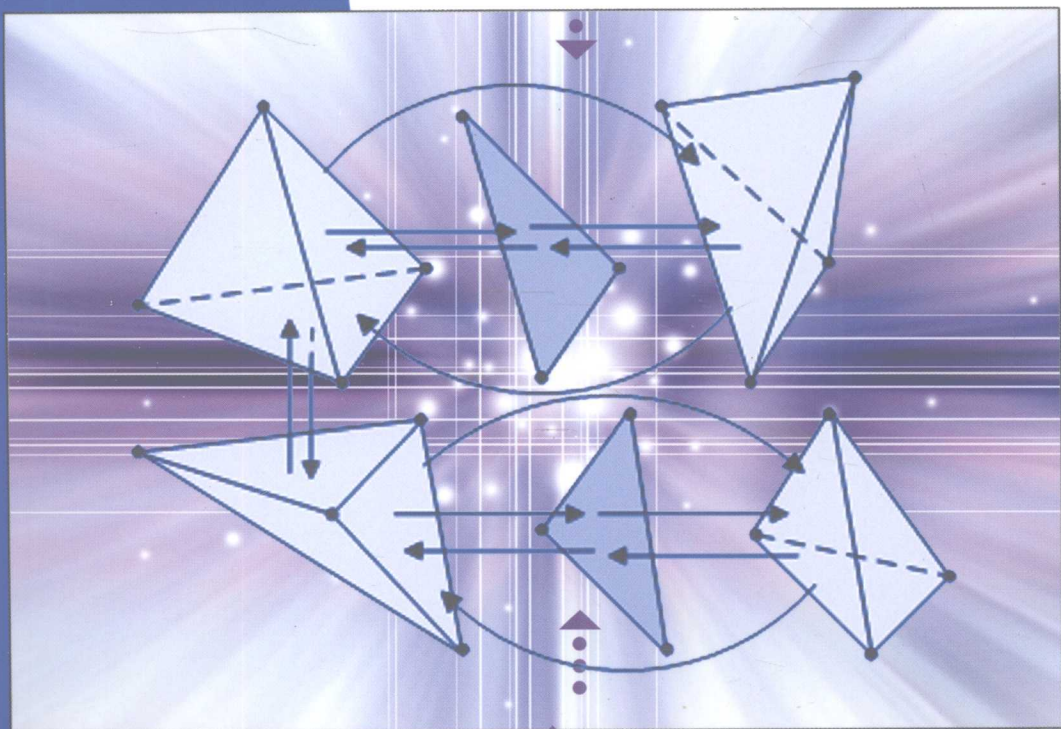


空间数据库

吴信才 编著



科学出版社
www.sciencep.com

空间数据库

吴信才 编著

科学出版社

北京

内 容 简 介

空间数据库是近年来热点研究领域,是一门前沿交叉学科。本书系统阐述空间数据库基本概念、原理、方法及技术的新发展,重点介绍空间现象抽象表达、空间数据模型、空间数据组织与管理、空间数据索引技术、空间数据查询与访问、时态空间数据库、空间数据元数据与空间数据共享、空间数据库设计、空间数据库新技术等内容。

本书内容全面、条理清晰、叙述严谨、实例丰富、针对性强,可作为地理信息系统、遥感、软件工程、测绘工程、通信工程等专业本科生和研究生教材,也可供地质矿产、地理信息、测绘遥感、城市规划、国土管理、环境科学及相关专业研究和开发人员参考和使用。

图书在版编目(CIP)数据

空间数据库/吴信才编著. —北京:科学出版社,2009
ISBN 978-7-03-024587-8

I. 空… II. 吴… III. 地理信息系统 IV. P208

中国版本图书馆CIP数据核字(2009)第075188号

责任编辑:韩鹏 刘希胜/责任校对:张怡君

责任印制:钱玉芬/封面设计:王浩

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>

双青印刷厂印刷

科学出版社发行 各地新华书店经销

*

2009年5月第一版 开本:787×1092 1/16

2009年5月第一次印刷 印张:23 3/4

印数:1—4 000 字数:547 000

定价:59.00元

(如有印装质量问题,我社负责调换〈路通〉)

前 言

地理信息系统以数字形式表达现实世界。空间数据库是地理信息系统的一个重要组成部分和核心技术。空间数据模型可支持现实世界中空间实体的表达及其相互之间的关系,是 GIS 系统进行空间数据组织和空间数据库设计的理论基础。国内外 GIS 研究和发展的实践表明,对空间数据模型的认识和研究在很大程度上决定着 GIS 系统和应用的成败。因此,国际 GIS 学术界和产业界对空间数据库的研究一直十分关注和重视。空间数据库愈来愈受到我国广大地学工作者、信息技术从业者及其他相关学科研究人员高度关注及积极投入。

