

博士后文库
中国博士后科学基金资助出版

基于本体的多尺度空间数据 模型及其一致性研究

黄 慧 著



科学出版社



博士后文库

中国博士后科学基金资助出版

基于本体的多尺度空间数据 模型及其一致性研究

黄 慧 著

科学出版社

北 京

内 容 简 介

本书系统地阐述了多尺度空间数据模型的构建及其一致性评价的实现。全书共分6章,第1章为绪论,介绍了课题的研究背景和意义等;第2章为多尺度地理空间认知,系本书的理论基础,在地理空间及尺度认知的基础上,提出多尺度地理空间的抽象模型;第3章为基于本体的多尺度空间数据模型,从本体的角度来分析空间数据的空间、时间及语义特征,并构建了多尺度空间数据模型;第4章为多尺度空间数据的一致性评价,在绘出了多尺度空间数据一致性概念、分类、评价标准的基础上,建立了多尺度空间数据的一致性评价模型;第5章为保持多尺度空间数据的更新,在介绍了多尺度数据更新研究内容、方法的基础上,给出了一套可行的多尺度空间数据更新方案、一致性的空间数据更新方案;第6章为应用实例,基于本书的理论构建了一个实验系统用于验证模型的效果。本书结构严谨,原理和方法结合紧密,丰富的图表和应用实例便于读者学习。

本书既可供地球信息科学、地学相关学科的科研人员、研究生及本科生等学习,又可供从事地理信息服务行业的开发人员及相关大专院校师生参考。

图书在版编目(CIP)数据

基于本体的多尺度空间数据模型及其一致性研究/黄慧著. —北京:科学出版社, 2017.3

(博士后文库)

ISBN 978-7-03-051922-1

I. ①基… II. ①黄… III. ①空间信息系统—数据模型—一致性—研究
IV. ①P208.2

中国版本图书馆 CIP 数据核字(2017)第 038113 号

责任编辑:李 迪 / 责任校对:郑金红

责任印制:张 伟 / 封面设计:刘新新

科学出版社 出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

北京京华虎彩印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2017 年 3 月第 一 版 开本:720×1000 B5

2017 年 3 月第一次印刷 印张:10 3/4

字数:200 000

定价:88.00 元

(如有印装质量问题,我社负责调换)

《博士后文库》编委会名单

主 任 陈宜瑜

副主任 詹文龙 李 扬

秘书长 邱春雷

编 委 (按姓氏汉语拼音排序)

付小兵 傅伯杰 郭坤宇 胡 滨 贾国柱 刘 伟

卢秉恒 毛大立 权良柱 任南琪 万国华 王光谦

吴硕贤 杨宝峰 印遇龙 喻树迅 张文栋 赵 路

赵晓哲 钟登华 周宪梁

《博士后文库》序言

1985年，在李政道先生的倡议和邓小平同志的亲自关怀下，我国建立了博士后制度，同时设立了博士后科学基金。30多年来，在党和国家的高度重视下，在社会各方面的关心和支持下，博士后制度为我国培养了一大批青年高层次创新人才。在这一过程中，博士后科学基金发挥了不可替代的独特作用。

博士后科学基金是中国特色博士后制度的重要组成部分，专门用于资助博士后研究人员开展创新探索。博士后科学基金的资助，对正处于独立科研生涯起步阶段的博士后研究人员来说，适逢其时，有利于培养他们独立的科研人格、在选题方面的竞争意识以及负责的精神，是他们独立从事科研工作的“第一桶金”。尽管博士后科学基金资助金额不大，但对博士后青年创新人才的培养和激励作用不可估量。四两拨千斤，博士后科学基金有效地推动了博士后研究人员迅速成长为高水平的研究人才，“小基金发挥了大作用”。

在博士后科学基金的资助下，博士后研究人员的优秀学术成果不断涌现。2013年，为提高博士后科学基金的资助效益，中国博士后科学基金会联合科学出版社开展了博士后优秀学术专著出版资助工作，通过专家评审遴选出优秀的博士后学术著作，收入《博士后文库》，由博士后科学基金资助、科学出版社出版。我们希望，借此打造专属于博士后学术创新的旗舰图书品牌，激励博士后研究人员潜心科研，扎实治学，提升博士后优秀学术成果的社会影响力。

