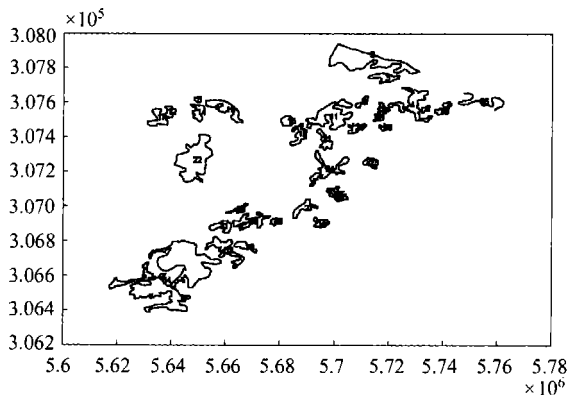


图3 岛屿群数据图形

Fig. 3 Islands Data Figure

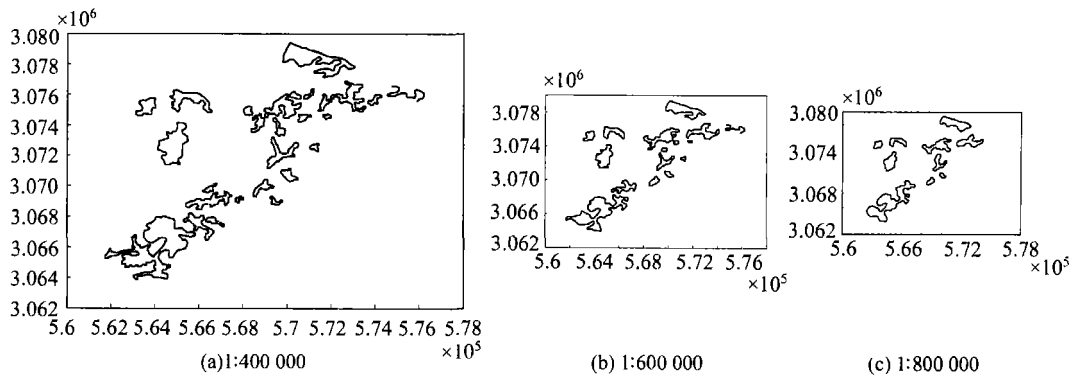


图4 岛屿群多尺度数据表达

Fig. 4 Multi-scale Representation of Islands Data

### 3 结论

从地物的结构特征提取到最佳地保持地图中的地理规律和特性,将是制图概括发展的重要方向。对于图斑数据类型,本文将图斑要素放在三维空间中进行分析,在定义了图斑要素之间的距离后,建立了图斑数据要素的层结构,以衡量图斑要素本身的重要性,从而得出能够保留原数据结

构特征的对象选取方法。最后实验验证了这一方法的有效性和可行性。

#### 参考文献

- [1] Toepfer. 制图综合[M]. 北京:测绘出版社,1982
- [2] 萨里谢夫 K A. 地图制图学概论[M]. 北京:测绘出版社,1982
- [3] 齐清文,刘岳. GIS 环境下面向地理特征的制图概括的理论和方法[J]. 地理学报,1998,53(4): 303-313
- [4] 毋河海. 地图信息自动综合基本问题研究[J]. 武

汉测绘科技大学学报, 2000, 25(5): 377-386

- [5] Scott T A. Scale Space Classification Using Area Morphology[J]. IEEE Trans. on Image Processing, 2000, 9(4)

第一作者简介: 卢林, 博士生, 现从事 GIS 数据多尺度表达研究。  
E-mail: lulin@ee.buaa.edu.cn

## Selection of Speckle-data Based on Layer-structure Construction in Cartographic Generalization

LULin<sup>1</sup> WUJitao<sup>2</sup> LIUZhongkan<sup>1,2</sup>

(1 School of Electronic Information Engineering, BUAA, 37 Xueyuan Road, Beijing 100083, China)

(2 School of Science, BUAA, 37 Xueyuan Road, Beijing 100083, China)

**Abstract:** In this paper, the definition and construction algorithm of layer-structure of speckle data are proposed based on the measure of information and structured selection of speckle-data as well as the spatial distributing and connection, with modeling the selection processing. The subjective influence as well as instability has being released or decreased, with improving the quality and effect of the generalization.

**Key words:** cartographic generalization; geographic characteristics; structured selection

**About the first author:** LU Lin, Ph. D candidate, majors in multi-scale representation of GIS map-data.  
E-mail: lulin@ee.buaa.edu.cn

### 欢迎订阅 2006 年《武汉大学学报·信息科学版》

《武汉大学学报·信息科学版》即原《武汉测绘科技大学学报》,是以测绘为主的专业学术期刊。其办刊宗旨是:立足测绘科学前沿,面向国际测量界,通过发表具有创新性和重大研究价值的测绘理论成果,展示中国测绘研究的最高水平,引导测绘学术研究的方。本刊为中国中文核心期刊, EI 刊源期刊。是国家优秀科技期刊,并获中国国家期刊奖,入选中国期刊方阵。

本刊主要栏目有院士论坛、学术论文、科技新闻等,内容涉及摄影测量与遥感、大地测量与物理大地测量、工程测量、地图学、图形图像学、地球动力学、地理信息系统、全球定位系统等。收录本刊论文的著名国际检索机构包括 EI、CAS、PY 等,其中 EI 收录率达 100%,其影响因子长期名列中国高校学报前列。本刊读者对象为测绘及相关专业的科研人员、教师、研究生等。

本刊为月刊,国内外公开发行,邮发代号 38-317,国外代号 M O1555。A4 开本,96 面,定价 8 元/册,每月 5 日出版。漏订的读者可以与编辑部联系补订。