空间数据库的发展可谓突飞猛进。空间数据组织管理方式从早期文件管理发展到今天的面向对象关系数据库管理,数据模型从传统的三大经典模型发展到今天面向对象模型及面向实体数据模型,空间数据共享和互操作变得越来越容易。作者凭借 20 多年来对地理信息系统的技术研究和实践,对空间数据模型、空间数据库技术悉心研究,取得一些为实践所认可的成果和经验,丰富了空间数据库的理论和实践。作者希望通过本书对前期空间数据库的研究成果进行比较全面的回顾和详细的总结,为初学者提供系统、全面的空间数据库知识,使他们能在尽量短的时间内理解空间数据库原理、方法,为从事 GIS 领域的工作人员和其他涉及空间数据处理的技术人员提供相关方法、技术。

本书在内容和结构上作了精心设计与安排,结构严谨,内容新颖,理论和实践结合紧密。书中大多数实例来自于教育部地理信息系统软件及其应用工程研究中心。本书按照理论部分、实践部分以及新发展部分展开系统论述。在理论部分,本书系统地阐述空间数据库基本概念、基本原理和基本方法,空间现象抽象表达,空间数据模型,空间数据索引技术,空间数据组织与管理,空间数据查询与访问,时态空间数据库,空间数据库设计。在实践部分,本书系统阐述空间数据库建设思路、方法及体系,以武汉市江夏区为例论述了 MapGIS7.0 空间数据库在土地利用规划中的应用。在新发展部分,介绍了分布式空间数据库技术、空间数据仓库技术。

本书相关内容是“十五”国家 863 项目“面向网络海量空间信息的大型 GIS”、“基于大型 GIS 的基础地质空间数据库信息管理系统开发”,国家科技攻关计划项目“城市数字化系统集成关键技术研究”,国家发展和改革委员会项目“面向网络分布式空间信息应用服务支撑平台”等研究成果的凝

练。书中重点阐述了空间数据中心概念、特点、体系结构、开发模式及实际应用，这也是本书的一大亮点。

全书主要参考了美国 Shashi Shekhar、Shanjay Chawla 编写，谢昆青、马修军、杨冬青等译的《空间数据库》，第 5 章主要参考了郭薇、郭菁、胡志勇编写的《空间数据库索引技术》等。此外还参阅和吸收了国内外有关论著的理论和技术成果，书中列出了主要参考文献，在此一并表示衷心的感谢。

本书由长江学者吴信才教授策划、组织编写及统稿。参加本书编著的人员还有：谢忠教授、周顺平教授、刘修国教授、郑贵洲教授、吕建军教授、张发勇副教授、刘福江博士。

由于时间紧，水平有限，难免出现错误和不足，敬请读者提出宝贵意见。

作者

2008 年 6 月

目 录

前言

第 1 章 绪论	1
1.1 空间数据库概述	1
1.2 空间数据库的发展	8
1.3 空间数据库与传统数据库比较.....	13
1.4 新型数据库系统.....	15
思考题	23
第 2 章 空间现象抽象表达	24
2.1 现实世界的认知.....	24
2.2 从现实世界到模型世界.....	28
2.3 空间实体描述.....	31
2.4 空间实体矢量表达.....	36
2.5 空间实体栅格表达.....	42
思考题	50
第 3 章 空间数据模型	51
3.1 空间关系.....	51
3.2 面向对象空间数据模型.....	57
3.3 二维空间对象模型.....	62
3.4 数字表面模型.....	71
3.5 三维空间数据模型.....	74
3.6 网络结构模型.....	89
思考题	91
第 4 章 空间数据组织与管理	92
4.1 文件组织与数据库.....	92
4.2 空间数据管理方式.....	94
4.3 空间数据引擎	100
4.4 空间数据与属性数据的连接	106
4.5 空间数据组织	107
4.6 栅格数据存储和管理	113
思考题.....	116
第 5 章 空间数据索引技术	117
5.1 空间索引的发展	117
5.2 简单网格空间索引	119
5.3 二叉树索引	123