2015年，国务院办公厅印发了《关于改革完善博士后制度的意见》（国办发〔2015〕87号），将“实施自然科学、人文社会科学优秀博士后论著出版支持计划”作为“十三五”期间博士后工作的重要内容和提升博士后研究人员培养质量的重要手段，这更加凸显了出版资助工作的意义。我相信，我们提供的这个出版资助平台将对博士后研究人员激发创新智慧、凝聚创新力量发挥独特的作用，促使博士后研究人员的创新成果更好地服务于创新驱动发展战略和创新型国家的建设。

祝愿广大博士后研究人员在博士后科学基金的资助下早日成长为栋梁之才，为实现中华民族伟大复兴的中国梦做出更大的贡献。



中国博士后科学基金会理事长

前言

多尺度空间数据的表达与处理是当今地理信息科学领域理论与方法研究的前沿性问题,也是国家空间数据基础设施和数字地球的重要内容。多尺度空间数据的质量及不同尺度空间数据间的一致性维护是多尺度空间数据成功应用的前提。随着地理空间信息的使用群体越来越大,应用范围越来越广,用户对多尺度空间数据的多元化及面向需求特征的要求越来越高,如何使用户在海量的多尺度空间信息中快速找到最合适的多尺度空间数据,是目前多尺度空间数据应用的主要问题,因此,需要在多尺度空间数据的描述中表达用户需求相关信息,并提供面向需求的多尺度空间数据的一致性评价方法,从而来比较不同数据源的多尺度数据与用户需求的接近程度。

在互联网时代,多尺度空间数据来源于分布在各地的数据提供者,有效使用这些多尺度空间数据必须实现多尺度空间数据间的共享和互操作。同时,对于这些多尺度空间数据间的一致性评价也要能根据用户需求进行多因素综合评价。本书从多尺度空间数据建模的本体出发,在对多尺度空间数据的尺度特征、空间关系本体、语义本体及时间本体进行研究分析的基础上,提出一种基于本体的多尺度空间数据模型,然后探讨了多尺度空间数据一致性评价指标体系和面向需求的空间要素集划分方法,在此基础上,提出了一个面向需求的多尺度空间数据一致性评价模型,并给出了一个多尺度空间数据更新方案,旨在对多尺度空间数据建模及维护更新提供参考。

本书是在教育部人文社会科学研究青年基金项目“农业产业集群的时空变异与生态效率研究——湖北省的实证研究”(12YJC790069)及博士后科学基金面上项目“农产品加工产业集群的组织结构优化与绩效提升策略研究”(2015M572172)的资助下,由笔者的研究成果凝练集结而成。尽管笔者以认真的态度开展研究工作,力求在理论及实践上取得创新性的研究成果,但由于从本体的角度来研究多尺度数据模型及其一致性问题尚属于一个全新的视角,可参考的研究成果不多,这无疑增大了本研究的难度;同时也由于笔者的研究能力与写作水平限制,因此,在理论、研究方法和学术观点等方面,可能存在许多有待改进及进一步完善之处,在此敬请读者批评指正。

黄 慧

2016年6月于武汉

目 录

第 1 章 绪论	1
1.1 研究背景和现状	1
1.1.1 空间数据的多尺度表达	1
1.1.2 传统和数字环境中的制图综合	8
1.1.3 本体的研究现状	10
1.2 立题依据和研究意义	12
1.2.1 问题的提出	12
1.2.2 研究的意义	16
1.3 研究内容与方法	17
第 2 章 多尺度地理空间认知	20
2.1 认知相关理论	20
2.1.1 认知科学的发展及影响	20
2.1.2 认知科学的研究内容	20
2.1.3 认知过程	21
2.2 地理空间认知	24
2.2.1 地理空间	24
2.2.2 地理空间认知内容	27
2.2.3 地理空间抽象过程	30
2.3 尺度认知与分析	32
2.3.1 尺度问题的由来	32
2.3.2 尺度认知	33
2.3.3 地理信息系统中相关尺度概念	36
2.4 广义尺度概念	39
2.5 广义尺度下地理空间表达框架	42
2.5.1 地理空间数据多尺度表达	42
2.5.2 地理空间表达的广义尺度框架	43
2.6 多尺度地理空间的认知与抽象模型	46
2.6.1 地理空间认知的尺度特征	46

