

一种多比例尺 GIS 方案及其数据模型

王晏民¹ 李德仁² 龚健雅²

(1 北京建筑工程学院测绘工程系, 北京市展览路 1 号, 100083)

(2 武汉大学测绘遥感信息工程国家重点实验室, 武汉市珞喻路 129 号, 430079)

摘要: 提出了一种分层分级分区式多比例尺 GIS 方案, 在提出的空间数据模型基础上, 按照地图综合方法, 发展了适合本文方案的多比例尺 GIS 空间数据模型, 并通过试验说明了新方案及其模型的正确性和可行性, 讨论了实施中几种情况的解决办法。

关键词: 多比例尺; GIS; 数据模型; 地图综合

中图法分类号: P208; TP311. 12

多比例尺 GIS 是指同时存贮、管理、处理、表达和传输多种比例尺空间数据的 GIS, 其核心就是建立多比例尺空间数据库的问题, 其目的是提供变焦数据处理的能力。实现多比例尺 GIS 可以有多种途径或方案, 具有代表性的方案有多库多版本、单库多版本^[1]、单库单版本^[2]和 LOD^[3~5]等 4 种。多库多版本中各比例尺数据独立采集, 造成人力、物力和财力的浪费; 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, 数据更新很麻烦, 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。单库多版本将多个版本的数据存放在一个数据库中有时是难以实现的; 各个版本是从主导版本派生出来的, 有的数据是重复的, 造成数据冗余。单库单版本只适合于地图综合中的简化和选择两种方法综合的结果, 对其他综合方法得出的结果无法存取; 存在海量数据存放在一个数据库的问题。而 LOD 只考虑了纵向分层, 没有考虑到横向分区。

1 分层分区分级方案

分层分级分区式方案的基本思想是, 先将目标区分成若干比例尺层次, 以最上层比例尺的空间数据作为主导版本(该版本可以是独立采集的,

也可以是较大比例尺版本用制图综合方法派生的数据), 用该版本向上派生更小比例尺的版本, 直到屏幕能够显示全图为止。用主导版本对下层比例尺版本的空间数据进行分区, 形成多个分区的大比例尺版本, 将这些分区版本作为各分区的主导版本; 再对分区主导版本向上派生, 向下分区, 直到满足要求为止。各分区版本可以是单独的数据库, 也可以是分布在网不同计算机上的数据库。

1.1 总体分层

首先将整个目标区域按比例尺大小分成若干个比例尺层次。图 1 中将全区分为 k 个层次。

1.2 层中分区

如果单个数据库的容量不够存贮一层中的空间(以及属性)数据, 就必须将该层数据进行分区, 可以是格网分区, 也可以是多边形分区^[6]。在此规定第一层只能有一个分区, 也就是第一层的数据必须能够存贮于一个数据库中, 其目的是便于调用全图。图 1 中的一个圆锥体表示一个分区, 第 1 层为一个分区, 第 2 层为 n 个分区等。

1.3 区内分级

由于层与层之间的比例尺相差较大, 若调用的比例尺与层的比例尺不一致, 就会产生缩放问题。放大调用空间数据时, 精度和详细程度一般都不能满足要求, 缩小调用时又会产生数据量大、详细度太高的问题。因此, 层与层之间应该加密

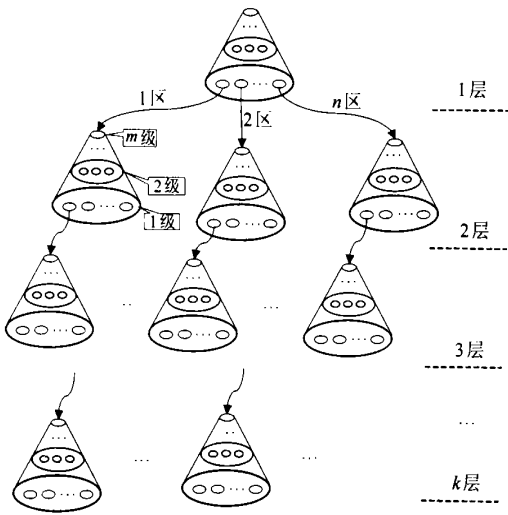


图 1 多比例尺 GIS 的分层分级分区式方案

Fig. 1 Layered, Leveled and Partitioned Scheme for Multi-scale GIS

一些中间比例尺的数据。加密方法是以分区数据为主导版本, 向上进行地图综合, 逐级加密比例尺的数据版本, 最上一级比例尺的数据版本能以合适的密度在屏幕上全图显示整个分区。分级时, 同一层不同分区应该有相同的分级, 便于分区之间各级版本数据的连接。图 1 中每个圆锥体分为 m 级。

