

# GIS 中的地理信息综合\*

王 桥

(中国科学院遥感应用研究所 北京 100101)

毛 锋

(中国科学院地理研究所 北京 100101)

吴纪桃

(武汉测绘科技大学 武汉 430070)

**摘 要** GIS中地理信息综合功能的研究和开发是当前GIS技术进步与应用发展所面临的一个新挑战。首先研究了GIS中地理信息综合功能的意义、基本内容及其实现途径,然后介绍了作者所提出的基于分形分析的自动综合算法和实例,最后分析了GIS地理信息综合功能研究中需要进一步解决的问题。

**关键词** 地理信息系统 地理信息综合 分形分析

## 1 引 言

近些年来, GIS技术和应用得到了迅速发展,应用规模不断扩大,需求日益增加,所处理问题也更加复杂化,这使得GIS中地理信息的有效利用和表达问题显得越来越突出。因此,人们不得不开始关注GIS中地理信息综合功能,希望通过这种功能,使花费巨大代价所获取的基本比例尺数据能够派生出多比例尺、多用途和多专题的更多有用信息,以降低数据采集、存储和管理的费用,扩大GIS的服务面;与此同时,提高地理信息的表达和利用的效率,满足GIS的可视化和各种分析、制图输出的需要<sup>[1-4]</sup>。目前,国际上一些学者和研究机构已越来越认识到研究和开发GIS中地理信息综合功能的重要性,并在实际工作中取得了一些阶段性的进展。例如, ZEISSZ在PHOCUS中开发的CHANGE模块已能够完成一些线状和面状要素的自动化简; INTERGRAPH在MGE中开发的MAP GENERALIZER模块初步提供了人机交互的分项综合功能;法国地理院和英国陆军测量局也正取得重要成果,以批处理方式满足了一些地理信息综合的特定需要。然而, GIS中地理信息综合所涉及的问题相当复杂,包括上述工作在内的有关研究实际上才刚刚起步,还远没有达到生产性实用程度。例如,美国的USGS系统只能进行人工编码定级式的

地理信息综合;德国的ATKS系统则采用了系列比例尺重复数字化的方法暂时回避了地理信息综合问题的处理。各种事实表明, GIS中地理信息综合功能的研究和开发是当前GIS技术进步所面临的新挑战,也是GIS应用发展中迫切需要解决的问题,值得进一步予以重视。

## 2 地理信息综合及其在GIS中的意义

### 2.1 GIS中的地理信息综合

GIS中的地理信息综合通常称为自动综合,是依据制图综合原理对地理信息进行化简、概括、合并、取舍、移位等,以获得所需比例尺或分辨率下地理信息的本质的、主要的和整体规律性内容的计算机信息处理过程。一般的说,地理信息综合比通常意义上制图综合的含义要广泛,它除了主要涉及图形信息外,还涉及属性信息、统计信息和有关专题信息。传统上,制图综合被认为是图形信息处理中一种最具主观过程的工作<sup>[1,2]</sup>,其中既有方法、规则的运用,也有各种复杂条件下高度思维的创造性劳动,还包含了大量经验的结晶,因此,难以抽象成形式化的模型,即使有些内容或过程可以用形式化语言表示,也难以建立相应的数学描述模型。显然,这种制图综合的知识、方法和过程的算法化和程序化的困难是造成目前GIS中的地理信息综合进展举步为艰的根本原因。

