

文章编号: 1000-0585(2001)05-0629-08

基于空间结构分析的地图概括研究

张青年

(中山大学城市与资源规划系, 广州 510275)

摘要: 在地理信息系统 (GIS) 中, 点、线、面等基本图形要素不仅直接表示了各种地理现象, 而且隐含地表示了地理现象的空间结构, 例如斑块结构、棋盘结构。在许多 GIS 应用中, 需要对图形数据进行概括处理, 以派生出较小比例尺的数据集或地图。由于地理现象的空间结构是地理规律和地理景观格局的重要反映, 因此需要识别数据库中隐含的空间结构, 并在概括后的数据库中有意识地反映这种空间结构。

关 键 词: 空间结构; 模式识别; 地图概括

中图分类号: P283.1 **文献标识码:** A

地理现象在空间分布上往往具有一定的格局或趋势, 例如森林分布的集中与分散趋势, 道路在空间分布上相互交织形成的网络结构。这种空间结构当然也应该反映在 GIS 图形要素中, 否则将难以反映真实世界的实际情况。在许多 GIS 应用中, 需要对图形数据进行概括处理, 以获取更小比例尺的图形信息。在这种情况下, 如何在缩小的图形上反映其原有结构就成了一个突出的问题^[1~3]。图形结构往往是地理现象的特定性质或者特定地理规律的反映。在地图概括中应该突出图形结构特征, 避免图形结构的歪曲。例如棋盘格状街道网概括时要多选取一些近于垂直的街道, 突出其矩形街区的特点。

1 空间结构的基本模式

地图图形的空间结构实质上是地理现象空间分布特点的反映。一般而言, 地理事物的空间分布是不均匀的, 表现出一定的特点与规律。这种规律反映在地图图面上, 就形成一定的结构模式, 例如斑状结构、树枝状结构。

各类地理现象在空间分布上具有不同的特点, 表现为点状分布、线状分布、离散面状分布和连续面状分布等空间分布类型。在此基础上, 还可以进一步区分其结构模式。这里主要讨论离散分布地理现象的结构模式。

其中, 点要素的离散分布研究较多, 一般可以区分为集群型、均匀型和随机型分布。均匀型的分布, 每个点与其它各个最邻近点的距离大致相等。均匀型分布的典型结构模式是规则格网结构、正六边形结构。集群型的分布, 则有一组或一组以上的点群, 每个点与其最邻近各点的距离很小, 而另外的很大区域上则没有点。集群型分布的典型结构模式是团块状结构, 例如我国北方农村居民地多呈团块状分布。随机型分布, 其中有些点比较集

收稿日期: 2001-05-20; 修订日期: 2001-09-23

作者简介: 张青年 (1968 -), 男, 湖北京山人, 博士, 讲师。现从事自动制图、网络地理信息系统等方面的研究与教学工作, 发表论文近 20 篇。

中，有些点比较分散。可以认为随机型分布是一种不规则结构模式，例如我国南方农村居民地的分散分布。

类似地，离散面状分布也可以区分为集群型、均匀型和随机型分布，相应地形成斑块状、棋盘状或不规则型结构。其中集群型分布较为常见，例如，湖泊、沙丘等面状地物都有成群分布的特点。

线状分布原则上也可区分为随机型、集群型和均匀型。但由于线状要素具有延展性，这种区分并没有多大意义。我们根据线状分布的形态结构特点，区分出一些常见的结构模式，例如孤立型分布、成组型分布、树状分布（例如河系树等）、网络型分布（例如道路网）和嵌套型分布（例如等值线）。其中树状分布不存在回路，网络型分布则有回路。例如，水系通常为树状结构，可进一步区分为树枝状、向心状、离心状、羽毛状、格状等；道路通常为网络型分布，又可以进一步细分为矩形、放射状、不规则形等。

2 空间结构的识别与描述

2.1 空间结构的自动识别

在 GIS 中，地理现象被编码为地理数据，各种地理现象通常是用点、线、面等基本要素来表示的。在由地理实体转换到数字模型的过程中，空间结构信息通常没有显式地表示出来。例如，河系通常表示为河段的集合，河流实体由河段的子集合表示，河流之间主流支关系由河段之间的相交关系间接反映，通过弧段 - 结点等拓扑关系隐含在数据库中。在应用 GIS 数据库进行地图概括时，往往涉及到对河系、建筑物群等空间结构单元进行操作。在空间结构单元隐含表示的情况下是无法完成这类操作的。这就需要识别数据库中隐含的空间结构，并将其显式地表示出来。