5.4	B 树索引	130
5.5	四叉树索引	145
5.6	可扩展的哈希索引	152
5.7	空间填充曲线	154
	思考题	157
第 6 章	空间数据查询、访问	159
6.1	标准数据库查询语言	159
6.2	空间查询语言	163
6.3	空间查询处理	170
6.4	查询优化	181
	思考题	190
第 7 章	时态空间数据库	191
7.1	地理信息的时空分析	191
7.2	时态数据库	203
7.3	时空数据模型	206
7.4	时空数据库	212
	思考题	216
第 8 章	空间数据元数据与空间数据共享	217
8.1	空间数据元数据	217
8.2	空间数据共享	229
	思考题	238
第 9 章	空间数据库设计	239
9.1	空间数据库设计概述	239
9.2	需求分析	243
9.3	概念结构设计	249
9.4	逻辑结构设计	260
9.5	空间数据库物理设计	272
9.6	空间数据库的实施和维护	281
9.7	空间数据库建库	283
9.8	武汉市江夏区土地利用规划空间数据库设计	294
	思考题	300
第 10 章	空间数据库新发展	301
10.1	分布式空间数据库技术	301
10.2	空间数据仓库技术	307
10.3	数据中心	313
	思考题	323
第 11 章	MapGIS7.0 空间数据库	324
11.1	概述	324
11.2	MapGIS 空间数据模型	324

11.3 MapGIS 空间数据管理	331
11.4 MapGIS 时空数据管理	336
11.5 MapGIS 长事务处理	340
思考题	343
参考文献	344
附录 武汉市江夏区土地利用规划空间数据库结构表	348

第1章 绪论

1.1 空间数据库概述

1.1.1 空间数据库基本概念

1. 数据

数据是指客观事物的属性、数量、位置及其相互关系等的符号描述。

在空间数据库中，数据可以是一个数，如某一点的高程值、一个多边形的面积等，也可以是一组符号组成的字符串，如一个地名、一个河流注记、一幅图像。空间数据库中的数据大多与地理位置有关，一般称为空间数据（Spatial Data）。空间数据不同于普通的数据，它具有空间性、时间性、多维性和大数据量等特点，而且数据之间不仅有传统的关联关系，更多的还有空间关系，这就给数据的处理和利用带来了更多的难度。

2. 空间

空间（Space）是一个应用很广泛的名词。空间是客观存在的物质空间，是一个复杂的概念，具有多义性，既有与时间对应的含义，也有“宇宙空间”的含义。空间可以定义为一系列结构化物体及其相互间联系的集合。日常语义上的“空间”是指事物之间的距离或间隔。从感观角度将空间看作是目标或物体所存在的容器或框架，因此空间更倾向于被理解为物理空间。空间知识的本质问题是一个古老的研究领域，关于空间的概念，不同的学科有不同的解释。哲学家、天文学家、物理学家对空间的论述众说纷纭。从物理学角度看，空间为宇宙在三个相互垂直的方向上所具有的广延性；天文学中的空间指宇宙空间，是时空连续体系的一部分；行星物理和相关地球物理中的空间常指地球高层空间和行星际空间；在数学中空间概念的范围很广，一般指某种对象（现象、状况、图形、函数等）的任意集合，其中要求说明“距离”或“邻域”的概念；从地理学的意义上讲，空间是人类赖以生存的地球表层具有一定厚度的连续空间域，是一个定义在地球表层空间实体集上的关系。GIS领域的空间是指地理环境或地球表层空间，是地理信息系统表达和研究的对象。为了在GIS中对地理空间（Geospace）进行描述，常常需要借助于抽象的数学空间表达方法。

3. 地理空间

地球表面上的一切地理现象、地理事件、地理效应、地理过程统统都发生在以地理空间为背景的基础之上。地理学的空间是一个定义地球表层空间实体集上的关系，在空间实体之间有无数种关系，定义一种关系就自然定义了一种空间，而这个空间又是和几何关系联系在一起的，并且，几何关系是所有这些关系中的基础关系。地理空间是一个

相对空间,是一个空间实体组合排列集(这些空间实体具有精确的空间位置),强调宏观的空间分布和空间实体间的相关关系(关系以各单个地理空间实体为联结的结点或载体)。地理空间是指空间参考信息的地理实体或地理现象发生的时空位置集。依附地理空间存在着各种事物或现象,它们可能是物质的,也可能是非物质的,与一定的地理空间位置有关,都具有一定的几何形态。在地理空间中,物体不仅反映事物和现象的地理本质内涵,而且反映它们在地理空间中的位置、分布状况以及它们之间的相互关系。地理空间十分复杂,其各组成部分之间存在内在联系,形成一个不可分割的统一整体,而且地理空间具有等级差别,同等级地理空间之间亦存在差异。

