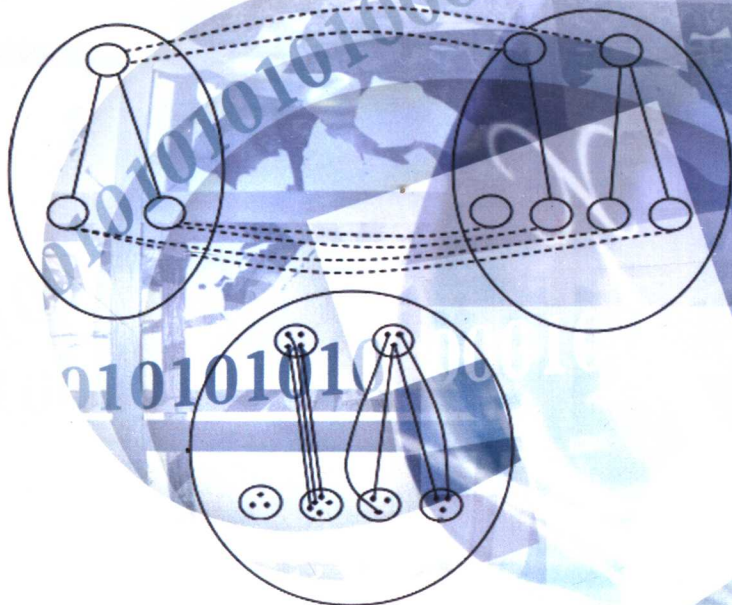





地理信息系统理论与应用丛书

地理空间数据库原理

● 崔铁军 编著



 科学出版社
www.sciencep.com

内 容 简 介

本书全面、系统地论述了地理空间数据库的基本概念、原理与方法,涉及地理空间数据库各个方面的主要内容。全书共分 11 章,内容包括:地理空间数据管理的发展过程和趋势、地理空间实体的计算机表示方法、基本数据结构、地理空间数据的物理组织、空间索引方法、空间数据模型、数据库体系结构、关系数据库接口技术、空间数据库引擎(SDE)、地理空间数据库管理系统、空间查询语言、地理空间数据库设计方法、地理空间数据库工程建立方法、资料收集和处理、空间数据获取及质量评价、地理空间数据仓库、元数据和空间数据互操作等。本书内容丰富、组织严谨,原理和方法结合密切,丰富的图表和应用实例便于读者自学。

本书既可作为高等院校地理信息系统专业或相关专业本科生和研究生的教材,也可供从事信息化建设、信息系统开发等有关科研、企事业单位的科技工作者阅读参考。

图书在版编目(CIP)数据

地理空间数据库原理/崔铁军编著. —北京:科学出版社,2007

(地理信息系统理论与应用丛书)

ISBN 978-7-03-018800-7

I. 地… II. 崔… III. 地理信息系统:数据库系统 IV. P208
TP311.13

中国版本图书馆 CIP 数据核字(2007)第 042930 号

责任编辑:韩 鹏 李久进/责任校对:邹慧卿

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>

源海印刷有限责任公司印刷

科学出版社发行 各地新华书店经销

*

2007 年 4 月第 一 版 开本:787×1092 1/16

2007 年 4 月第一次印刷 印张:23

印数:1—4 000 字数:523 000

定价:59.00 元

(如有印装质量问题,我社负责调换〈环伟〉)

序

自古以来，地图都是地理空间信息的主要载体和传播工具。随着社会发展和科学技术进步，人们认知的地理空间的范围越来越大，对地图品种和数量的需求与日俱增，作为地理空间信息载体的地图的存储、管理与分发成为突出的问题。

对于纸质的模拟地图，通过地图库存储与管理，这种方式一直沿用至今。

目前，以地理空间数据库、计算机地图制图和计算机网络技术为支撑的数字化地图制图已经取代了传统的手工地图制图，并正向以地理空间信息服务为核心的信息化数字化地图制图和地理信息系统转变，地图制图的思想观念、技术手段、产品形式和服务方式等都在发生深刻的变化，地理空间数据的存储与管理也一直是业界关注的问题。