借鉴视觉认知中的模式匹配观点，我们将地理数据库中地理实体/空间结构单元的识别也理解为一个模式匹配过程。在人的视觉认知过程中，感知到的环境刺激形成视觉影像 (retinal image)，经过与记忆库中的模型匹配，进而识别感知到的物体。类似地，在 GIS 环境下的地理实体识别过程中，存在地理现实、数字模型、理论模型三个对象^[4]。其中，数字模型是对地理现实数字化编码的结果，理论模型是从地理现实抽象而来的简化模型。地理实体及空间结构单元的识别实际上是数字模型与理论模型匹配的过程，见图 1。如果匹配成功，就从地理数据库中识别出相应的地理实体或者空间结构单元。例如，从点状分布的居民地数据库中，识别出团聚状分布单元、散列式分布单元等。我们不考虑地理现实

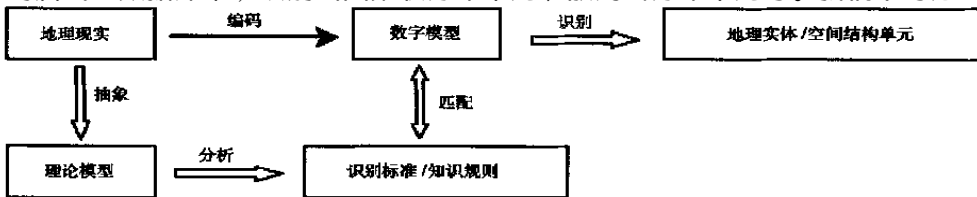


图 1 地理实体/空间结构单元的识别过程

Fig. 1 Identification of geographical entities and/or spatial structural units

到数字模型的转换问题，将这个认知过程简单地划分为分析与识别两个阶段。分析阶段形成理论模型及其识别标准。通过对地理现象的抽象，形成地理现象的理论模型，或者称为理论地理实体，它是一种理想化的地理模型。例如，将棋盘状街道网抽象为规则格网，将岛链抽象为大小相似、排列间距相似的一串图斑。理论模型虽然不存在于现实世界，但高度概括了现实世界中地理现象的共性。因此，建立理论模型是识别地理实体的重要基础。例如，在岛链模型的基础上，可以识别数据库中各个岛链单元。

不过，直接使用理论模型存在许多困难，必须将其分解为若干准则与数字模型匹配。例如，对于岛链模型，应该确定岛屿面积的相似度、排列间距的相似度、排列方向的相似度等准则。根据这些准则，将数据库中符合这些标准的图斑串识别为岛链。

识别阶段是将地理实体的识别标准及知识规则与数字模型匹配，在匹配成功的情况下识别出相应的地理实体。这个识别过程可能由人工完成，也可以编制程序代码由计算机自动完成并辅以必要的后处理。应该指出，不同地理学家对同一地理现象的理解不尽相同，因此可能形成不同的理论模型和识别标准。在不同的识别标准和知识规则引导之下，同样的地理数据可能被识别为不同的地理实体。

2.2 空间结构的参数描述

前已提及，地理实体抽象形成的理论模型通常不是直接与数字模型匹配。因此，必须在建立理论模型的基础上，通过对理论模型的分析，建立地理实体的识别标准和知识规则，作为地理实体识别的依据。对于地图概括而言，主要是使用各种参数对空间结构单元的图形特征进行描述。

空间结构的图形特征可以从 5 个方面进行分析 (Pinchemel 1995)^[4]：

- 极性 (polarity)，向心性 (centrality)，集中性 (concentricity)，外围 (periphery)
- 线性 (linearity)，同轴度 (axiality)，发散性 (radiality)，偏侧性 (laterality)，曲折性 (angularity)，垂直 (perpendicularity)
- 直角形 (orthogonality)，三角形 (triangularity)，等边性 (equilaterality)
- 内部 (interiority)，外部 (exteriority)
- 对 称 (symmetry)，不 对 称 (dissymmetry)，各 向 异 性 (anisotropy)