\* 国家自然科学基金(49571061)及测绘遥感信息工程国家重点实验室开放基金(WKL(95)0301)项目资助。

收稿日期: 1997-06-18; 收到修改稿日期: 1997-08-19

### 2.2 GIS 对地理信息综合的需要

GIS 中地理信息处理过程总要依赖于一定的比

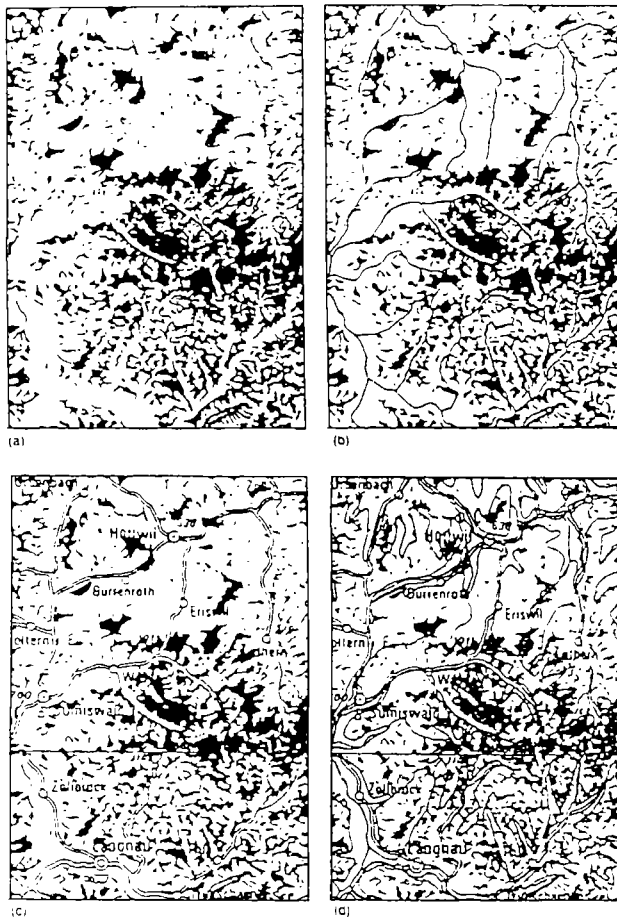


图 1 地理信息综合的必要性举例(据 E. Spiess, 1996)

(a) 单层要素图 (b) 与水系叠置图

(c) 与交通线、居民点叠置图 (d) 多层要素叠置图(经综合)

Fig.1 (a) A single component map without reference data; (b) With reference data; (c) and (d) The more additional components the more generalization becomes a necessary

例尺。农业、交通、地质、规划、管理等方面研究和应用的任务不同,对地理信息比例尺(详细程度)的要求

亦不同,GIS 的开发者和用户总是希望在最容易表达和理解空间过程的尺度下来处理专业问题<sup>[2]</sup>。但是,GIS 中的地理信息是以固定的基本比例尺存储于数据库中的,为每一种所需比例尺的数据都建立数据库显然是不现实的。因此,有必要在 GIS 中开发地理信息综合功能以支持多比例尺操作,即利用基础数据库本身来生成各种所需比例尺或分辨率的数据。这不仅可以节省大量人力、物力、全面降低数据采集、存储、维护、更新的费用,而且可以提高现有数据库的价值和整个 GIS 的效率。

由于纸质地图的普及性、方便性、易读性等特点,各种比例尺的纸质地图仍然是 GIS 所输出数字化产品的主要形式。而在此过程中(如由 GIS,数据库中基本比例尺数据产生较小比例尺地图、产生新的专题地图或多层要素叠置图等如图 1),必然涉及到地理信息综合,只有依据制图综合的原理和方法进行数据处理,才能够正确处理内容结构关系,把握适当的信息负载量,客观体现制图区域内地理现象的本质规律,并满足制图规范及地图印刷等方面的要求。如果不开发专门的地理信息综合功能,一般的图形编辑功能是很难实现这些的。

一般认为,GIS 数据屏幕输出过程中的地理信息综合问题远不如在制图输出过程中那样严重,其实不然。在相当多的情形下,GIS 用户和开发者都需要在 GIS 固定大小的窗口内显示或操作带有全局性的数据,如在大范围内显示和研究区域要素的分布、密度、空间关系,在大范围内观察和比较一定专题的整体性内容和特征等等。尽管 GIS 已提供了比较完善的屏幕缩放功能(Zoom),可以方便地进行比例尺变化,但由于受计算机屏幕分辨率等方面的限制,当比例尺或分辨率变小时可视化程度也随之降低,数据显示和操作效果很难令人满意(图 2)。因此,从 GIS 可视化的需要看,研究和开发地理信息综合功能也是很有必要的。