地理空间数据的管理技术经历了许多变化，地理空间数据从由文件管理发展到用数据库管理可以追溯到 20 世纪 60 年代中期，到目前为止，它已从第一代的层次、网状数据库系统，第二代的关系数据库系统发展到第三代的以面向对象为主要特征的新一代数据库系统。与此同时，空间数据仓库技术也备受重视，它作为一种支持决策过程的、面向主题的、集成的、稳定的、不同时间的地理空间数据的集合，是一个多种异构数据源在单个站点以统一的模式组织的存储，其根本目的是服务于决策支持。

地理空间数据库具有地理空间信息科学技术、计算机科学技术等多学科交叉的特点，地理空间数据库的研究必须具备扎实、雄厚的信息科学技术知识和很强的地理空间抽象能力及丰富的工程实践经验。崔铁军教授于 1979~1983 年在解放军测绘学院专修计算机地图制图，毕业后留校任教，后在德国联邦国防军大学获工学博士学位，1998 年进解放军测绘学院博士后科研工作站，长期从事地理空间数据库和地理信息系统理论、方法和技术的科研、教学和工程实践，在地理空间数据库方面有较深的造诣和丰富的实践经验，为撰写《地理空间数据库原理》一书奠定了坚实的基础。

该书内容丰富，既包含了数据库的基本理论，又扩展到了地理空间信息范畴；既把重点放在地理空间数据库，又扩展到了空间数据仓库；既突出了地理空间数据库的核心——空间数据模型，又把地理空间数据库的设计和建立放在了重要位置；既论述了地理空间数据库的体系结构，又介绍了关系数据库的接口技术和空间数据库引擎，是一部完整地论述地理空间数据库的好书，值得一读。有幸，我作为第一个读者，受益匪浅。

目前，我国有关地理空间数据库方面的书还太少，真诚地期待着有更多的青年学者撰写这类著作，共同繁荣地理空间信息科学这块前程似锦的园地。

中国工程院院士 王家耀

2006 年夏

前 言

地理空间数据库技术是地理信息系统数据组织的核心技术，也是地理科学、测绘科学、计算机科学和信息科学相结合的产物。地理空间数据库技术已经代替传统的文件管理方式，逐步成为地理空间数据管理的主流技术。由于地理空间数据的特殊复杂性，地理空间数据管理在为计算机和信息科学做贡献的同时，也如饥似渴地汲取计算机主流技术的各项最新成果，成为计算机科技领域中应用研究技术内容最丰富的分支之一。

地理空间数据库的主要任务是研究地理空间物体的计算机数据表示方法、数据模型以及在计算机内的数据存储结构和建立空间索引方法，如何以最小的代价高效地存储和处理空间数据，正确维护空间数据的现实性、一致性和完整性，为用户提供现实性好、准确性高、完备、开放和易用的地理空间数据。地理空间数据库是理论性和实践性很强的学科，理解起来也非常抽象。帮助有兴趣的读者更好地了解地理空间数据库的基本概念和建设方法，是作者编写本书的用心所在。

多年来，撰写一部有关地理空间数据库原理的书是作者的愿望，但由于现代科技高速发展，使地理空间数据库内容也更新很快，章节很难固定下来，加上人到中年工作繁忙，能专注写书的时间有限，而且由于知识积累与资料准备等原因，一直未能如愿。

作者长期致力于地理空间数据库理论、技术和方法的研究，工程实践，教学和研究生培养等工作。从1987~1992年为地图数据库，特别是1:25万地图数据库技术方案论证做了大量工作，主要负责地图数据库管理系统的结构设计、功能模块划分和部分软件的编写及总体系统调试工作。1998年在博士后工作期间，承担大型地理信息系统研制任务，负责系统的需求分析、总体设计和详细设计等工作。近几年，先后参加和主持了多项国家和军队科研课题，从底层自主设计与开发了基于Oracle关系数据库的空间数据库管理系统，实现了对海量多源空间数据的管理与综合应用，主要功能有空间数据库定义，空间数据的录入、编辑与处理，检索与查询，可视化，备份和恢复等。所取得的研究成果，为本书的撰写奠定了坚实的基础。因此，可以说，这本书也是上述工作的结晶。

