

逐层分解选取指标的河系简化方法

张青年

(中山大学地理科学与规划学院, 广州 510275)

摘要: 地图综合理论与方法的研究重点之一, 是依据上下文环境进行结构化选取和化简。水系作为地图的基本要素之一, 在地图综合中必须依据水系的类型和河网密度进行化简, 以保持河网密度差异等宏观特征。本文提出了一种基于支流数量差异的选取指标分配与河系简化方法, 将选取数量在河流的各个子流域上按比例分配并递归分配到下一层次的子流域, 在各个子流域依据长度和间距选取河流, 从而实现顾及密度差异的结构化选取。在软件开发基础上进行了选取试验, 并对照手工综合的河系图进行了分析。结果表明, 基于支流数量差异的比例选取方法切实可行, 能够有效地保持树状河系的特点和河网密度的区域差异。

关键词: 河系简化; 层次分解; 密度差异; 地图综合

文章编号: 1000-0585(2007)02-0222-07

1 引言

近年来地图综合理论与方法的研究取得了许多进展, 极大地提高了地图综合处理的自动化程度。地图综合研究由早期的曲线简化等简单处理深入到地物群化简等复杂问题^[1~7], 依据上下文(context)环境进行结构化选取和化简成为地图综合的研究热点。例如, Ruas 将城市空间划分为若干个城市区块(urban situation), 对同一个城市区块内各个建筑物的取舍和移位操作进行协调和控制^[3,4]。Regnauld 用最小生成树(MST)分割建筑物群, 在建筑物选取时考虑整个建筑物群分布特征的保持^[5]。Sester 用人工神经网络自动调整建筑物的选取, 直到选取的建筑物分布与原建筑物群分布一致^[6]。艾廷华等在 Delaunay 三角网及其对偶 Voronoi 图模型上定义了分布范围、分布密度、分布中心及分布轴线等描述参量, 实现了保持空间分布特征的群点化简^[7]。结合上下文的分析能够有效地顾及背景图形, 协调地图综合中图形之间的矛盾。但是, 这类研究目前局限于大比例城市地图(特别是建筑物)的地图综合, 处理过程非常复杂。直到现在, 基于上下文的地图综合尚未引起足够重视, 许多研究仍局限于选取数量的开方根规律等基本模型的应用。

水系作为地图的基本要素之一, 是地图综合中不可避免的要害, 已有较多研究成果^[8~12]。确定河流选取指标的模型主要有一元回归模型、多元回归模型、开方根规律模型等。一元回归模型依据河网密度确定河流选取指标(河流长度)。多元回归模型不仅考虑单位面积内河流的长度(河网密度), 而且兼顾单位面积内河流的条数, 依据河长与条数确定河流的选取标准(河流条数)。方根模型则依据方根规律确定新比例尺下的河流选取指标

收稿日期: 2006-08-08; 修订日期: 2006-11-18

基金项目: 国家自然科学基金项目(40671164、40101024); “985”工程 GIS 与遥感的地质应用科技创新平台(105203200400006)

作者简介: 张青年(1968), 男, 湖北京山人, 博士, 副教授。主要从事地图学理论方法与地理信息系统应用研究。E-mail: zqznus@163.com

(河流条数)。这些模型有利于确定选取资格(河长)或定额(数量), 但无法根据区域特点进行差异化选取。若河网密度存在区域差异, 则需事先进行密度分区, 再分区计算指标并选取河流。河流结构选取模型主要有等比数列法、模糊数学模型等。等比数列法建立河长和间隔两个等比数列, 依据河流的大小和河流的密度两个因素来衡量河流的重要性, 对不同长度等级的河流按不同的间隔进行取舍。模糊数学模型考虑河流的长度、密度、相对重要性和河网类型等因素, 建立模糊综合评判矩阵进行河流选取。结构选取模型能够确定哪些河流应该选取, 但涉及河流间隔等多种指标的计算, 处理过程复杂, 难以自动化实现。因此, 有必要发展新的、计算简便、又能兼顾河流的上下文环境的河系简化方法。