这 5 组参数从不同的侧面反映了空间结构形态的特点。例如，岛链主要呈线性分布；城市外围居民地分布较集中，具有集中性；梳状水系河流的分布具有偏侧性，主要分布在河流的某一侧。

在实际应用中总是考虑具体情况，从不同侧面选择若干个参数。例如，对于线性分布的岛屿，使用图斑面积、间距、排列方向等参数描述岛链模型，将面积相似、间距相似、排列方向一致的一组岛屿识别为岛链，见图 2。间距、方向等参数的计算可以采用多种辅助数据模型，例如 Delaunay 三角网^[5]。



图 2 规则排列的岛链 (灰色图斑)

Fig. 2 Regular island chain (grayed patches)

3 基于空间结构分析的地图概括

地图概括中最重要的反映地理现象的相互联系与分布规律。否则,只注重局部形态的概括,而忽视全局性规律的反映,就会出现“只见树木,不见森林”的窘境。空间结构是反映地理规律的重要内容。与单个地理现象相比,空间结构更好地反映了全局性的信息。正是基于这一认识,作者认为地图概括必须考虑空间结构的特点和规律,在地图概括过程中有意识地反映研究区域内空间结构的特点与规律。这就要求在空间结构识别与分析的基础上进行概括处理。因此,首先要识别各种空间结构单元及其分布特点,例如点群、面群、树状结构单元及其空间分布差异;然后在地图概括过程中采取一定方法尽量保持这些空间结构单元及其分布特点。这里主要探讨概括过程中的一些处理方法。下面按点、线、面三种分布类型来探讨基于空间结构分析的地图概括方法。

3.1 点状分布的概括

点状分布常见的是集群型,例如沙丘、温泉等往往成群分布。对于点状分布,我们只探讨点群的概括。在全局层次上,通过空间聚类等方法分析点群的空间分布规律,并在概括后的地图上保持地物成群分布的特点,主要是保持点群集中与分散的趋势,以及不同区域点群密度的对比。在点群层次上,将每个点群作为一个整体来考虑,识别点群的分布范围与内部结构,并在概括后的地图上保持点群的原有分布范围及内部结构特点。

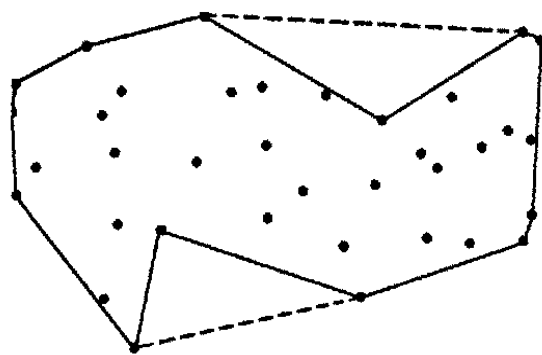


图3 不规则点群的分布范围确定

Fig. 3 Boundary polygon of irregular point cluster

点群层次的概括包括点群分布范围、内部结构的概括。点群有团块状、星形、不规则形等不同形态,应根据其具体特点采用相应方法确定其分布范围。团块状点群的分布范围是一个凸多边形,可用点群的凸壳逼近该边界多边形。而对于星形、不规则形的点群,其边界多边形存在较大幅度的锯齿状弯曲,因此不适合于用凸多边形来逼近,而应该按照锯齿状轮廓线来逼近。我们在点群凸壳的基础上调整相关顶点而得到锯齿状轮廓线。方法是,根据视觉分辨率(0.2mm)和概括前后的比例

尺比值计算出一个阈值,作为锯齿状轮廓线顶点偏离凸壳边界的距离阈值。根据该阈值选取凸壳内部的点替代凸壳顶点,它们之间的三角形内不存在其它点,于是得到锯齿状弯曲的轮廓线,见图3。在确定点群分布范围的基础上,再对点群的轮廓界限进行概括。我们按间距取舍边界点,实现边界点的稀疏化。

点群概括的另一方面是内部点的取舍,这可能要考虑点群的内部结构。某些情况下,点群内部存在明显的结构线。例如城市周围的居民地沿放射状交通线分布时可形成星形结构,城市边沿建筑物沿环形街道网分布时可形成同心环状结构,沙丘可能沿盛行风向呈平行条带状分布。星形结构线、同心环状结构线、平行条带结构线都是一些相对规则而容易引起注意的点群内部结构形式,应该以适当方式保持在概括后的图面上。这里涉及的问题