2.6.2 多尺度的认知空间	47
2.6.3 地理空间的多尺度抽象过程	48
第3章 基于本体的多尺度空间数据模型	50
3.1 本体与地理本体	50
3.1.1 本体的概念	50
3.1.2 地理本体	52
3.1.3 本体语言	52
3.2 本体与尺度	56
3.2.1 尺度特征	56
3.2.2 尺度的本体认知	56
3.2.3 多尺度地理空间信息本体认知	58
3.3 空间关系的本体分析	59
3.3.1 空间关系分类	59
3.3.2 空间关系特征	61
3.3.3 空间关系的形式化表达	62
3.3.4 空间关系推理	68
3.3.5 空间关系本体	68
3.4 时间关系的本体分析	69
3.4.1 时间本质的认知	70
3.4.2 时间关系的形式化描述	70
3.4.3 时间关系推理	71
3.4.4 时间本体	72
3.5 语义关系的本体分析	74
3.5.1 语义关系分类	74
3.5.2 多尺度空间数据间语义层次关系	75
3.5.3 多尺度空间数据语义层次匹配	76
3.5.4 基于本体的空间信息语义互操作	77
3.6 多尺度空间数据建模本体分析	78
3.6.1 多尺度数据模型特征	78
3.6.2 基于本体的多尺度空间数据建模	78
3.7 基于本体的多尺度空间数据模型	80
3.7.1 基于本体的多尺度空间数据模型结构	80
3.7.2 多尺度地理本体	82
3.7.3 基于本体的多尺度空间数据组织	84

第4章 多尺度空间数据一致性评价	87
4.1 多尺度空间数据一致性概念	87
4.1.1 一致性的定义	87
4.1.2 一致性描述的层次	89
4.1.3 一致性度量	90
4.2 多尺度空间数据一致性分类	91
4.2.1 基于内容的分类	91
4.2.2 基于对象关系的分类	94
4.2.3 基于对象抽象方式的分类	94
4.3 多尺度空间数据一致性评价标准	95
4.3.1 相关影响因素	96
4.3.2 拓扑一致性评价	97
4.3.3 方位一致性评价	110
4.3.4 结构一致性评价	112
4.3.5 语义一致性评价	113
4.3.6 时间一致性评价	115
4.3.7 一致性评价指标体系	115
4.4 本体驱动空间要素权重划分	117
4.4.1 本体驱动空间要素划分标准	117
4.4.2 本体驱动空间要素分类体系	118
4.4.3 空间要素类一致性评价权重判定	120
4.5 多尺度空间数据一致性评价模型	122
4.5.1 面向需求一致性评价模型	122
4.5.2 评价模型体系结构	122
4.5.3 一致性评价过程	123
4.5.4 评价结果的表达与分析	126
第5章 多尺度空间数据更新	127
5.1 研究现状	127
5.2 多尺度空间数据更新研究内容	128
5.2.1 更新的内容	129
5.2.2 更新数据源选择	131
5.2.3 多尺度空间数据库更新方式	132
5.3 多尺度空间数据更新关键技术	133
5.3.1 更新要素的空间匹配	133

5.3.2 变化检测	134
5.3.3 不同尺度地理要素的关联	134
5.3.4 空间目标自动制图综合	134
5.4 多尺度空间数据更新方案	135
5.4.1 总体设计思想	135
5.4.2 单一尺度空间数据更新	136
5.4.3 多尺度空间数据更新	137
第6章 应用实例	139
6.1 实验系统设计	139
6.1.1 软硬件环境	139
6.1.2 逻辑结构	140
6.2 实验数据准备	141
6.3 多尺度空间数据更新实例	143
6.4 多尺度空间数据一致性评价实例	145
参考文献	147
编后记	158

第 1 章 绪 论

尺度问题是一个古老的问题,是包括地理信息科学在内的各种学科中最重要、但迄今为止仍未解决的问题之一 (Li, 1999; Bruegger, 1995),它是人们认知地理对象、地理空间和地理现象的基础 (Montello, 1993; UCGIS, 1996)。作为认知主体的人类由于不同的文化背景,认识客观世界的角度不同,而对于不同的应用目的也需要用不同的方式来表达同一地理现象。因而空间数据的多尺度表达与处理是符合人类思维习惯的一种自然表示方法。随着信息和知识时代的来临,人类对地理信息的需求结构向知识化方向发展,传统的地理信息系统也逐步向地理信息数据的集成和网络地理服务转移,空间数据的信息服务将成为主要需求。人类需求的多元化决定了空间数据处理和显示的多样性,并成为地理信息系统实际应用的核心问题之一。