地理空间若想精确定位于地球上,还必须承认它有欧氏空间基础,有相对于地球坐标系的绝对位置。GIS中的地理空间是指经过投影变换后,在笛卡儿坐标系中的地球表层特征空间。它是地理空间的抽象表达,是信息世界层面的地理空间。地理空间由地理空间定位框架及其所连接的地理空间特征实体组成。通过地图投影,地理现象的宏观特性和空间位置的精确特征紧密有机地联系在一起。其中地理空间定位框架即大地测量控制,为建立所有地理数据的坐标位置提供通用参考系统,将所有地理要素同平面及高程坐标系连接。地理空间特征实体则为具有形状、属性和时序性的空间对象。地理空间的数学描述可以表达为 $S=(O, R)$,其中设 E_1, E_2, \dots, E_n 为 n 个不同类的地理空间实体, R 表示地理空间实体间的相互联系、相互制约关系, $O=\{E_1, E_2, \dots, E_n\}$ 表示地理空间中各个组成部分(实体)的集合。也就是说,可以简单地将地理空间理解为一个空间目标组合排列集,其每个目标都有具体位置、属性和时间信息以及与其他对象的拓扑关系和语义关系。

4. 空间数据

空间数据是对空间事物的描述,空间数据实质上就是指以地球表面空间位置为参照,用来描述空间实体的位置、形状、大小及其分布特征诸多方面信息的数据。空间数据是一种带有空间坐标的数据,包括文字、数字、图形、影像、声音等多种方式。空间数据是对现实世界中空间特征和过程的抽象表达,用来描述现实世界的目标,记录地理空间对象的位置、拓扑关系、几何特征和时间特征。位置特征和拓扑特征是空间数据特有的特征。此外,空间数据还具有定位、定性、时间、空间关系等特性。郭仁忠(1997)认为空间数据由关于空间对象形态和位域的几何数据、关于空间对象主题的属性数据、关于空间位域上的空间变量的统计数据或模型参数三个部分组成。

由于现实世界的复杂性、模糊性和不确定性,以及人类认识世界和描述世界的能力的局限性,用空间数据来描述和表达地理空间实体时会产生误差,空间数据只能从有限的方面去描述事物,反映事物的某些特性,在一定程度上接近真实值,不可能也没必要全面、详尽、逼真地复制事物本身。另外,空间数据是以如人工统计、仪器测量、社会调查等多种方式获得的,就必然产生各种误差,如人为差错、仪器的系统误差等。空间数据在GIS中是采用笛卡儿坐标系表达的,在计算机中只能用离散的方式来表示空间物体,与自然界真实的地物是不一样的,不能达到百分之百的精确。

空间数据是数字地球的基础信息,数字地球功能的绝大部分都以空间数据为基础。当前空间数据已广泛应用于社会各行业、各部门,如矿产勘查、资源清查、土地利用、

城市规划、环境管理、航空航天、交通、旅游、军事等。随着科学和社会的发展，人们已经越来越认识到空间数据对于社会经济发展、人们生活水平提高的重要性，这也加快了人们获取和应用空间数据的步伐。

5. 数据库

数据库技术是 20 世纪 60 年代中期开始发展起来的一门数据管理自动化的综合性新技术，是计算机科学的重要分支。数据库的英文是 Database，其意义为数据基地，即统一存储和集中管理数据的基地。有些类似资料库，实际上资料库的许多特征都可以从数据库中找到。在资料库中，各类资料都有严格的分类系统和编码表，并存放在规定的资料架上，为管理和查找资料提供了极大的方便。当资料的数据形式存放于计算机时，它已经失去直观性，更需要建立严密的分类和编码系统，实现数据的标准化和规范化。数据库是长期储存在计算机内的、有组织的、可共享的数据集合。数据库中的数据按一定的数据模型组织、描述和储存，具有较小的冗余度、较高的数据独立性和易扩展性，并可为各种用户所共享（萨师焯等，2000）。数据库是数据管理的高级阶段。作为文件之上的更高层次的数据组织，数据库并不是众多文件的简单集合，数据库中的文件及其内的数据是有内在联系的；数据库系统透过文件，全面有效地实施数据管理。

数据库的基本结构如图 1-1 所示。最中央为存储的数据体（亦称库存数据），其外为数据管理程序，实施对数据的管理、维护和操作。然后，通过各种应用程序来实现它们与外部的联系，各种输入和输出介质便是实现这种联系的工具。