是结构线的识别和结构线上点的取舍。结构线上的点分布较密集，而周围其它点分布较稀疏，因此结构线在视觉上较突出。正因为如此，我们以最小间距原则为基础构造结构线。即按最小支撑树（MST）构造方法^[6]，将各个点与其最近的点连接起来，然后按间距阈值分割 MST。由于结构线上点的间距较小，因此它们能被分割到同一棵子树上，见图 4。在子树分割的基础上，再对子树上的点按间距取舍。这一方法可实现结构线上点的稀疏化，以及结构线之外的点的取舍。

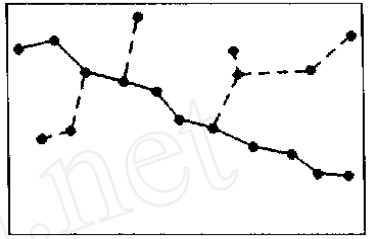


图 4 MST 与结构线

Fig. 4 MST and structural line

3.2 线状分布的概括

线状分布较常见的是树状分布和网络型分布，河系的树状结构和道路的网状结构是最常见的例子。此外，线状分布还有一些其它结构模式，例如成组分布的断裂带、嵌套型分布的等值线等等。这里只讨论树状分布与网络型分布的概括。

3.2.1 树状分布 河流、谷地线、山脊线等地理现象通常呈树状分布。树状结构的概括应反映空间密度差异，保持树状结构的主干骨架线，保持其局部结构模式特点。以河系为例，空间密度差异通过分区概括来控制，主干骨架线选取通过河流等级体系分析来确定，格状、羽毛状等局部结构模式的概括以模式识别为依据。

河流、谷地线、山脊线等各种树状结构都可以建立其等级体系，确定各河流（谷地线、山脊线）在树状结构中的相对地位。以树状结构的等级体系为依据，有利于在概括过程中保留主流（主谷、主脊）舍去支流（支谷、支脊），从而保持其主干骨架线。以河流为例，河流等级体系的建立宜采用 Horton 规则，该等级体系中的河流是从源头到河口的一个完整地理实体。根据地理数据库中河段之间的连接关系，可以确定河流之间的父子关系，进而自动建立河流等级体系^[7]。不过，概括过程中往往还要考虑其它因素，如河流长度、汇流面积等。河流取舍的依据是河流长度、集水面积、河流等级的综合指数，记为：

$$weight = f(length, catchment - area, grade)$$

在特定条件下，树状结构可形成一些特殊的模式，即特化的树状结构，例如平行状、格状、羽毛状、梳状、向心状等结构模式。这些结构模式都是特定的地质构造、岩性、地形环境条件的产物，具有特定的地理指示意义。在概括时应该适当多选取一些能反映这些结构模式的分支。考虑到特殊树状结构模式，河流取舍的综合指数记为：

$$weight = f(length, catchment - area, grade, pattern)$$

根据经验知识，可以建立起各种河系结构模式的原型，实现其自动识别。例如，羽毛状水系的明显特点是，近似平行的若干条支流从河流两侧汇入同一主流，主流两侧支流分布大致均衡（图 5）。根据同侧相邻支流延伸方向的一致性，就能自动识别一些羽毛

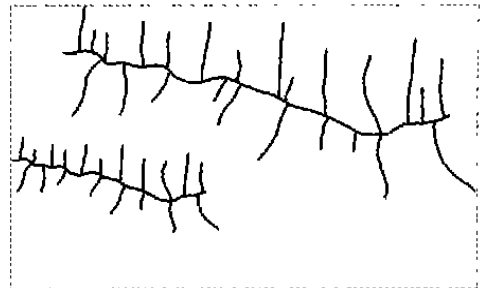


图 5 概括前后的羽毛状水系

Fig. 5 Generalization of feather-like drainage

状水系。当然, 这种自动识别是不完全的, 需要进行一些调整。在结构模式识别的基础上, 将能够进一步区分河流的相对重要性。例如, 如果一条支流是一个羽毛状水系中的一根“羽毛”, 它的重要性系数应该适当提高(乘以一个大于1的系数), 以增加它被保留的机会, 这有利于保持羽毛状水系的特点。