1.1 研究背景和现状

1.1.1 空间数据的多尺度表达

由于人类的认知能力和接收信息量的有限性,不可能观察地理世界的所有细节,因此尺度必定是所有地理信息的重要特征。加之许多地理现象和过程的尺度行为并非按比例线性或均匀变化,相应地需要研究地理实体在不同尺度上的表达,以及实体表达随尺度变化的规律。因此,如何为用户提供多尺度的空间数据,并对多尺度的空间数据进行有效的存储、处理和输出成为地理信息系统伴随信息和知识时代到来而出现的新课题。随着数字地球、数字省、数字区域和数字城市研究与应用的深化,多尺度空间数据模型理论与实现技术的研究已经成为美国地理信息系统 (Geographic Information System, GIS) 研究的热点之一 (龚健雅等, 2000; 陈军和蒋捷, 2000; Jones and Kidner, 1996)。

国内外的地理信息组织和专家都将空间数据的多尺度表达和处理列为重点研究课题:美国国家地理信息与分析中心 (National Center for Geographic Information & Analysis, NCGIA) 于 1988 年在其创新研究计划中提出研究空间数据的多重表示问题 (Buttenfield and Delotto, 1989a; Buttenfield, 1993; NCGIA, 1993); 1992 年 SMALLWORLD 的 Richard G. Newell 等将空间数据多尺度处理与表示列入 GIS 领域十大困难问题之一 (Newell and Theriault, 1992); 1996 年 6 月美国大学地理

信息科学协会 (University Consortium for Geographic Information Science, UCGIS) 也将该问题列为未来十年地理信息科学的十个优先研究领域之一 (UCGIS, 1996); 1997 年 NCGIA 的 VARENTUS 基金将“地理细节的形式化概念”列为高度优先的认知研究项目, 以研究信息认知中的尺度、详细程度及多尺度表示等多方面的问题 (NCGIA, 1997; Mark et al., 1999); 在一系列国际 GIS 与地图学大会上, “Multi-Scale GIS”均被列为中心议题 (Jones, 1991; Egenhofer et al., 1994c; Timpf and Frank, 1995; NCGIA, 1997; UCGIS, 1998); 在 1995 年 (西班牙巴塞罗那) 和 1997 年 (瑞典斯德哥尔摩) 两届国际制图协会会议上, 多篇与该问题有关的论文亦被列为大会重点宣读的行列 (Oosterom and Schenkelaars, 1995; Govorov, 1995; Woodsford, 1995); 2001 年 NCGIA 再次将空间数据的多种表达方式问题列为美国当代 GIS 研究的 19 个方向之一 (张永忠译, 2001); 国际摄影测量学会 (ISPRS) 数据综合与数据挖掘工作组联合国际制图协会 (ICA) 地图综合委员会也于 2002 年 7 月在加拿大渥太华以“空间数据的多尺度表达”为主题召开学术会议, 研讨与多尺度表达有关的基础理论及应用领域内的有关问题 (ISPRS, 2002)。中国地理信息系统协会第二届年会 (1996) 的“数据库和数据模型”分会场的热门话题之一也是多比例尺 GIS。在国家构建“数字中国”地理空间基础框架的总体战略中也将空间数据的多尺度表达列入地理空间数据库关键技术研究的核心理念之一。同时, 国家自然科学基金委员会也将地理空间尺度及基于尺度的智能化自动综合列为优先资助领域。