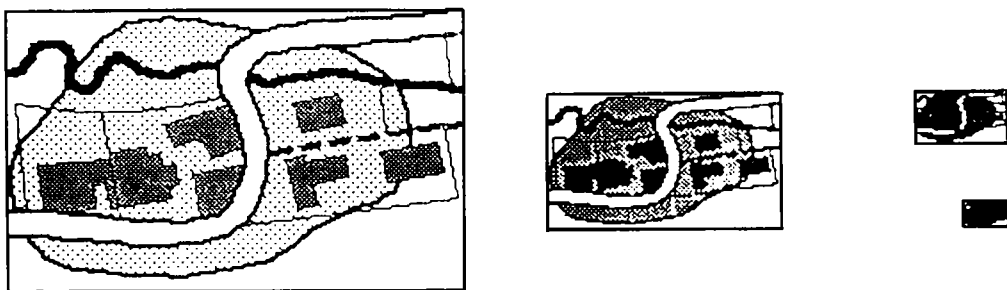


图 2 Zoom 功能对可视化的影响举例

Fig.2 A example of the visualization is affected by Zoom function

事实上, GIS 对地理信息综合功能的需要还表现在数据更新、数据质量和精度控制等诸多方面, 另外, 地理信息综合还可以去除图像噪声(这些噪声不仅包括几何方面, 而且也包括属性、概念及其它多余信息), 突出其基本特征和内容, 这无疑有助于对 GIS 中的图形图像进行各种分析和操作。正如 Mark (1991)所指出的, 综合不仅对图形显示具有重要意义, 对于空间分析也具有重要的意义。

### 3 实现地理信息综合功能的基本途径

#### 3.1 模型综合

空间物体可以有多种描述形式, 如图形方式、模型方式等等。所以人们认为, 相对于面向图形的综合(制图综合)也需要有面向模型的综合(模型综合)。模型综合主要体现在以下方面: 一是进行空间数据压缩, 以节省存储空间、提高执行和表达效率; 二是对数据集的精度或分辨率压缩, 以保证一定比例尺或一定用途下数据集的一致性(如把时间序列集的单位由月转换为年); 三是复杂目标的重构和数据合并, 以缩减有关专题的解空间; 四是从一种比例尺数据模型派生多种数据模型(如多精度、多分辨率、多比例尺等数据模型)。模型综合可视为制图综合的预处理, 它基本上是一种模型驱动的确定性行为, 其结果往往是坐标被过滤到图形所需要的分辨率。

针对模型综合的需要, 现已发展了一些支持随机比例尺下数据插入、删除、更新、查寻的方法<sup>[2,5]</sup>, 还提出了无缝存储结构、无比例尺地理数据库的概念, 包括四叉树、样条树等在内的数据结构也常被用来进行数据合并和分割; 此外, 模型综合过程中实体和关系的描述与处理一般还采用了 GIS 中的拓扑空间和结构空间方法。需要注意的是, 现有模型综合方法大都是面向离散地理现象的, 所以应发展适合于处理在时间或空间上连续的地理现象的方法。

#### 3.2 制图综合

制图综合是面向图形的地理信息综合, 在 GIS 环境下, 它是一种利用计算机去完成人工制图综合的基本操作的信息自动处理过程。当然, 无论是方法、规则还是标准都仍以人工制图综合为依据。尽管不少学者对此持有异议, 认为数字化综合不应沿用人工制图综合的模式, 但目前还没有找出其它可

行的途径。

进行自动制图综合首先需要在理论上回答为什么综合? 综合什么? 何时综合? 怎样综合? 并提出要达到的目标框架<sup>[1]</sup>; 然后给出实现这些目标的方法步骤和 workflows (它可以因目标需要的不同而不同); 最后进行程序化和计算机实现。在实现过程中一般有两种方式, 即批处理方式和交互方式。

批处理方式就是让计算机去完成综合的主要过程, 除了给出整体性参数外, 基本不进行人工干预。目前, 只有在个别方面或相当低的水平上才能达到, 例如线状要素综合。事实上, 由于地理信息综合所涉及问题千变万化, 很难预先估计出所有情况和操作, 因此实现百分之百的批处理综合几乎是不可能的, 但随着研究进一步深入, 特别是从形式的、几何的层次深入到语义的、拓扑的层次, 综合过程批处理的比例会进一步加大。

批处理方式实现的困难使人们把研究的重点转向交互方式。这种方式就是把大量机械性、非创造性的工作交给计算机完成, 而复杂、思维性工作由人机对话方式处理。通常采取的策略是将综合过程分解成一个个子过程, 用模块化或面向对象的方法设计一批相应的基本操作(如取舍、概括、简化、合并、夸大、移位等)组成通用工具包, 然后根据需要从工具包提取单项操作或进行“装配”使用。显然, 交互处理方式需要建立一个能够支持“自上而下”或“自下而上”进行综合处理的友好的用户界面, 同时, 还要求用户具有基本的综合知识和能力。