笔者在树状水系结构分析的基础上曾提出一种结合河流长度、所在层次、以及支流数的综合选取指数, 依据该综合指数将河流排序并按先后顺序选取, 实现顾及密度差异的水系简化^[13]。本文采用一种新的方法, 基于河系的层次分解和各个层次上不同子流域之间选取数量的比例分配进行河系简化, 以期更有效地保持河流的密度差异。本文首先分析了河系层次结构与区域密度差异, 认为河系内部的密度差异体现为同一层次的不同子流域的河流密度差异, 据此提出了在同一层次的各个子流域间分解选取数量指标的河系简化方法。将河流选取数量指标逐层分解到它的各级子流域, 从而确定各个层次、各个子流域的河流选取数量, 并结合长度、间距指标确定具体选取哪些河流。最后使用该方法进行了案例研究, 验证了该方法的有效性。

2 河系层次结构与区域密度差异

大多数河系都具有树状结构, 河系内部存在密度差异。因此, 结合河系的层次结构, 分析其主干河流及密度差异, 是确定河流选取指标的重要依据。河网的形态类型等因素也是取舍河流的重要依据, 但不易量化处理, 本文暂不考虑。

2.1 河系层次结构

树枝状、格状、羽毛状等树状河系是最主要的河系类型。一些河系虽然局部存在辫状支流, 但仍可表示为树状结构。树状河系具有层次结构, 表现为主流-支流-次级支流之间的父子关系。在空间数据库支持下, 能够根据河段之间的连接关系自动构建河系树, 从而自动建立河流之间的层次结构。

在河系树结构中, 河流所在的层次越高, 其相对重要性越大, 在地图综合中被优先选取。事实上, 在手工地图综合中, 河系简化是在选取主干河流的基础上逐渐加密河流实现的^[8]。但是河流所在层次并不能作为河流选取的主要依据, 即便是一级支流也可能由于长度不够而被舍去。

在河系树的同一层次上, 各条河流具有不同的长度, 据此可确定其相对重要性。河流长度是选取河流的基本依据之一。一元回归模型、多元回归模型、等比数列模型等都考虑了河流长度指标。以河流长度为分界尺度选取河流是水系综合的基本方法。例如, 把 1 cm 的长度作为河流的选取标准, 地图上大于 1 cm 的河流即可选取, 这个标准以下的则一般应舍去。但是, 河流长度不能全面衡量河流的重要性, 还需考虑河流所在区域的河网密度、河网类型。长度相似的两条河流在不同的地理环境中其重要程度可能差之甚远。

2.2 区域密度差异

河网密度存在区域差异, 保持河流密度差异是河系简化基本目标之一。在手工作业中, 通常对河流进行密度分区, 在各密度区采用不同的选取标准。人为划定的密度区界线

有一定的随意性。各密度区的选取标准通常不是一个固定值, 而是一个范围值, 需要根据河系类型和密度级别灵活运用。河流间隔也可反映河网密度, 是河流选取的一个辅助指标。例如按等比数列法确定特定长度和间隔的河流的选取间隔。在资料缩小图上间隔小的河流的选取间隔通常增大较多, 而资料图上间隔大的河流的选取间隔增加较少。然而, 在自动制图环境下, 密度分区和河流间距指标的计算较难实施。例如, Delaunay 三角网可用于计算河流的相邻点对之间的距离^[11], 但计算量大。相邻两条河流之间的间距主要取决于两条河流在若干关键点上的距离, 但难以自动判断采用哪些点位计算河流间距。

本文不划分河网密度区, 直接以河系为单元计算河网密度差异。不同河系之间存在河流密度差异, 同一河系内部不同子流域之间同样存在密度差异(图1)。在自动制图环境下, 河系中的河流数量和总长度易于计算。直接以河系为单元, 比较各个河系中河流数量的差别(除以流域面积), 就能确定不同河系的河流密度差异; 类似地, 比较同一河系中各个子流域的河流数量差别(除以子流域面积), 就能确定同一河系内部的河流密度差异。在缩编地图的过程中, 通常只需按比例分配河流选取数量。显然, 依据资料图上的各个河系的河流数量比例将总的选取指标(河流条数)分配到各个河系, 就能保持新编图上各个河系的密度对比; 依据资料图上某个河系内部各个子流域的河流数量比例将该河系的选取指标(河流条数)分配到各个子流域, 就能保持新编图上各个子流域的密度对比; 各个子流域又存在次一级的子流域, 按同样道理可以进行递归分配。树状河系存在典型的层次结构, 将河流选取数量进行层次分解, 逐层分配到各级支流, 将能有效地保持河流的区域密度差异。