随着人们对尺度问题重要性认识的加深, 空间数据的多尺度问题的研究范围也越来越广泛和深入。1993 年 NCGIA 为空间数据的多尺度表达问题定义了数据模型、多重表示之间的连接、所实现视图的维护、空间模拟、综合问题等 5 个研究领域 (Buttenfield, 1993)。UCGIS (1998) 认为未来的 GIS 应该是尺度依赖的, 应该为用户提供尺度管理工具。为此应该研究: 开发能了解并描述尺度行为的空间数据模型; 先进的、高质量的对尺度影响的理解能力; 具有描述数据尺度的新颖方法; 尺度变换的智能化方法等 (UCGIS, 1998)。2002 年 ISPRS 数据综合与数据挖掘工作组和 ICA 地图综合委员会召开的学术会议列出了: 多尺度/多重表示数据库; 数据库/模型综合; 基于多尺度表示, 从图像和矢量数据集中提取目标; 利用层次、多尺度结构进行图像解译和分类; 图像与 GIS 数据的集成与匹配; 三维可视化与综合; 几何与语义数据的匹配/一致; 多尺度数据库中传递更新与版本控制机制; 基于空间位置的服务和小目标的地图显示; 与国际互联网、互操作性和空间数据基础设施等有关的多尺度方面; 矢量域提取目标的算法在图像域的应用; 质量评估技术等 12 个研究主题 (ISPRS, 2002)。

从理论与技术方法研究的角度, 对空间数据多尺度问题的研究主要集中在以下 4 个方面。

1. 多尺度空间数据的存取与表达

空间数据的多重表示是地理信息系统中一个重要的研究主题 (Buttenfield, 1993)。由于人类不同的文化背景和专长, 认知客观世界的角度也不同, 因而在 GIS 数据库中存储现实世界实体的多种表示是十分自然的。实现空间数据的多重表示可以有多种途径或方案, 目前具有代表性的方案有多库多版本、单库多版本 (齐清文和张安定, 1999)、单库单版本 (Oosterom and Schenkelaar, 1995) 和 LOD (Lodestar, 1997; 王家耀, 2001) 等 4 种。

多库多版本方案是在数据库中存储来自同一个现实世界实体的多种表示, 即建立对应于多种比例尺的多个数据库, 如图 1-1A 所示。数据库通过采集不同比例尺的数据来建立, 图形显示时通过控制当时屏幕比例尺的变化, 轮换调入和释放相应尺度的数据实现。该方法是一种静态的空间数据组织方法, 在一定程度上缓解了系统对不同详细程度数据的迫切需求与自动综合相对落后的矛盾, 对早期 GIS 的发展和应用起了很大的推动作用。其问题是各比例尺数据独立采集, 造成人力、物力和财力的浪费; 在图形显示时, 一种比例尺转换到另一种比例尺时会出现明显的不协调和不连续的现象; 由于各种比例尺的数据之间没有任何联系, 数据更新困难, 各比例尺的数据必须分别进行更新, 系统数据的一致性很难保证。

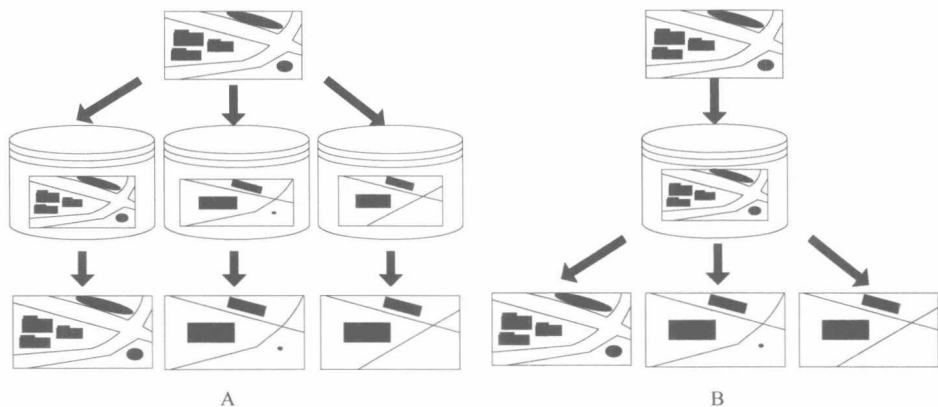


图 1-1 空间数据多重表达方案
A. 多库多版本方案; B. 单库多版本方案

单库多版本是所谓的主导数据库的方法, 即维护一个单一的、有较高详细程度的主导数据版本, 任何其他较概略的尺度数据都是从该版本的数据派生出来的, 如图 1-1B 所示。当一个地理实体发生变化时, 其变化能通过同一地理要素之间的连接关系传播到相关的其他实体, 从而更新尺度的数据, 实现多重表示。这是目前认为比较现实可行的方法, 许多专家学者对这种方案的具体实施进行了研究。Oosterom (1995) 利用反作用树 (reactive-tree) 技术, 在开放式的面向对象数据