#### 3.3 基于知识的综合

无论是批处理方式还是交互方式进行综合, 都需要提高智能化程度, 以从思维的角度模拟综合过程, 因此基于知识的综合越来越引起重视。基于知识的综合就是运用人工智能、专家系统技术作决策支持的自动综合。目前, 进行决策的程序语言和推理工具基本上已经具备, 关键的问题是知识的获取和表达, 即如何把有关综合知识转化为适于智能化处理需要的形式(例如转化成“如果—则”规则等)。

地理信息综合知识一般有 3 种类型, 几何知识、结构知识和过程知识。可从多种途径来获取这些知识, 如从有关教科书和技术参考资料、现有地图和制图规范、专家制图经验等。知识获取及其表达亦可以采用多种方式, 例如, 一种较低层次的方式是用查找表列出“如果—则”规则, 形成逻辑化的程序; 另一种较有用的方式是基于超图数据结构的语义网络——

用结点表示概念或事实,用弧段表示概念之间的关系;还有一种是机器学习方式,通过对大量知识进行编辑(归纳、推理、统计等),生成决策树;此外,还有神经网络方式,通过交互式综合系统自动获取和学习知识的方式等。从长远看,基于知识的智能型综合是解决地理信息综合的根本出路,但目前的研究进展并不乐观,现有方法、技术和计算机发展水平(如平行推理能力等)决定了智能化综合的实现将是一个相当困难和漫长的过程。

## 4 有关算法的新探求

### 4.1 问题的提出

地理信息自动综合的前提是算法化。因此有关算法的探求一直是地理信息自动综合研究的主题。

在算法研究方面,比较集中和成功的探索是线状要素的自动综合算法<sup>[2,3]</sup>,这主要是因为线状要素的重要性——GIS图形的80%几乎都表现为线状要素或由线状要素组成<sup>[2]</sup>。但是,线状要素自动综合仍有不少本质性的问题没有解决,其中一些是整个地理信息自动综合中带有共性的问题。例如:(1)图形综合的基本要求是保持其形状结构特征,但图形的形状结构特征如何体现,目前尚缺乏适于计算机处理的量化指标;(2)对于指定的比例尺跨度,图形究竟应综合至什么程度,目前还缺乏客观的标准和方案;(3)阈值控制直接关系到综合的程度和质量,但这一算法实现的关键环节目前还主要以手工和经验为主;(4)综合后原有图形不同部分复杂程度的对比需得到应有的保持,但目前还缺乏有效的技术措施。事实上,从深层次上分析上述问题的存在可以发现,没有能够从数量化的