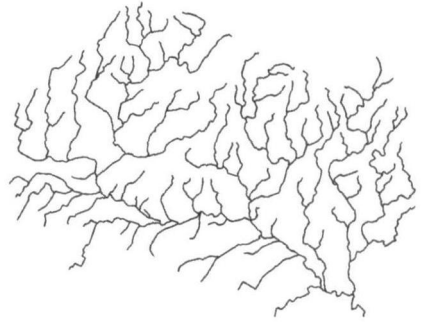


图1 河系内部不同子流域之间的河流密度差异

Fig 1 Density differences between the sub-drainages of a river

3 基于层次分解的河系简化

主要有三个步骤: 首先构建实体, 建立河系树; 然后确定选取数量; 最后将河流选取数量分配到河系的各个子流域, 并依据河系的层次结构逐层分解选取数量指标; 在同一层次上, 依据河流长度和间隔选取河流, 实现河流的结构化选取。

3.1 构建河系树

在空间数据库中, 河流通常是以河段为单元存储的, 河流被表示为一个相互邻接的河段的序列。这意味着首先要在河段的基础上建立河流实体。然后根据河段之间的父子关系可以建立河流的层次结构, 建立河系树^[14, 15]。本文采用文献[15]的方法直接在河段的基础上以顺着河源向下游追踪的方式构建河流实体和河系树。

构建河流实体和河系树的基本原则是从河源向下游追踪, 在多条河流的汇合处, 将长度较大的上游段作为河流的正源与下游河段合并, 使河流实体不断向下游伸展, 直到它不再作为河流正源为止(图2)。由于历史原因, 某些河流并不以最长的河源作为正源, 构建河流实体时还应考虑河流名称。如果河流数据含有河流名信息, 应将同名河段合并到同一个河流实体中。在追踪过程中不断更新河流的等级值(支流数), 见图2。

在构建河流实体的同时也识别了河流之间的主支流关系, 于是就逐步建立起河系树。

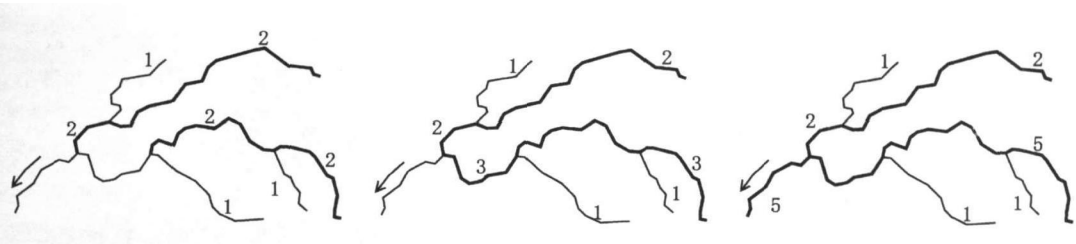


图 2 从河源向下游追踪构建河流实体

Fig 2 Constructing river entities by tracing them downward from sources

如前所述向下游追踪时，在河流交汇处，作为正源的河流与下游河段合并，该河流实体是主流；而在此汇入但不是正源的河流则是支流。随着合并过程的继续，一些河流不再作为主流与下游河段合并，而是作为更高级别河流的支流。这样，随着追踪过程的继续，最低级别的河流最先停止向下游伸展而到达其终点，继续向下游伸展的河流的级别越来越高，河系树的各个层次逐步形成。

3 2 确定选取数量

选取数量可以用方根模型确定。方根模型是关于新编图与资料图上某类制图物体数量规律的一种模型。该模型认为新编图与资料图上同一类制图物体数量之比与这两个图件的比例尺分母之比正相关，同时还受到符号尺度系数和物体重要性系数的影响。计算河流选

取数量的方根模型可采用下述形式：
$$n_F = n_A \left(\frac{M_A}{M_F} \right)^x \tag{1}$$

式中， n_A 是资料图上的物体数量， n_F 是新编图上的物体数量， M_A 是资料图的比例尺分母， M_F 是新编图的比例尺分母， x 为经验系数，包含了符号尺度系数和物体重要性系数的综合影响。