库管理系统 Postgres 中建立了一个交互式的多尺度 GIS-GEO++。Timpf (1998) 在其博士论文研究中基于图论也建立了这样的实验系统 DAG。一些商用 GIS 软件如 Smallworld GIS、Intergraph DynaGen 等已开发了部分类似功能,但都未从空间数据处理模型上作整体考虑,目前仍需要进一步研究。该方法努力的方向是开发更强大、更智能化的支持多重表示的数据结构,包括面向对象的方法和语义数据模拟,以及基于面向对象的方法建立智能识别处理程序,并为每一要素添加丰富的语义数据(吴凡,2002)。因为空间实体不仅有图形表示,还有描述其性质的一般性说明信息,即语义部分。只有对空间目标的图形部分和语义部分都进行多尺度的处理和表示,才能真正完成从主导数据库导出不同尺度数据集的任务。此外还必须自动建立同一空间实体不同尺度之间的相关和互动机制,才能满足基于尺度的有效的综合分析和辅助决策需要。

单库单版本方案是在数据库中用一种称为反应数据结构的存储结构存放系统中最详细尺度的数据,其他比例尺的数据信息隐含在该存储结构中,需要某种比例尺的数据时直接从存储结构中实时提取,不必存储这些数据。这是一种比较理想的然而也是最难以实现的方法,是地图学家努力的方向(王家耀,2001)。该方案可以在较短的时间内提取并显示任意比例尺的空间数据,但只适合于地图综合中通过简化和选择两种方法得到的综合结果,对其他综合方法得出的结构就无法用这种数据结构来存取。

LOD (level of detail) 技术是一种采用不同的细节程度来描述并存储对象的一组模型的方法,也称为细节分层技术(王家耀,2001)。该技术符合人的视觉规律,即在远近不同的位置上观察同一地区或同一地物,人眼所能观察到的该地区或地物的详细程度是不同的。所以,从技术角度上讲,我们可以根据人眼的视觉规律,为同一地区或同一地物构建一组不同详细程度的数据模型。在计算机中显示时,根据地区或地物离视点距离的大小,调入相应详细程度的数据模型来生成视图。LOD 方案实际上是方案 1 和方案 2 的结合,该方案常用于栅格数据组织和虚拟环境建模中。

2. 空间数据的多尺度处理模型

数据模型设计是信息系统设计的核心和首要问题,它表达了系统设计人员对客观世界的认知和抽象过程。经过诸多专家学者的多方面研究,数据模型的发展取得了很多成果。但由于长期以来 GIS 空间数据模型的设计没有对尺度及其尺度效应给予应有的重视,因而先天性不足,使现有 GIS 没有尺度维的处理和表达能力(Li, 1999),故 GIS 未来发展和应用面临的一个主要问题是研究新的空间数据模型(Mennis et al., 2000)。由于不同尺度下的空间目标有不同的运动状态和关系,目标的维数也会发生变化,因此多尺度空间数据处理模型的建立异常复杂(Li, 1999)。