数据库的应用领域相当广泛，从一般事务处理，到各种专门化数据的存储与管理，都可以建立不同类型的数据库。

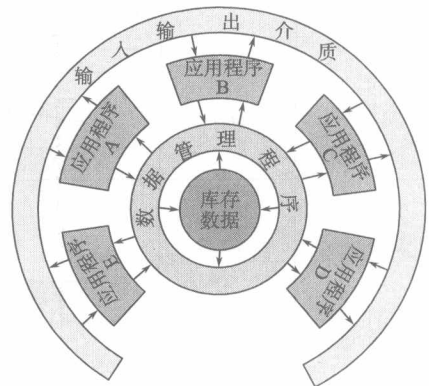


图 1-1 数据库的基本结构

6. 空间数据库

空间数据库是描述与特定空间位置有关的真实世界对象的数据集合，在此，我们把些对象称为空间参考对象。任何真实世界的对象可能表示成数据库中的对象，但并不是任何对象都和地理位置有关，这取决于我们所要表达的信息模型及应用。只有当对象在数据库中需要考虑其空间位置时，它们才成为空间参考对象，才和空间位置相关。空间数据库既要能处理空间参考对象类型，也要能处理非空间参考对象类型。而如何表示空间或地理现象即空间参考对象的关键是其数据模型。数据模型的设计除与应用有关外，还与提供支持模型的基本概念、方法等有密切联系。空间数据的表示则与计算机表示数据的精度和计算机的存储空间有关。

1.1.2 空间数据类型

空间数据是 GIS 的核心，也有人称它是 GIS 的血液，因为 GIS 的操作对象是空间数据，因此设计和使用 GIS 的第一步工作就是根据系统的功能，获取所需要的空间数据，并创建空间数据库。

GIS 中的数据来源和数据类型繁多，概括起来主要有以下几种类型，如图 1-2 所示。

	点	线	面
类型数据			
面域数据	点状要素 	线状要素 	面状要素
网络数据	区域中心 	境界线 	行政单元
样本数据	气象站 	航线 	样方分布区
曲面数据	高程点 	等值线 	概略等值区
文本数据	地点名称 	线状要素名称 	区域名称
符号数据	点状符号 	线状符号 	面状符号

图 1-2 空间数据类型 (黄杏元等, 2001)

(1) 地图数据。来源于各种类型的普通地图和专题地图，这些地图的内容丰富，图上实体间的空间关系直观，实体的类别或属性清晰，实测地形图还具有很高的精度。

(2) 影像数据。主要来源于卫星遥感和航空遥感，包括多平台、多层面、多种传感器、多时相、多光谱、多角度和多种分辨率的遥感影像数据，构成多源海量数据，也是

GIS 的最有效的数据源之一。

(3) 地形数据。来源于地形等高线图的数字化，包括已建立的数字高程模型 (DEM) 和其他实测的地形数据等。

(4) 属性数据。来源于各类调查报告、实测数据、文献资料、解译信息等。空间数据根据表示对象的不同，又具体分为七种类型，它们各表示的具体内容如下：①类型数据，如考古地点、道路线、土壤类型的分布等。②面域数据，如随机多边形的中心点、行政区域界线、行政单元等。③网络数据，如道路交点、街道、街区等。④样本数据，如气象站、航线、野外样方分布区等。⑤曲面数据，如高程点、等高线、等值区域等。⑥文本数据，如地名、河流名称、区域名称等。⑦符号数据，如点状符号、线状符号和面状符号 (晕线)。

1.1.3 空间数据特征

1. 时空特征

通过大量研究，认为地学信息具有多维结构，如图 1-3 所示，一般由空间、属性和时间三部分构成，那么对其进行建模就是要对空间、属性和时间三者的提问域和答案域都要作出回答。

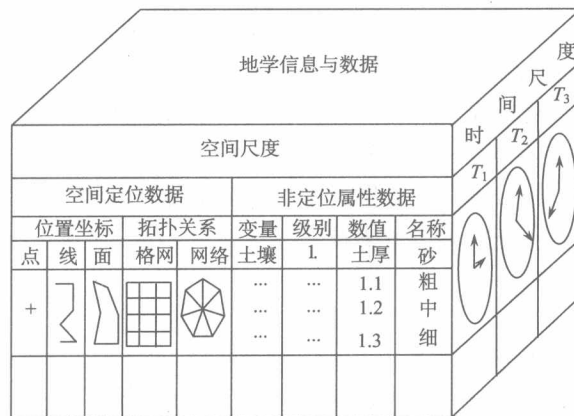


图 1-3 空间数据多维结构 [据黄杏元 (2001) 修改]