2 多比例尺数据模型

多比例尺模型应该满足两个要求: ① 能够表达各种综合方法所得的结果; ② 数据更新时能够向上传递, 使各级比例尺的数据都得以更新, 保证系统的一致性。本文以笔者提出的矢量要素拓扑链纯关系模型^[7]为基础, 采用分层分级分区式方案, 设计了满足上述两个要求的多比例尺数据模型。

2.1 要素对象的分类

1) 简单要素。只有一个几何对象构成的要素对象称为简单要素对象, 简称简单要素。几何对象是指构成空间对象的基本几何形状, 二维几何对象有点、线和面三种基本形状。例如, 一口井由一个点构成, 一幢房子由一个多边形构成, 一段路由一条曲线构成等。这里将点、线和面 3 种基本几何对象在构成要素时合并成了一种简单要素, 用各自的属性码来区分它们是点、线还是面, 使数据库的结构大大简化。

2) 聚合要素。由多个几何对象构成的要素对象称为聚合要素对象, 简称聚合要素。例如, 一

个烟囱是由一个点和一个多边形构成的聚合要素, 一座山可以用多个离散高程点来表示, 也可用多个多边形构成的等高线表示, 还可用多条直线构成的不规则三角网(TIN)表示。定义了聚合要素对要素的扩充是很有意义的, 例如, 在几何对象中增加像元和栅格, 那么用聚合对象就很容易定义一个面状要素的影像表达和栅格表达, 形成了多源数据一体化的结构。

3) 复合要素。由多个要素对象构成的要素对象称为复合要素对象, 简称复合要素, 有的文献称其为复杂要素^[8]。例如, 城市是由道路网、房屋、公园等多种要素构成的复合要素。

2.2 制图综合方法

制图综合方法的分类有多种方式, 比较常见的有简化、移位、夸张、合并和淘汰等。

1) 简化。将大比例尺中形状复杂的空间对象在小比例尺中简化为比较简单的空间对象, 主要是对各种曲线或折线按照比例尺的要求提取特征点^[7]。

2) 移位。为了消除矛盾, 将大比例尺中的要素在小比例尺中移动位置。如道路旁边的房屋, 比例尺缩小到一定程度时就会与道路相交, 为了消除这一矛盾, 可将房屋往旁边移位。

3) 夸张。将那些按比例尺无法表示的重要要素用夸张的手法表示出来。如北京市在全国地图上的表示等都是经过夸张的。

4) 合并。将大比例尺中的几个有关联的要素在小比例尺中合并成一个要素, 也称聚合(aggregation)或综合(generalization)。例如, 将多个房子和院落合并成一个村庄。由于合并的情况很复杂, 很难找出适合于各种情况的通用合并方法, 所以它是综合自动化的瓶颈。

5) 淘汰。将大比例尺中的一些要素在小比例尺中去掉, 或者说将大比例尺中的一些同类要素有取舍地转到小比例尺中。例如, 高程注记的密度要求 $2 \sim 3\text{cm}$ 一个点, 大比例尺转到小比例尺时就要淘汰一些高程点, 以保持密度。

2.3 概念模型

图 2 是按照前述要求设计的多比例尺 GIS 概念模型中与多比例尺有关的部分。模型将多比例尺的关系定位在要素对象之间, 几何对象之间的关系不受比例尺的影响。从图 2 中可看出, 聚合要素与复合要素之间没有综合关系, 也就是说, 复合要素不能综合成聚合要素; 所有要素之间的综合都是一对 0 或一对一的关系。下面讨论各要素间的综合关系。