对多尺度矢量数据模型的研究已有较长的历史,其中较有代表性的是 Molenaar (1989) 提出的 Formal 数据模型。该模型以点为基本存储单位,基本思路是点构成弧,弧构成线面实体,可称为点模型。Bruegger 和 Frank (1989) 形式化地描述了一种独立于空间数据维数的方法,用于构造多类结构相关的空间实体表示。同一实体在不同的层次上表达不同的空间分辨率和详细程度,并用实体之间的层次关系互相连接各个层次 (Bruegger and Frank, 1989)。李霖 (1997) 分析研究了数据库描述空间目标的特点和方式,针对语义层次定义了集合元素的聚合算子,并以此聚合算子为基础,提出了多尺度空间目标聚合模型及幂集的查询模型。Medeiros 等 (1996) 提出一个多层次视图地理参考模型,通过将数据存储在多版本数据库中来管理数据的多重表达。该模型允许每个用户在多版本数据库中定义其各自的视图,并在选择其感兴趣的要素类的同时保持它们与基础数据库之间的连接。Timpf (1998) 在其博士论文中提出一个基于空间剖分的地图立方体模型 (map cube model)。该模型将整个地图空间剖分为网络 (networks)、容器 (containers)、面域 (areas)、要素基元 (elements) 4 个层次,并用树结构来维护不同层次下数字地图的表达,同一层次下的要素集合形成了该层次 (比例尺) 下的地图。这种空间层次剖分方法一方面处于理论设计阶段;另一方面,立方体模型将地图空间完整剖分,进行分层抽象,不具有实用性。Leung 等 (1999) 提出的面向对象的通用 GIS 概念模型,引入空间层次、实体层次、要素层次的层次概念来研究要素的多层次抽象问题,在实体层存储所有的实体 (包括简单实体和复合实体) 和实体间的空间关系,所有的实体按自上而下的方法排列成层次结构。该模型虽然考虑了要素的多尺度表达,但只是简单带过,没有考虑多尺度表达之间的尺度关系,而且其总体上仍处在概念模型的阶段,并没有真正实施。Borges 等 (2001) 所设计的 OMT-G (object-oriented model for geographic application) 模型认为,面向地理实体应用的数据模型应该根据用户的认知及空间关系的需要对地理实体的几何形状作出明确的表达,能够表达一个特定地理对象的几种不同的抽象视图,但只在概念模型层次上区分了依比例尺绘制和不依比例尺绘制的两种抽象层次,而且没有考虑它们之间的联系,其他数据模型 (如 GISER、GMOD、MODULO-R) 考虑了多视图对地理实体的影响,但并没有针对专门地理要素的多尺度特征进行抽象。王晏民在操作速度方面对 Molenaar 的 Formal 数据模型进行了改进,提出了以弧为基本存储单位的弧模型,其基本思路是弧构成实体,实体也作为临时存储单位。但其不足之处在于没有显示表达出同一要素不同尺度下的对应关系 (王晏民, 1996, 2002)。张锦提出了一个适用、完整的超图对象模型 HOOM,将空间现象及其表达分为 5 级或 5 层结构,包括:几何对象模型、地理对象模型、时态对象模型、地理表现对象模型和空间计算对象模型 (张锦, 1999, 2004)。这 5 级模型的集成组合构成了超图对象模型 HOOM。并用超图对象模型来组织特征数据,但其

研究只是初步的。郭武斌(2009)针对车辆导航领域的多尺度空间数据模型设计了用于提高物流车辆导航路径分析速度与精度的多尺度空间数据模型,周捍东等(2014)设计了一种支持多尺度、多层次的空间数据模型,解决特大型城市复杂公交网络的数据组织问题,为多尺度数据模型的应用进行了有益的探索。至今,学术界尚未能找到合适的方法和技术来形式化定义和构建多尺度数据模型,也没有形成一套形式化操作算子规则来管理数据(Kilpelainen, 2000a),提出的地理要素多比例尺数据模型也还停留在概念设计阶段。

3. 面向空间数据多尺度处理的自动综合

地理信息的自动综合是实现空间数据多尺度表示的核心技术。面向空间数据多尺度处理的自动综合的目标是实现智能化的空间数据多尺度处理。它主要包括3个方面的研究:自动综合的概念框架,综合操作算子和综合知识规则的形式化描述。

在自动综合概念框架研究方面,Brassel和Weibel(1988)曾提出了一个概念框架以识别手工综合过程的主要步骤,并将这些概念转换到数字领域。它将操作、知识、阈值均存储在程序库中,是一个集成专家系统技术的自动综合模型。McMaster和Shea(1992)扩展了Brassel和Weibel的模型。该模型将综合分解为3种操作:从哲学意义上考虑为什么(why)要综合;从地图学观点评估何时(when)该综合;选择适当的空间和属性变换为如何(how)综合提供技术。但这样一个完整而复杂的系统至今尚未实现。智能化综合的基础是基于知识的推理。为了减少每次派生新地图时的人工引导,Weibel(1991)曾提出了“智能增强”的方法。但基于知识的方法仍然缺乏形式化的制图知识和面临如何获取知识的问题(Weibel et al., 1995; 郭庆胜, 1998; Kilpelainen, 2000b)。