空间数据的空间特性是指空间实体的空间位置及其与其他空间实体的空间关系，指明地物在地理空间中的位置，用于回答“Where”提问。空间位置可以用绝对空间位置和相对空间位置来表示。绝对空间位置用来表示地物本身的地理位置，通常用笛卡儿坐标、地理经纬坐标、空间直角坐标 (x, y, z) 、平面直角坐标和极坐标等表示地理空间实体在一定的坐标参考系中的空间位置。相对空间位置用来表示多个地物之间的位置相互关系，通过距离和方向描述相对于其他参照系或地物的空间位置，如某观测站位于某高地 135° 方向 500m 处。空间关系则是地理空间实体之间存在的一些具有空间特性的关系，如拓扑关系、顺序关系和度量关系等。空间特征是空间数据最基本的特征，空间数据记录了地理空间实体对象的空间分布位置和几何形状等空间信息，所以在使用空间

数据时首先需要考虑的就是空间特征。

属性特性是指地学现象的数量、质量和分类等属性信息，包括用来描述地物的自然或人文属性的定性或定量指标的成分。用于回答“*What*”和“*How*”提问。可以用名义量、顺序量、间隔或比率来表示。例如，表述一个城镇居民点，若仅有位置坐标(x, y)，那只是一个几何点，要构成居民点的地理空间数据，还需要经济（人口、产值等）、社会（就业率等）、资源和环境（污染指数等）等属性数据。

时态特性指地理数据采集或地理现象发生的时刻或时段，这部分数据称为时态特征数据或时态数据。同一地物的多时段数据，可以动态地表现该地物的发展变化。时态特征数据可以按时间尺度划分为短期（如地震、洪水、霜冻）、中期（如土地利用、作物估产）、长期（如城市化、水土流失）和超长期（如地壳变动、气候变化）等类型。

2. 多维特征

地理空间数据具有多维结构的特征。地理空间实体或地理现象本身具有各种性质，空间目标的属性特征也称为主题、专题，是与地理空间实体相联系的、具有地理意义的数据和变量，用于表达实体本质特征和对实体的语义定义，一般分为定性（如类型、名称、特征值）和定量（如数量、等级）两种。属性之间的相关关系则反映实体间的分类分级语义关系，主要体现为属性多级分类体系中的从属关系、聚类关系和相关关系。地理空间数据不仅能描述空间三维和时间维，也可以表现空间目标的属性以及数据不同的测量方法、不同来源、不同载体等多维信息，实现多专题的信息记录。例如，在一个坐标位置上，既包括地理位置、海拔高度、气候、地貌和土壤等自然地理特征，也具有相应的社会经济信息如行政界线、人口、产量、交通等。此外，一些空间对象或地理目标（如河流）同时又作为其他空间目标的分界线，也是空间数据多重属性的表现。在进行空间数据分析过程中，要重视并充分考虑地理空间数据的多维结构及其对空间关系的影响，为地理系统的综合研究提供技术支持。

3. 多尺度性

尺度是空间数据的重要特征之一。空间数据的多尺度特征可从空间多尺度和时间多尺度两个方面进行理解。正是由于空间数据具有多尺度特征，导致空间数据的综合难度加大，不利于数据管理和共享。

地球系统是由各种不同级别子系统组成的复杂巨系统，各个级别的子系统在空间规模和时间长短方面存在很大差异，而且由于空间认知水平、精度和比例尺等的不同，地理实体的表现形式也不相同，因此多尺度性成为地理空间数据的重要特征。在空间数据中多尺度特征包括空间多尺度和时间多尺度两个方面。空间多尺度是指空间范围大小或地球系统中各部分规模的大小，可分为不同的层次，时间多尺度指的是地学过程或地理特征有一定的自然节律性，其时间周期长短不一。空间多尺度特征表现在数据综合上，数据综合类似于数据抽象或制图概括，是指数据根据其表达内容的规律性、相关性和数据自身规则，可以由相同的数据源形成再现不同尺度规律的数据，它包括空间特征和属性的相应变化。多尺度的地理空间数据反映了地球空间现象及实体在不同时间和空间尺度下具有的不同形态、结构和细节层次，应用于宏观、中观和微观各层次的空间建模和