通过地图综合样图的统计分析可以得到 n_A 和 n_F ，从而确定 x ，例如取 1~ 5 之间的一个值。 x 的取值受河流密度、新编图的制图目的等因素影响。例如，将某幅 1: 150 万的资料图缩编为 1: 400 万的科学参考图时 x 取 2；缩编为 1: 400 万的教学参考图时 x 取 3^[18]。

3 3 逐层选取河流

具体包括四个步骤：将选取数量分配到各个河系；在各个河系内部将选取数量分配到各个子流域；在子流域内部递归分配选取数量；在同一个层次上依据河流长度和河间距选取河流。下面分为数量分配和河流选取两个方面进行说明。

(1) 选取数量的分配

首先将选取指标分配到各个河系。各个河系分配到的数量按下述公式计算：

$$n_{Fi} = n_F \times \frac{n_{Ai}}{n_A} \tag{2}$$

式中， n_A 为河流总条数， n_F 为总的选取数量， n_{Ai} 为第 i 个河系的河流数量， n_{Fi} 为第 i 个河系的选取数量，选取数量按四舍五入处理为整数。

按公式 (2) 计算，支流少的小河系的选取数量可能是 0；如果存在许多小河系，计算得到的选取数量之和将小于应选取的总数 n_F ，假定差值为 Δn_F 。将选取数量为 0 的小河系按其主流长度排序，将前 Δn_F 个小河系的选取数量更新为 1。

然后在河系内部分配选取指标到各个子流域，并将选取指标逐层分解到次一级子流域。仍按公式 (2) 计算各个子流域的选取数量，但各个符号的含义略有变化。以河系的

一级支流(子流域)为例, n_A 为该河系的河流总数, n_F 为该河系的总选取数, n_{Ai} 为第 i 个子流域的河流数量, n_{Fi} 为第 i 个子流域的选取数量。在河系内部分配选取指标时, 同样会出现计算得到的选取数量不够的问题。河系内部补选数量的分配与河系之间的情况略有不同, 见下一节。

(2) 河流的选取

在上述选取数量分配过程中, 如果某个河系或者子流域的选取数量为 1, 直接选取其主流即可; 如果选取数量大于 1, 先选取其主流, 再进一步递归分解剩余选取指标; 如果选取数量等于 0, 则依据补选结果进行处理。如果某个河系在补选后的选取数量仍为 0, 则直接舍去整个河系, 否则选取其主流。在河系内部, 选取数量为 0 并且主流长度大于阈值的子流域是补选的候选对象, 是否补选其主流还要考虑河间距, 避免补选的河流与其他已选的河流的间距过小。具体方法是, 首先选取那些选取数量大于 1 的子流域的主流, 再计算候选对象的主流与同侧的已选主流的间距, 这些主流是同一河流的同级支流。计算同侧支流的河口之间的直线距离作为二者的河间距, 将河间距小于阈值的子流域从候选对象集中剔除。然后在候选对象集中按河流的支流数和长度的降序优先选取支流较多和长度较大的河流, 一次只能补选一条河流, 补选一条河流之后立即更新其他候选对象的河间距值。也可考虑采取随机方式而不是按长度的降序选取河流, 这样可以使选取的河流随机分布在其上一级主流的沿线, 而不是集中于某些部位, 但随机补选方式下可能有一些明显较长的河流被舍弃。继续在候选对象集中补选河流并更新河间距, 直到达到补选数量为止。

4 河系简化试验

在 Java 环境下, 我们编制了水系综合插件(Plugin), 实现了河系树的自动构建和选取指标的逐层分配。在进行河系简化时, 首先构建河系树并计算各条河流的支流数, 然后采用方根规律计算河流的选取数量, 最后逐个河系、逐个子流域递归分配选取指标并选取子流域的主流直到选取数量达到要求为止。本次试验的部分样图见图 3。