综合操作算子是从制图综合过程中提取的基本综合操作元素,每一次综合操作由一个或一组综合操作算子来完成。对综合操作算子的研究,众多学者提出了不同的体系。主要有:Beard和Mackaness(1991)的九算子模型:选择、分类、符号的改变、符号的感受层面的提高、尺寸区别的加大、省略、组合、位移和夸大;Shea和McMaster(1992)的十二算子模型:简化、光滑、聚合、混合、合并、收缩、精选、典型化、夸大、增强、位移和分类;Schlegel和Weibel(1995)的七算子模型:选择、删除、化简、光滑、混合、夸大和位移。划分的种类如此繁多,可见地图自动综合到底需要哪些算子,究竟应如何分类仍在争议之中。

制图综合可以定义为当比例尺不断变小时,简化图面上地物表达复杂程度的过程。对地物表达的简化需要遵循一定的规则和知识,以便达到与人类自然认知获得知识的一致性表达。自动制图综合所需的知识类别包括:①关于所表达要素的知识;②关于几何、语义质量的知识;③关于数据集表达的知识;④关于产品目的(比例尺/分辨率)的知识;⑤关于综合算法的知识;⑥关于综合活动(操作、

控制、顺序)的知识;⑦关于符号化的知识(Lagrange, 1997)。制图规则的产生必须考虑目标之间的空间和属性关系,而这些关系之间的差别可能很大,从而使规则产生的过程非常复杂。为此,Kang等(2000)研究并形式化了一套模型的综合规则,试图简化这个过程。智能制图综合的核心是基于知识对综合过程的理解和形式化。这些知识包括深层知识和浅层知识:浅层知识是关于客观事物的表象及其与结论之间关系的知识,深层知识则是有关事物本质及其因果关系的内涵、基本原理等类型的知识(徐洁磐等,2000)。McMaster和Buttenfield(1997)强调了形式化地图综合知识的重要性,并列举了浅层知识与深层知识的例子。目前,基于符号智能的产生式规则在应用上会有局限性,因为自动综合包含的一些主观成分不容易形式化为逻辑规则,导致一个应用或区域的综合规则可能不适用于另一个。因而面向应用的自动制图综合必须考虑上下文和应用背景关系。

4. 多尺度空间数据的一致性

多尺度空间数据库的一个关键问题是维护多重表示空间目标之间的一致性。所谓数据一致性,是指同一客体的不同数据之间在属性继承关系、图形拓扑关系等方面不存在任何逻辑上的矛盾。而“数据不一致”就是对数据一致性限制条件的侵犯,它导致用户数据查询结构的不稳定、不规则或与预料的结果相去甚远的现象(齐清文和张安定,1999)。不同尺度表达层上空间目标的一致性是很重要的,因为它允许在一个较概略层上的查询能够给出与在详细层上相同查询有非常相似的结果。最明显的例子是多尺度的土地利用查询,其不同尺度上的土地利用面积的查询结果应该一致。因此评估不一致性的方法和维护策略是这一研究的重点内容(Egenhofer et al., 1994a; Oosterom, 1997; Carvalho, 1998)。这对于多重表示的数据库的更新也具有十分重要的意义。在对多尺度空间数据的一致性研究中,较具代表性的有:MAGE系统中大比例尺地图应用三角网数据结果表达,在应用化简和夸张等综合操作时,通过维护三角剖分的拓扑来实现空间关系的拓扑一致性(Bundy et al., 1995)。Egenhofer等提出一个评估多重表示拓扑一致性的框架(Egenhofer et al., 1994a; Carvalho, 1998),该框架提出的理论依据是任意目标的拓扑,以及任意目标之间的拓扑关系随着不同尺度层次必须保持相同或持续降低其复杂性和细节。这个方法基于拓扑关系模型,通过描述空间目标的内部、外部和边界的相交内容是否改变,以及描述相交的构成是否改变来判断其一致性。郭庆胜等(2005a, 2005b, 2006a)基于基本空间拓扑关系组合描述的方法分析了线与线、线与面、面状目标之间的拓扑关系,并进行组合推理,得出其拓扑关系的组合推理表。该方法能细分空间对象间的拓扑关系。从空间抽象的角度来看,只有充分区分空间拓扑关系后,才能进一步描述空间拓扑关系的抽象问题,为多尺度数据库和多重表达的拓扑一致性(或等价性、相似性)评价提供理论基础。