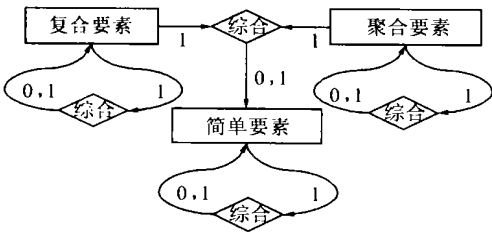


图 2 多比例尺 GIS 概念模型的结构

Fig. 2 Structure of the Model of Multi-scale GIS

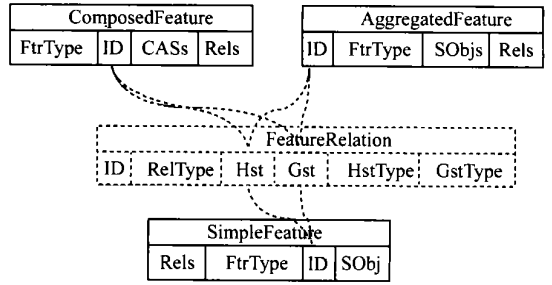


图 3 要素之间关系的逻辑结构

Fig. 3 Logical Structure of the Relation Among Features

1) 简单要素之间。其一对 0 的关系表示简单要素经过综合变没了, 该关系可表达淘汰法对简单要素的综合结果。必须说明, 只有简单要素才可以被完全淘汰, 聚合要素和复合要素不能直接被完全淘汰, 必须蜕化成简单要素后才能被淘汰。其一对一的关系表示简单要素经过综合还是简单要素, 该关系可表达简化、移位和夸张等 3 种方法对简单要素的综合结果。

2) 聚合要素之间。其一对 0 的关系表示聚合要素经过综合变成了非聚合要素(即简单要素)。其一对一的关系表示聚合要素经过综合还是聚合要素, 该关系可表达简化、移位、夸张和淘汰(选择)4 种方法对聚合要素的综合结果。

3) 复合要素之间。其一对 0 的关系表示复合要素经过综合变成了非复合要素(即简单要素)。其一对一的关系表示复合要素经过综合还是复合要素, 该关系可表达全部 5 种方法对复合要素的综合结果, 但对于淘汰法只能是淘汰复合要素中的部分要素, 说选择更合适一些。

4) 聚合(复合)要素与简单要素之间。其一对 0 的关系表示聚合(复合)要素经过综合还是聚合(复合)要素。其一对一的关系表示聚合(复合)要素经过综合蜕化成了简单要素, 该关系主要是表达合并方法对聚合(复合)要素的综合结果, 在合并的同时也可能伴随着简化、移位、夸张和淘汰。

2.4 逻辑模型

图 3 是从图 2 所示概念模型的逻辑结构。SimpleFeature 表用来存贮简单要素对象, 它由主键 ID、要素类型字段 FtrType、几何对象字段 SObj 以及与其相关的其他要素字段 Rels 组成。AggregatedFeature 表用来存贮聚合要素, 其 SObj 是一个不等长字段, 存贮构成聚合要素的几何对象的形态码、个数及其地址码。如 SObj 字段中的内容为 0, 2, 4, 19, 2, 3, 3, 45, 88, 可理解为该聚合要素由 2 个点状要素(在点表中的第 4 和 19 个点)和 3 个面状要素(在多边形表中的第 3、45 和 88 个多边形)组成。ComposedFeature 表用来

存贮复合要素对象, 其 CASs 是一个不等长字段, 存贮构成复合要素的要素对象的类别、个数和地址码。如 CASs 字段中的内容为 0, 1, 55, 1, 3, 12, 18, 32, 2, 2, 46, 77, 可理解为该复合要素由 1 个简单要素(在简单要素表中的第 55 个要素)、3 个聚合要素(在聚合要素表中的第 12、18 和 32 个要素)以及 2 个复合要素(在复合要素表中的第 46 和 77 个要素)组成。在 3 个要素表中, 要素类型字段 FtrType 此时不仅表示要素的类型, 还可说明比例尺的级别。FeatureRelation 表由主键 ID、关系类型字段 RelType、主要素字段 Hst、次要素字段 Gst、主要素类别字段 GstType 和次要素类别字段 GstType 组成。Hst 是关系的主角或施与者, Gst 是关系的配角或接受者。

综合关系可用“综合成...”来表示。在 FeatureRelation 表中, 可规定 RelType = 11, 表示“Hst 综合成 Gst”的关系。将这个关系记录到 Hst 的要素表的 Rels 字段, Rels 中的关系类型为 11、综合的结果要素为 Gst 的地址码、结果要素的类别码为 Gst 的值。例如, 在 SimpleFeature 表中, 某个记录的 Rels 字段值为 11, 34, 0, 可解释为本简单要素综合成 SimpleFeature 表中第 34 号记录的要素。如在 AggregatedFeature 表中, 某个记录的 Rels 字段值为 11, 58, 0, 可解释为本聚合要素综合成 SimpleFeature 表中第 58 号记录的要素。