3.2.2 网状分布 网状结构最常见的是道路网。道路具有明确的语义信息, 例如高速公路、国道、省道、县级公路、乡镇公路, 通常用赋权图来表示。对于道路网而言, 语义信息是其取舍的首要依据, 例如高速公路优先于国道被选取。但这不是绝对的, 例如, 连接两个子网的道路, 即使其等级较低, 其相对重要性仍较高。这就要求对道路网结构进行分析, 已有的分析方法包括节点强度计算(祝国瑞, 徐肇忠, 1990)、最小成本连通树分析(Beard & Mackaness 1993)、最短路径连通树分析(Thomson & Richardson 1995)^[8]等。这些分析有利于确定道路交汇点、道路段在网络结构中的相对重要性, 为道路取舍提供依据。

这里考虑的是另一个问题, 即道路网图形结构的保持。街道网、公路网有时表现出非常突出的结构模式, 因此应该反映在概括后的图面上。街道网有矩形、放射状、不规则形和混合型。其中, 矩形及混合型街道网在北方城市较常见, 西安、北京等城市较典型, 广州等城市只表现在局部。公路网也有矩形、放射状、不规则形和混合型。例如, 在成都平原上, 高级道路以灌县为中心呈扇形向外辐射; 在平行岭谷地区, 主要道路沿谷地穿行, 次要道路翻越山岭, 多交织成四边形网状。

图形结构要素的识别是道路网图形结构概括的基础。例如, 矩形道路网的构成要素是两组相互垂直的街道, 这可以通过道路延伸方向分析来识别; 放射状道路网的构成要素是中心点、发散于中心点的一组放射状道路及围绕中心点的另一组环状道路, 可以依据网络节点(道路交汇点)的度数来识别中心点。在此基础上, 较多地选取道路网图形结构要素, 将有利于保持其平面图形的特点。对于矩形街道网, 应注意选取相互垂直的两组道路, 舍去其它影响矩形结构形式的道路; 对于呈放射状的道路网, 除选取发散于一点的放射状道路外, 还应选取围绕发散点的另一组环状道路, 共同构成梯形道路网; 对于混合型的道路网, 则应根据组合的图形, 按保持各自特征的原则选取道路。

3.3 面状分布的概括

面状分布有集群型、均匀型和随机型。其中, 集群型离散面状分布往往构成突出的地理景观, 在概括后的图面上应该保持地物成群分布的特点。面状地物群的识别和概括方法与点群相似。识别时可采用空间聚类、最小连通树层次分割等方法^[6], 概括时则要进行边界要素及内部结构的处理^[9]。但由于面目标有面积和形状, 因此要考虑其影响, 不能简单地借用点群概括方法。下面简单探讨面状地物群概括方法。

3.3.1 面状地物群边界图斑的取舍 面群边界图斑的取舍与点群类似, 但是要考虑图斑面积的影响。也就是说, 比例尺缩小后, 原来的某些边界图斑可能由于面积过小而无法表示, 因此不能再作为边界图斑。我们的做法是, 先按一定阈值舍去面积过小的边界图斑; 然后在剩余的图斑中识别边界图斑。取舍图斑时采用的面积阈值略小于正常阈值, 这是考虑到边界图斑表示了面群范围而具有相对重要性, 因此允许它们略小于正常面积阈值。识别边界图斑后, 生成凸壳边界并按面群形状进行调整, 得到新的面群边界。然后沿边界线按间距取舍边界图斑, 并将略小的边界图斑夸大到面积阈值。

3.3.2 面状地物群内部结构的概括 面群内部结构的概括包括两方面,即面群骨架和内部结构线。面群骨架是面积相对较大图斑,选取面积阈值以上的图斑就能保持面群骨架。面积阈值的确定受视觉分辨力的影响,同时也要考虑地物的性质及区域特征。例如我国北方耕地连片分布且面积较大,可取较大阈值;而南方耕地分布分散且地块小,应取较小阈值。