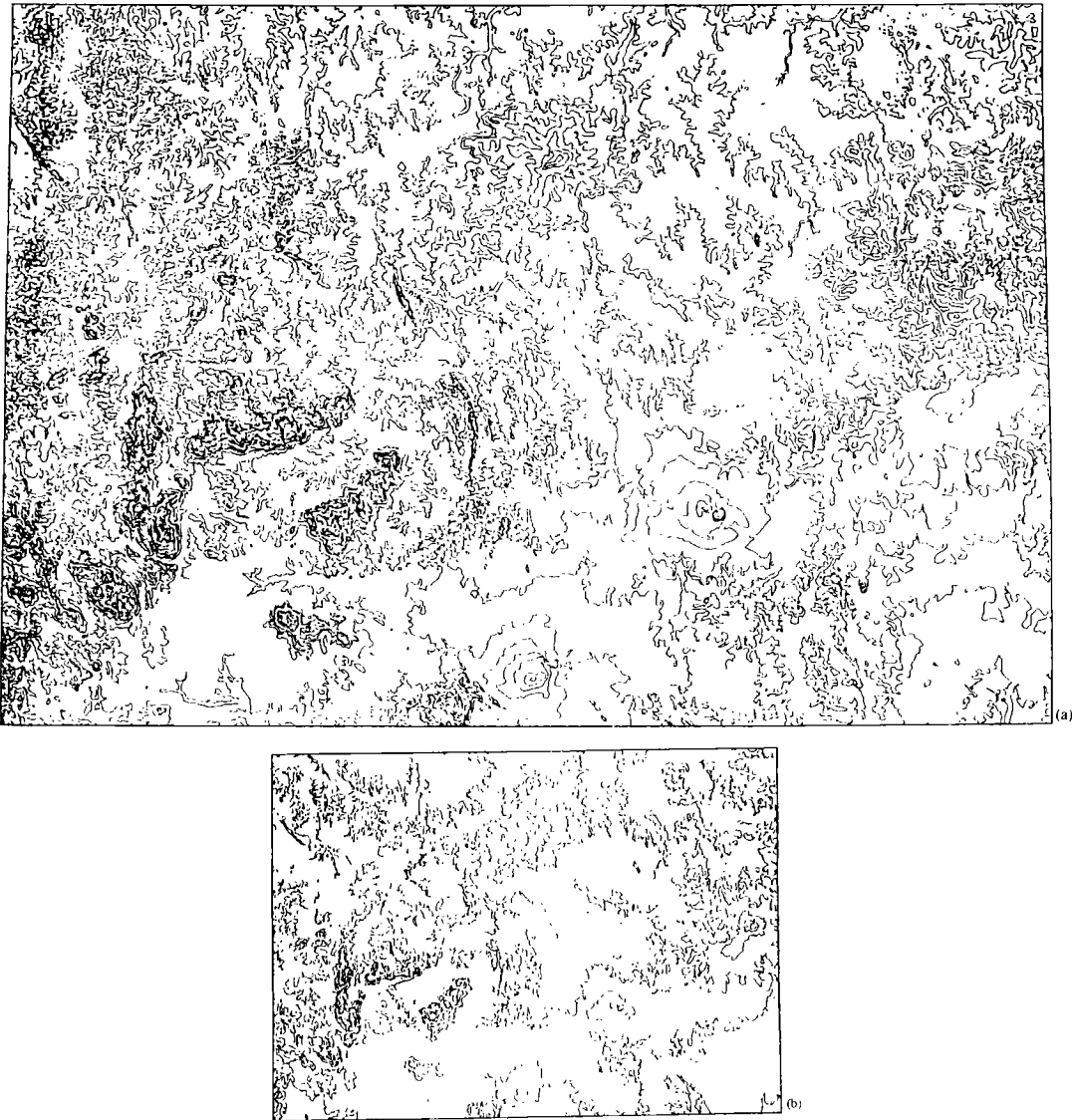


图3 算法应用实例

Fig.3 The example of the proposed method put to use

角度找出图形形状结构特征随比例尺变化的客观规律是一个带有根本性的原因。

#### 4.2 基于分形分析的算法研究

分形理论是近来发展起来的专门研究自然界复杂现象的非线性理论,在地学领域已得到大量和成功的应用。根据分形理论的自相似原理和分维概念,我们可以研究地学图形随尺度变化的数量规律,从而找出所需要的图形形状结构特征随比例尺变化的客观规律<sup>[6]</sup>。

设所考虑的自动综合对象为数字化曲线:  $L = \{P_i | i = 1, 2, \dots, n\}$ , 综合前后比例尺分母分别为  $M_1$ ,  $M_2$ ; 相应的曲线长度为  $L_{M_1}$ ,  $L_{M_2}$  (它们实际上描述了综合前后曲线的复杂程度, 其中,  $L_{M_1}$  可直接由输入数据  $L$  确定;  $L_{M_2}$  可由 Beckett 公式求出)。根据分形分析方法, 我们可得出  $\ln(s) - \ln L(s)$  图及反映尺度(比例尺)与曲线复杂程度之间变化关系的表达式<sup>[6]</sup>:

$$\ln L(s) = (1 - D)\ln s + \ln k \quad (1)$$

(1)式中  $s$  为描述曲线  $L$  所用尺度;  $k$  为常数;  $D$  为描述曲线复杂程度的量化指标: 分维数。

将  $L_{M_2}$  代入(1)式, 可唯一确定尺度  $s_2$ 。容易知道,  $s_2$  即是描述综合后曲线所需的尺度, 当采用特定的曲线简化算法(如 Douglas-Peucker 算法等)确定(1)式时,  $s_2$  既为应用该算法进行自动综合所需的阈值。由此我们建立了一种基于分形分析的阈值自动生成算法, 从而也找到一种基于图形形状结构特征随比例尺变化的客观规律的自动综合算法。