本书是在作者所讲授地图数据库课程内容的基础上编著而成。第3章基本数据结构主要参考了韩丽斌教授编著的《地图数据库原理与技术》(1994年，内部教材)；第5章地理空间数据模型和第6章空间数据库体系结构的部分内容参考了毋河海教授编著的《地图数据库系统》。其他章节也参阅、吸收了国内外有关论著的理论和技术成果，书末仅列出了部分参考文献，按出版社的要求，未公开出版的文献没有列在书后参考文献中，而在正文当页下方作脚注，这里向所有参考文献作者致谢！

在本书撰写过程中，研究生邹方磊、和万礼和张威等协助完成了初稿校对等工作。陈应东副教授和郭健副教授等提出了宝贵的意见。在此，作者向他们表示衷心的感谢。

值此成书之际，作者要感谢导师王家耀院士、Kurt Brunner教授、刘家豪教授、刘光运教授、韩丽斌教授和已故杨启和教授的培养教育；感谢解放军信息工程大学测绘

学院训练部和地图学与地理信息工程系的几任领导的一贯支持；感谢地图数据库教研组董延春、郭黎、姚慧敏、吴正升和张斌等教师和历届博士生、硕士生在地空间数据库研究方面所作出的不懈努力。本书的撰写得到科学出版社朱海燕和韩鹏编辑的热情指导和帮助，在此表示衷心的感谢。

本书内容横跨多个学科，加之作者水平所限，书中定有不少疏漏谬误之处，恳请读者与专家们批评指正。

作者

2006年7月于郑州

目 录

序 前言

第 1 章 地理空间数据库导论	1
1.1 地理空间数据	3
1.2 空间数据管理演变过程	6
1.3 空间数据库系统.....	12
1.4 地理空间数据库系统与其他课程的关系.....	15
1.5 地理空间数据库的研究内容与发展趋势.....	17
第 2 章 地理空间现象计算机表达	20
2.1 空间实体及地图表示.....	20
2.1.1 实体的维数和延展度	20
2.1.2 空间变量和空间实体的属性	20
2.1.3 空间实体的地图表示	21
2.2 空间实体的数据描述.....	22
2.2.1 空间实体的数据抽象	22
2.2.2 基于实体对象的描述	23
2.2.3 基于场的描述	26
2.3 空间实体矢量数据表示.....	31
2.3.1 空间实体的几何表示	32
2.3.2 空间实体的属性描述	35
2.3.3 地理空间关系的表示	38
2.4 空间实体的栅格表示.....	39
2.4.1 栅格格式及其结构	41
2.4.2 栅格数据编码方法	43
2.4.3 栅格数据的操作	49
2.5 矢栅结构的比较及转换算法.....	51
2.5.1 栅格结构与矢量结构的比较.....	51
2.5.2 相互转换算法	52
2.6 空间数据的基本特性.....	57
第 3 章 基本数据结构	59
3.1 线性表结构.....	59
3.1.1 线性表	59
3.1.2 栈和队列.....	60
3.1.3 数组	61

3.2	链表	61
3.2.1	线性链表	61
3.2.2	循环链表	63
3.2.3	双重链表	63
3.3	串	63
3.4	树	64
3.4.1	树	64
3.4.2	二叉树	65
3.4.3	线索树	67
3.4.4	树的二叉树表示	70
3.4.5	树的应用	70
3.5	图	71
3.5.1	基本概念	71
3.5.2	图的存储结构	73
3.5.3	图的运算	75
第4章	空间数据的物理组织	79
4.1	文件组织的基本概念	79
4.1.1	操作系统的文件管理	79
4.1.2	逻辑记录与物理记录	80
4.1.3	地址与指针	81
4.1.4	分页与系统缓冲区	82
4.1.5	文件组织	83
4.1.6	动态存储管理	84
4.2	流水文件	86
4.3	顺序文件	86
4.3.1	如何确定关键字值的顺序	87
4.3.2	顺序文件的存储组织	87
4.3.3	顺序文件的查找	87
4.3.4	顺序文件的维护	89
4.4	索引文件	89
4.4.1	索引顺序文件	90
4.4.2	索引无序文件	90
4.4.3	B-树	91
4.4.4	B+树	95
4.4.5	Hash文件	96
4.5	空间数据索引	102
4.5.1	空间数据索引概述	102
4.5.2	空间索引与B+树索引	110
4.5.3	空间填充曲线的索引	111