面群内部的结构线则是一个规则排列的图斑序列,例如河口三角洲外围滩地的扇形分布、鸡足状分布等。在这类结构线上,图斑排列规则有序,是地图图面上的鲜明特征,应尽可能表现在概括后的地图上。当结构线上的图斑过于紧密或图斑过小,低于最小可分辨距离、最小表示面积等标准时,必须作为一个整体进行取舍、合并、移位、夸大等概括处理。内部结构线概括后应尽量保持其排列方向、间距及空间分布范围。

对于图斑间距较小的紧密结构线单元,应作取舍或合并处理。湖泊等要素只能取舍,植被土壤图斑则可合并。取舍或合并图斑应按一定间距进行,即每隔若干个图斑舍去一个图斑,或每两个图斑合并为一个图斑,以保持线状结构单元的分布范围和内部间距。对于由多个小图斑构成的结构线单元,必要时可以适当夸大各个图斑,以便保留在概括后的图面上。

4 结论

各种地理现象的空间分布都表现出一定的结构或规律,例如斑块结构、条带状分布等。当地理世界转换为地理信息系统(GIS)中的数字世界时,这些空间结构信息没有显式地描述出来。为了在缩小的地图上反映这些结构信息,就需要识别数据库中隐含的空间结构,并在概括后的数据库中有意识地反映这种空间结构。点、线、面等不同空间分布类型具有不同的空间结构模式,在地图概括中应采取不同的处理方法。目前需要积极引进计算机视觉、计算几何、图论等各种分析方法,以识别和处理各种空间结构。空间结构的识别与处理增加了计算工作量,但这将有利于在概括后的地图上保持研究区域内各种空间结构的特点和规律。

参考文献:

- [1] 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.
- [2] 陈述彭, 岳天祥, 励惠国. 地学信息图谱研究及其应用[J]. 地理研究, 2000, 19(4): 337~343.
- [3] 齐清文, 刘 岳. 非连续分布面状地理现象的图形自动概括方法[J]. 地理研究, 1996, 15(1): 1~8.
- [4] Boffet A. A framework for automated spatial analysis based on spatialization principles[EB/OL]. <http://www.geo.unizh.ch/ICA/Documents/Workshop99/papers.html>, 1999.
- [5] Wanning P. Automatic generalization in GIS[M]. The Netherlands: ITC Publication Series, 1997.
- [6] Regnault N. Recognition of building cluster for generalization[A]. In: Proceedings of the 7th International Symposium on Spatial Data Handling[C]. 1996. 185~198.
- [7] 毋河海. 河系树结构的自动建立[J]. 武汉测绘科技大学学报, 1995, 20(增刊): 7~14.
- [8] Thomson R C, Richardson D E. A graph theory approach to road network generalization[A]. In: Proceedings of the 17th ICA[C]. Barcelona: 1995. 1871~1880.
- [9] 张青年, 秦建新. 面状分布地物群识别与概括的数学形态学方法[J]. 地理研究, 2000, 19(1): 93~100.

On the identification and generalization of spatial structure in GIS

ZHANG Qing-nian

(Department of Urban and Resources Planning, Zhongshan University, Guangzhou 510275, China)

Abstract : Spatial structural patterns are very important reformation inflecting regional characteristics. However, they are seldom explicitly expressed in geo-data bases. In order to express them in generalized maps derived from geo-data set, spatial structural units need to be identified beforehand. This involves in the classification, identification, description and generalization of spatial structural patterns.

This paper includes three sections. The first two parts discuss the classification and identification of spatial structure patterns. Spatial distribution can be subdivided into three categories, i. e., point, linear, and areal phenomena, which can be further divided into uniform, agglomerate, or random point/areal patterns, and isolated, grouped, nested, dendriform, or netlike linear patterns. Spatial structural units can be identified by means of comparison between digital model and theoretic model, which are derived from geographical entity. Theoretic models of structural patterns are usually described by a group of parameters, such as polarity, linearity, and symmetry.

The third section studies the method to generalize spatial structure of point, linear, and areal phenomena. Among point and areal distribution, clusters are the most important structural units. Clusters can be generalized by means of boundary adjustment and inner structure simplification. As for linear distribution, dendriform and netlike patterns are the most important units. The generalization of both of them involves in trunk lines selection and typification of characteristic patterns, which need to be dealt with different methods according to concrete situation.

Key words : spatial structure; structure identification; map generalization