图 3a 与图 3b 的数据取自文献 [8]。其中, 图 3a 是湖北省西部 1:20 万水系图的局部, 共有 68 条河流。图 3b 是手工缩编制作的 1:50 万水系图, 共选取了 27 条河流。与图 3b 对应, 图 3c 是按层次分解选取指标的方法编制的 1:50 万水系图。将公式 (1) 中的指数 x 取 2, 得到选取指标为 27, 与图 3b 的选取额度相同, 补

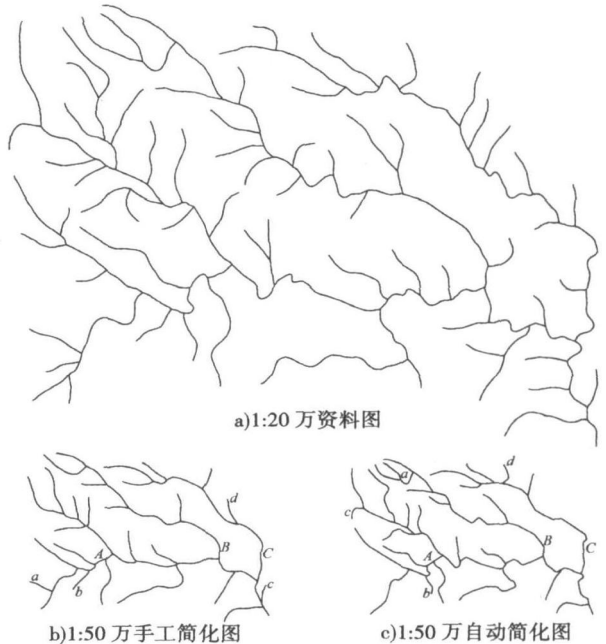


图 3 河流选取试验 (图 3c 是采用本文方法进行综合的结果, 图 3a 与图 3b 引自文献 [8])

Fig 3 A case study of selecting rivers (Generalized map in Fig 3c was created by layered decomposition method; Fig 3a and Fig 3b were copied from literature [8])

选河流时依据河流的支流数和长度的降序优先选取支流较多和长度较大的河流。试验中实际选取的河流数量正是 27 条。由于地图综合结果评价方面的研究尚不成熟^[16], 我们采取目视比较方法评价河流选取结果。可以看出, 图 3c 选取的河流与图 3b 基本一致, 选取额在各个子流域的分配合理, 选取效果较好。具体而言, 子流域 A 在图 3c 中的选取数量比在图 3b 中的数量小 1 (图 3b 中的河流 a 在图 3c 中没有选取), 子流域 B 在图 3c 中的选取数量比在图 3b 中的数量多 1 (图 3c 中的河流 a 在图 3b 中没有选取), 子流域 C 在图 3c 中的选取数量与在图 3b 中的数量持平。总的看来, 河流密度的区域差异在缩编后的图上保持得较好。不过, 位于河系树底层的部分河流的选取不尽合理。具体地说, 图 3c 中的河流 a 和 c 应该舍弃, 而河流 b 和 d 由图 3b 中的河流 b 和 d 取而代之更为合理。这反映了依据河流长度补选河流存在不足之处, 需要进一步考虑河流间距。

5 结论

地图综合理论与方法的研究重点之一, 是依据上下文环境进行结构化选取和化简。结构化选取涉及到结构分析和指标选择等复杂问题。对于水系而言, 需要结合河网的形态类型、河流密度差异等因素进行结构化选取。本文从保持河网密度差异出发, 在计算河流选取总额的基础上, 采取逐层分解选取额的方法, 实行河流的逐层选取。选取指标首先在各个河系之间分配, 然后在各河系的各个子流域之间进行分配。各个河流和子流域分配的选取指标取决于该河系或子流域的总河流数。在 Java 环境下基于层次分解进行了河流选取试验。试验结果表明, 基于层次分解的河流选取方法能够较好地保持河网密度的区域差异, 但位于河系树底层的部分河流的选取不尽合理。因此, 在确定各子流域的选取额之后, 具体选取哪些河流的问题尚待进一步研究。

参考文献:

- [1] Weibel R, Dutton G. Generalising spatial data and dealing with multiple representations. In: Longley P A, Goodchild M F, Maguire D J, et al. *Geographic Information Systems—Principles and Technical Issues* (2nd ed), Volume 1. New York: John Wiley & Sons, 1999. 125~ 155
- [2] Weibel R. Generalization of Spatial Data: Principles and Selected Algorithms. In: van Kreveld M, Nievergelt J, Roos T, et al. *Algorithmic Foundations of Geographic Information Systems*. Berlin: Springer, 1997. 99~ 152
- [3] Ruas A, Mackaness W A. Strategies for urban map generalization. In: *Proceedings of the 18th ICA/ACI International Cartographic Conference*. Stockholm: ICC, 1997. 1387~ 1394
- [4] Ruas A. A method for building displacement in automated map generalization. *International Journal of Geographical Information Science*, 1998, 12(8): 789~ 804
- [5] Regnauld N. Contextual building typification in automated map generalization. *Algorithmica*, 2001, 30: 312~ 333
- [6] Sester M. Optimization Approaches for Generalization. [http://www. geo. unizh. ch/ICA/](http://www.geo.unizh.ch/ICA/)
- [7] 艾廷华, 刘耀林. 保持空间分布特征的群点化简方法. *测绘学报*, 2002, 31(2): 175~ 181
- [8] 何宗宜. *地图数据处理模型的原理与方法*. 武汉: 武汉大学出版社, 2004
- [9] 毋河海. 自动综合的结构化实现. *武汉测绘科技大学学报*, 1996, 21(3): 277~ 284
- [10] 何宗宜, 阮依香, 尹为利, 等. 基于分形理论的水系要素制图综合研究. *武汉大学学报(信息科学版)*, 2002, 27(4): 427~ 431
- [11] 艾自兴, 毋河海, 艾廷华, 等. 河网自动综合中 Delaunay 三角的应用. *地球信息科学*, 2003, (2): 39~ 42
- [12] Thomson R, Brooks R. Exploiting perceptual grouping for map analysis, understanding and generalization: The case of road and river networks. In: *Graphics Recognition: Algorithms and Applications*, Volume 2390 of Lecture Notes in Computer Science. Berlin: Springer, 2002. 141~ 150

- [13] 张青年. 顾及密度差异的河系简化. 测绘学报, 2006, 35(2): 191~ 196
- [14] 毋河海. 河系树结构的自动建立. 武汉测绘科技大学学报, 1995, 20(增刊): 7~ 14
- [15] 张青年, 全洪. 河系树的建立及其应用. 中山大学学报(自然科学版), 2005, 44(6): 101~ 104
- [16] 吴洁, 李霖, 郭仁忠, 等. 数字环境下地图综合结果自动评价模型的研究与应用. 测绘科学, 2002, 27(2): 34~ 39

Drainage generalization by layered division of the number of retained rivers in river trees

ZHANG Qing-nian

(School of Geography, Sun Yat-sen University, Guangzhou 510275, China)

Abstract: The theoretic and methodological research of map generalization is an important field in cartography and geographical information systems. The difficulties of map generalization lie in selectively omitting some features while maintaining the overall characteristics of their spatial distribution. Simplifying drainages is an important aspect of map generalization. Structural patterns and density differences are salient characteristics of drainages, which should be maintained in generalized versions. This paper proposed a new method to generalize dendritic drainages while maintaining the density differences between sub-drainages, which is based on the layered division of the number of retained rivers in river trees. We analyzed the layered structure of dendritic drainages and their density differences, and found the density differences inside a drainage come from the differences of the number of tributaries of its sub-drainages. A method to allocate the number of retained rivers is accordingly proposed, which divides the overall retained number of sub-drainages of a drainage according to the ratio of the number of the tributaries. The structured drainage generalization is executed in three steps. Firstly, the river entities and river trees are constructed, and the number of tributaries for each river is counted. Then the number of rivers to be retained in the generalized version is computed according to the Square Root Law. Lastly the number of retention is divided among all of the drainages according to the number of their tributaries, the mainstream of a drainage with big length and spacing is selected, and the allocation and selection are recursively executed in the lower layer until the number of the selected rivers reaches the limit of the drainage. The method was implemented as plugins in Java Environment. Experimental results are compared with the hand-made generalized maps. A case study of drainage generalization showed that the method created acceptable results and the density differences among drainages were maintained effectively.

Key words: drainage generalization; layered division; density differences; map generalization