分析应用。

4. 海量数据特征

GIS 地理空间数据的数据量极大。它既有空间特征（地学过程或现象的位置与相互关系），又有属性特征（地学过程或现象的特征）。空间数据不仅数据源丰富多样（如航天航空遥感、基础与专业地图和各种经济社会统计数据），而且更新快，空间分辨率不断提高。随着对地观测计划的不断发展，每天可以获得上万亿兆字节的关于地球资源、环境特征的数据，使对海量空间数据组织、处理和分析成为目前 GIS 亟待解决的问题之一。

空间数据量是巨大的，通常称海量数据。之所以称为海量数据，是指它的数据量比一般的通用数据库要大得多。一个城市地理信息系统的数据量可能达几十吉字节，如果考虑影像数据的存储，可能达几百吉字节。这样的数据量在其他数据库中是很少见的。地理信息系统的海量数据，带来了系统运转、数据组织与储存、网络传输等一系列技术困难，自然也给数据管理增加了难度。正因为空间数据量大，所以需要在二维空间上划分块或者图幅，在垂直方向上划分层来进行组织。

1.1.4 空间数据库作用

随着遥感、GIS、计算机技术、通信技术等技术的发展，空间数据库技术不断走向成熟，目前已经在不同行业中得到广泛的应用。空间数据库技术的发展带动了我国空间信息产业化发展。根据统计，我国信息领域的投资有 30% 左右为地理信息获取、加工处理及应用方面的投资。一批国家级的地理数据库已经建立，包括国家基础地理信息系统 1:100 万和 1:25 万数据库，海洋信息相关的资源、环境、灾害等的数据库，气候气象数据库，环境信息监测数据库，矿产资源数据库，1:50 万土地利用数据库，1:10 万土地资源数据库等。空间数据库是地理信息系统中空间数据的存储场所。在一个项目的工作过程中，空间数据库发挥着核心的作用。

1. 空间数据处理与更新

地理信息数据一般时效性非常强，因此就要求人们不断更新数据库。空间数据更新是通过空间信息服务平台用现势性强的现状数据或变更数据更新数据中非现势性的数据，达到保持现状数据库中空间信息的现势性和准确性或提高数据精度；同时将被更新的数据存入历史数据库供查询检索、时间分析、历史状态恢复等（张新长，2005）。因此空间数据的更新并不是简单的删除替换。这其中又涉及数据的整体更新、局部更新、采集途径、时效性、保持原有数据的不变、更新数据与原有数据正确连接等多方面问题，这些都是空间数据库发展中亟待解决的问题之一。

2. 海量数据存储与管理

地理数据涉及地球表面信息、地质信息、大气信息等多种极其复杂的信息，描述信息的数据量十分巨大，容量通常达到吉字节级。空间数据库的数据量远远大于一般数据

库的数据量。空间数据库的布局和存取能力对地理信息系统功能的实现和工作的效率影响极大。如果在组织的所有工作地点都能很容易地存取各种数据,则能使地理信息系统快速响应组织内决策人员的要求;反之,就往往会妨碍地理信息系统的快速响应。空间数据库为空间数据的管理提供了便利,解决了数据冗余问题,大大加快了访问速度,防止了由于数据量过大而引起的系统“瘫痪”等。它可以充分利用 RDBMS 安全用户管理、数据备份等功能,实现空间数据和属性数据真正的无缝连接,提高数据管理和应用效率,便于数据共享,也为 GIS 采用完全的 C/S 模式提供了基础。

3. 空间分析与决策

空间数据库技术不仅实现了在 DBMS 中存储空间数据的目的,而且能够支持空间数据的结构化查询和分析,可以高效地把这些空间信息在 GIS 软件的工作空间中复原出来。空间数据库,作为源数据库,可通过对原始数据进行日常操作性的应用,提供简单的空间查询和分析。用户在决策过程中,通过访问空间数据库获得空间数据,在决策过程完成后再将决策结果存储到空间数据库中。如果获取空间数据很困难,就不可能进行及时的决策,或者只能根据不完整的空间数据进行决策,其结果都可能导致地理信息系统不能得出正确的决策结果。可见,空间数据库在地理信息系统中的重要性是不言而喻的。空间数据仓库则根据主题内容通过专业模型对不同的空间数据库中的原始业务数据进行抽取和聚集,给用户集成提供面向主题分析的决策支持环境。空间数据仓库根据用户的需求来组织源信息和提供数据,侧重于综合分析,而一般数据库的用户只能根据数据库中现存的数据来选择所需数据,侧重于一般性的数据处理。