因为数据更新时, 只需要向上传递更新信息, 以便同时更新保证数据的一致性, 所以综合关系是单向关联的, 只能用大比例尺要素查询相邻小比例尺的同名要素。如果需要反向查询, 就打开 FeatureRelation 表, 将 Rels 中的关系释放到表中。如果要提高查询不同比例尺要素的速度, 可在 3 个要素表中分别增加一个比例尺级别字段, 该字段只需 1byte 的存贮空间, 这时 FtrType 字段就不必说明比例尺级别了。

2.5 分区

1) 分区的表示。按照一个分区一个数据库的原则,数据库名、目录名、计算机名(或IP地址)这三者组合起来可确定数据库的物理地址。而作为分区,还要知道数据库在整个尺度空间的位置,因此,还要记录数据库的主导版本层号,以及表示分区范围和位置的最小定界矩形(MBR)。数据库不仅要记录自己的物理地址,还要记录父数据库的物理地址。父数据库的物理地址、本数据库的主导版本层号、MBR及多边形等数据都可放在关于数据库的属性表中。子数据库可用图3中的SimpleFeature表来确定,也可另建分区表。在SimpleFeature表中,可规定 $FtrType = fq$ (分区),表示该要素是下一层主导版本数据库的一个分区子数据库,ReIs中记录该分区子数据库的物理地址,SObj是分区多边形(边界环)的地址码。

2) 父子区之间综合关系的表示。由图1可知,第 i 层分区的第 m 级(最上级)比例尺版本与第 $i-1$ 层的第1级比例尺版本是相邻的比例尺版本,父子区之间的综合关系实际上就是这两个版本之间的综合关系。在图3的FeatureRelation表中,父子区之间的综合关系可规定 $RelType = 12$,其余字段的规定与同一分区的综合关系一样,还是Hst综合成Gst,此时Gst不是本数据库中的要素地址码,而是父数据库中的要素地址码。父数据库的地址在关于数据库的属性表中查找。三个要素表中的ReIs字段记录父子层间的综合关系是关系类型码为12,其余的规定和同区的综合关系一样。例如,在SimpleFeature表中,某个记录的ReIs字段值为12, 37, 0,可解释为本简单要素综合成父数据库的SimpleFeature表中第37号记录的要素。再如,在ComposedFeature表中,某个记录的ReIs字段值为12, 55, 2,可解释为本复合要素综合成父数据库的ComposedFeature表中第55号记录的要素。

3 试验

为了检验前面提出的多比例尺GIS数据模型的正确性和可行性,笔者研制了一个多比例尺GIS软件的原型系统。用黑龙江1:400万、1:75万和1:25万三种比例尺地形图对系统进行调试,检验了方案和模型的正确性和可行性^[7]。限于篇幅,略去软件研制和试验过程,只针对试验过程中出现的一些情况,提出了以下解决办法。

1) 跨区对象之间的联系。在本方案中,可利

用子区对象与父区同名对象之间的联系来确定跨区对象之间的联系。例如,穿过湖北省的长江与穿过湖南省的长江在省级地图中是分区存储的,它们所对的父区(全国地图)对象是同一条长江。

2) 分区边缘操作。在分区的边缘进行显示、网络分析、叠置分析、缓冲区分析等操作时,会同时用到两个以上分区的数据。由于各分区不在同一个数据库中,操作起来很麻烦。可采用工作区的方法来进行边缘操作。具体做法是:开辟一个工作区,将有关数据库都纳入该工作区,在工作区中确定一个工作库,将有关数据库中要用到的数据(边缘数据)映射到工作库中,建立工作库与有关数据库的联系,所有的边缘操作都在工作库中进行。

3) 任意比例尺显示。在屏幕上按任意比例尺显示空间数据时,有两种显示方式:一是用最近的下级比例尺缩小显示;二是用最近的上级比例尺放大显示。无论哪种方式,都会产生图面承载密度不一致的问题。经试验,认为用下级比例尺缩小显示效果稍好。