该算法主要有如下几方面特点: (1)能够根据曲线本身的形状结构特征自动提取综合所需的量化指标和进行自适应自动综合; (2)因为分维  $D$  是尺度变换下的不变量, 所以在比例尺从  $M_1$  变化到  $M_2$  的过程中, 图形形状结构特征能够得到客观的保持; (3)能够从数学机制上明确阈值与图形复杂程度、综合程度的数量对应关系, 从而对综合程度量化和客观控制、图形不同部分复杂程度对比的保持等问题的解决提供了有效途径。

图3给出了所提出算法的应用实例, (a)是综合前的海南省 1:5万 DTM 数据(经缩小); (b)是由(a)自动综合后的 1:20万 DTM 数据。

## 5 结论及讨论

研究和开发 GIS 中地理信息综合功能对于更有效的表达、理解和应用地理信息具有重要意义。GIS 有关领域的学者已开始认识到它的必要性和迫切性,

同时也认识到了所面临问题的艰巨性和复杂性。以下几方面的问题尚需得到进一步的研究和解决:

(1) 明确 GIS 环境下地理信息综合的含义、目标和障碍及其地理信息综合可能自动实现的范围、比例和途径;

(2) 完善和集成现有的各种水平的、分散的、甚至不相容的地理信息综合技术, 使之尽快形成 GIS 中的实用功能;

(3) 发展面向对象的数据模型、数据结构和数据结构动态维护技术, 以为自动综合提供更有效和更灵活方便的支持;

(4) 开发更具启发式的交互功能(如可视化的参数选择功能、综合决策模拟功能等), 以提示各种综合方案和可能的结果, 协助综合过程中的判断和决策;

(5) 发展多面向的自动综合方法(如面向多种空间单元类型、多种图形要素、多种应用专题等), 特别是多种类型要素相互关系处理方法(如移位、冲突处理、合并等);

(6) 运用各种可以利用的知识获取和表达方法将地理信息综合知识抽象成可为专家系统推理机能接受的具有严格性、准确性和系统性的“模型语言”, 提高地理信息综合中智能化的比例和水平。

## 参 考 文 献

- 1 MaMaster, R. B., Shes, K. S. Generalization in digital cartography. Washington D. C.: Association of American Geographers, 1992.
- 2 Muller, J. C. et al. GIS and generalization: methodology and practice. London: Taylor & Francis, 1996.
- 3 Grunreich, D. Generalization in GIS environment. Proceedings of the 16th ICA Conference. Cologne, 1993, 203—210.
- 4 Muller, J. C. Generalization of spatial databases. in Maguire, D. J., Goodchild, M. F. and Rhind, D. W. (eds.), Geographic Information System. London: Longman, 1991, 457—475.
- 5 Brassel, K., Weibel, R. A review and conceptual framework of automated map generalization. *Int. J. Geographical Information System*, 1988, 2(3):229—244.
- 6 王桥. 一种新分维估值方法作为工具的自动制图综合. *测绘学报*, 1996, 25(1): 10—16.

## 作 者 简 介

王桥,男,39岁,副教授。1996年于武汉测绘科技大学,获GIS方向博士学位;已发表论文40余篇,出版专著2本;目前在中国科学院遥感应用研究所GIS研究室从事GIS理论与应用方面的博士后研究工作。

## Research on Geographical Information Generalization in GIS

Wang Qiao

*(Institute of Remote Sensing Application, Chinese Academy of Sciences, Beijing 100101)*

Mao Feng

*(Institute of Geography, Chinese Academy of Sciences, Beijing 100101)*

Wu Jitao

*(Wuhan Technical University of Surveying and Mapping, Wuhan 430070)*

**Abstract** Geographical processes are scale dependent and numerous applications in GIS require changing degrees of detail or scale, thus geographic information should be made available at multiple levels of abstraction. In this sense, people are now beginning to realize the lack of generalization functionality with respect to the development of value-added products from the initial database and the update of existing databases, particularly when multiple scales are involved. Geographical information generalization is a new challenge to technical progress of GIS. This paper focuses its basic meaning and realization firstly, then introduces a new automatic algorithm based on fractal analysis, which is proposed by the authors, and discusses the further problems of generalization in GIS lastly.

**Key words** Geographical information system, Geographical information generalization, Fractal analysis