4. 空间信息交换与共享

空间数据库的应用范围非常广,对于不同的用户群,其要求和使用方式以及所需数据也非常不同。虽然根据不同用户要求,空间数据库系统选用不同的专题地理信息数据库和不同的数据模型,但是随着网络技术的发展,空间数据库系统能够支持网络功能,使得信息的交流与共享变得更加便捷,较好地解决了海量地理信息存储的不便,大大扩展了空间信息的共享范围。相对于各行各业对空间信息技术的需求,空间信息产业的开发与应用突飞猛进,空间数据库技术作为一种综合的新技术、新方法已延伸到国民经济建设的各个领域。借助空间数据库系统,空间信息的应用范围更加广泛,实效性更能得到保障,准确性得以提高,信息的共享程度能得到加强。

1.2 空间数据库的发展

1.2.1 空间数据库的发展历史

空间数据库系统是一个存储空间和非空间数据的数据库系统,在它的数据库模型和查询语言中能提供空间数据类型,可以进行空间动态索引,并提供空间查询和空间分析的能力。空间数据库技术经历了多年发展和演变,大体经历了以下几个发展阶段。

1. 20 世纪 50 年代后期到 60 年代中期

20 世纪 50 年代后期到 60 年代中期, 由于应用领域拓宽, 计算机不仅用于科学计算还大量用于数据管理。这一阶段的数据管理采用文件管理系统。计算机软件的操作系统中已经有了专门的管理数据软件, 即所谓的文件系统。文件系统的处理方式不仅有文件批处理, 而且还能够联机实时处理, 在这种背景下, 数据管理的系统规模、管理技术和水平都有了较大幅度的发展。在这种管理方式中, 文件管理系统是操作系统的一部分, 是通用的文件管理, 而不是专门的数据管理软件, 空间数据依然保留自身的文件格式, GIS 平台负责响应不同文件格式的空间数据请求, 对于流行的空间数据格式, GIS 平台都能支持。因此基于文件方式的空间数据库的建立就是空间数据目录下空间数据文件的组织。

2. 20 世纪 70 年代

对于空间数据库的研究, 在 20 世纪 70 年代随着计算机辅助制图和遥感图像处理的研究就已经开始, 但在空间数据管理的发展历史中, 空间数据的数据库管理也有很多实验。最初 GIS 数据管理是以文件管理的方式, 所有的数据都存储在自行定义的数据结构与操纵工具的文件中, 人们尝试将空间数据进行数据库管理的方法是将空间数据中的点、线、面分别存储和管理, 点可以进行结构化数据管理, 线和面用相邻两点进行结构化数据管理。这样能够完成对空间数据的数据库管理, 但效率十分低下。这种方式不利于空间数据管理和共享, 空间数据库技术处于初期阶段, 技术上不成熟。虽然数据库技术在 20 世纪 80 年代已经广泛使用, 对文本数值型数据管理方面已经比较成熟, 但是对于结构复杂、数据量庞大、具有拓扑关系的空间数据还是无能为力的。

3. 20 世纪 80 年代

20 世纪 80 年代, 随着数据库管理系统和 GIS 技术的发展, 在关系数据库技术走向成熟、应用迅速扩展的形势下, 以 ESRI 公司的 Arc/Info 为代表的矢量 GIS 技术, 基于关系数据库技术理念, 并部分直接地采用关系数据库技术, 提出地理空间数据管理的数据模型——地学关系模型 (Geo-relational Model), 成功地开发出基于这种模型的矢量 GIS 数据库系统。关系数据库与图形文件采用混合管理方式。图形数据一般采用文件 (如 Arc/Info Coverage) 或者专门图库 (如 Arc/Info ArcStorm) 的形式进行管理, 图形文件存储点、线、面图形要素, 主要采用拓扑结构编码存储。属性数据利用关系数据库管理系统 (RDBMS) 的关系表来存储, 并通过唯一的标识符 (ID 号) 建立空间数据和属性数据间的关联。这种文件-数据库混合管理方式在已有的应用系统中取得了巨大成功, 并取得应用上的成功。这是地理空间数据库系统技术发展史上第一次革命性飞跃, 矢量 GIS 数据库技术由此开始自成体系, 从那时起, 不同的 (矢量) GIS 厂家开发了多种具体的 GIS 数据库系统, 也取得一定的进步, 如 MapInfo 公司的 MapInfo、Intergraph 公司的 MGE 等基本上都支持工业标准的关系数据库。基于地学关系模型的地理空间数据库系统至今仍然应用于大多数 GIS 系统中。但这种混合管理方式弊端是明显的, 一方面, 它增加了数据维护和管理的难度, 特别是对以文件形式存在的空间数