4) 最上级的比例尺版本操作。由于分级版本与主导版本是同一分区的,我们又要求最上级版本比例尺的数据不经缩小即可在屏幕上全图显示整个分区,显示时会频繁调用不同分区数据库的数据,影响操作速度。操作最上层比例尺版本时,将与当前分区相邻的分区版本都映射到工作数据库中,即可提高速度。

5) 各分区各级比例尺的一致性。同一层各分区的分级比例尺版本必须一致,否则跨区操作将无法进行。不同层的各分级比例尺可不一致。

参 考 文 献

- 1 齐清文,张安定.关于多比例尺GIS数据库多重表达的几个问题研究.地理研究,1999,18(2)
- 2 Oosterom P V, Schenkelaar V. The Development of an Interactive Multi-scale GIS. International Journal of Geographical Information Systems, 1995, 9(5)
- 3 高俊,夏运均,游雄,等.虚拟现实在地形环境仿真中的应用.北京:解放军出版社,1999
- 4 王家耀.空间信息系统原理.北京:科学出版社,2001
- 5 Lodestar S H. An Octree-based Level of Detail Generator for VRML. Siggraph Symposium on VRML, 1997
- 6 王晏民.空间矢量数据的双字节存储法.武汉大学学报·信息科学版,2002,27(增刊):80~83
- 7 王晏民.多比例尺GIS矢量空间数据组织.[博士学位论文].武汉:武汉大学,2002

- 8 龚健雅. 当代GIS的若干理论与技术. 武汉: 武汉测绘科技大学出版社, 1999
- 9 毋河海, 龚健雅. GIS空间数据结构与处理技术. 北京: 测绘出版社, 1997
- 10 王晏民, 李德仁, 龚健雅. 多尺度GIS集中式数据模型. 黑龙江工程学院学报, 2001, 15(1): 19~22

第一作者简介: 王晏民, 教授, 博士. 现主要从事地理信息系统、摄影测量与遥感的研究. 代表成果: 摄影测量与大地测量三维联合平差; 全球定位系统与解析航空摄影测量相结合等. 已发表论文30余篇.

E-mail: wangyanmin@0451.com

A Multi-scale GIS Scheme and Its Data Model

WANG Yanmin¹ LI Deren² GONG Jianya²

(1 Dept. of Surveying and Mapping Engineering Beijing Institute of Architecture Engineering
1 Zhanlan Road, Beijing, China 100083)

(2 State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing,
Wuhan University, 129 Luoyu Road, Wuhan, China, 430079)

Abstract: This paper proposes an applicable multi-scale GIS scheme and realize it with an appropriate spatial data model. The advantages and disadvantages of four typical multi-scale GIS schemes are analyzed. Based on the analysis, a new scheme—layered, leveled and partitioned scheme is proposed. In the scheme, the object district is divided in several scale layers. Then the minimum scale layer is divided into several smaller scale levels. The nearby larger scale layer is partitioned into several sub districts. Similarly, each sub district is divided into smaller levels, and nearby larger layer is partitioned into sub districts, and so on. Each district with its levels is stored in one database. Based on the spatial data model designed by the authors, according to map generalization methods, a multi-scale GIS data model which is suitable for the new scheme is successfully developed. For the purpose of verifying the validity and feasibility of the new scheme and model, a prototype of multi-scale GIS software is written. Experiments prove that the new scheme and model are really valid and feasible.

Key words: multi-scale; GIS; data model; map generalization

About the first author: WANG Yanmin, professor, Ph. D. His major research orientations include GIS, photogrammetry and remote sensing. His major achievements are 3D combined adjustment for photogrammetry and geodesy; aero-photogrammetry combined with GPS etc. He has published more than 30 papers.

E-mail: wangyanmin@0451.com

(责任编辑: 袁丰)

李德仁院士荣获湖北省科学技术突出贡献奖

本刊讯 在7月14日隆重召开的湖北省科学技术奖励大会上, 中共中央政治局委员、湖北省委第一书记俞正声向中国科学院院士、中国工程院院士、武汉大学教授李德仁颁发了本省最高奖——湖北省科学技术突出贡献奖。李院士成为自该奖项设立以来, 获此殊荣的第三位科学家。据悉, 李院士已经决定, 将本次获得的50万元奖金中可以私人支配的10万元, 以及个人另外拿出的10万元, 合计20万元一并捐给武汉大学。