4.5.4	网格文件	113
4.5.5	点和区域的 R 树索引方法	116
第 5 章	空间数据模型	120
5.1	实体模型	120
5.1.1	模型	120
5.1.2	实体模型	120
5.2	数据模型	123
5.2.1	层次模型与树结构	124
5.2.2	网络模型与图结构	129
5.2.3	关系模型与二维表结构	136
5.3	面向对象数据模型	141
5.3.1	面向对象的基本概念	142
5.3.2	面向对象数据模型	144
5.4	面向对象空间数据模型	148
5.4.1	地理要素数据模型	149
5.4.2	地理要素分层模型	156
5.4.3	地理空间分块模型	158
5.4.4	地理要素空间关系模型	159
5.4.5	空间数据多尺度模型	161
5.4.6	面向对象空间数据模型	162
5.5	时空数据模型	165
5.6	三维数据模型	166
5.6.1	三维空间数据库的功能	167
5.6.2	三维数据结构	167
5.7	几种常见国内外软件空间数据模型	169
5.7.1	Arc/Info 数据模型	169
5.7.2	MapInfo 数据模型	173
5.7.3	Geostar 数据模型	174
5.7.4	Oracle Spatial 的空间数据模型	175
第 6 章	空间数据库体系结构	178
6.1	空间数据库系统	178
6.1.1	空间数据库	178
6.1.2	空间数据库硬件系统	180
6.1.3	操作系统	182
6.1.4	数据字典	186
6.1.5	空间数据库管理系统	187
6.1.6	空间数据库管理员	189
6.1.7	空间数据库用户	190
6.2	数据库系统的体系结构	191

6.2.1	数据库的抽象层次	191
6.2.2	映射与数据独立	194
6.2.3	数据语言	194
6.2.4	应用程序对数据库的访问	197
6.3	空间数据库系统的体系结构	198
6.3.1	基于文件系统的体系结构	198
6.3.2	基于文件系统与数据库的混合体系结构	201
6.3.3	基于数据库管理系统的体系结构	202
6.3.4	空间数据库系统的集中式体系结构	203
6.3.5	数据库系统的客户/服务器体系结构	203
6.4	分布式空间数据库系统	207
6.4.1	空间数据的分布	208
6.4.2	分布式空间数据库系统的模式结构	210
6.4.3	分布式空间数据库系统的体系结构	211
第7章	关系数据库接口技术与地理空间数据库引擎	213
7.1	关系数据库接口技术	213
7.1.1	开放数据库互连 ODBC	213
7.1.2	数据访问对象 DAO	214
7.1.3	OLE DB	217
7.1.4	ActiveX 数据对象 (ADO)	219
7.1.5	基于 PRO * C 的 Oracle 数据库访问	221
7.1.6	基于 Oracle 的数据库 OCI 访问	222
7.2	地理空间数据库引擎	226
7.2.1	SDE 的基本概念	227
7.2.2	SDE 的发展现状	228
7.2.3	SDE 的特点	229
7.2.4	SDE 的研究内容	230
7.3	国内外地理空间数据库引擎技术分析	232
7.3.1	ArcSDE	233
7.3.2	SuperMap SDX+	235
7.3.3	MapGIS SDE	237
7.3.4	ORACLE SPATIAL	237
第8章	地理空间数据库管理系统	240
8.1	地理空间数据库管理系统功能概述	240
8.2	空间数据库定义	241
8.3	空间数据库操作	242
8.4	空间数据操作功能	247
8.4.1	空间数据获取	247
8.4.2	空间关系建立	248

8.4.3	空间数据的检索和查询	250
8.4.4	空间数据编辑功能	257
8.4.5	空间数据可视化	259
第9章	地理空间数据库系统设计	261
9.1	空间数据库设计的内容与要求	261
9.1.1	空间数据库的设计内容	261
9.1.2	空间数据库的设计要求	262
9.2	地理空间数据库系统设计方法	263
9.2.1	信息建模	264
9.2.2	语义建模	268
9.2.3	实体及联系建模	271
9.3	空间数据库设计过程	273
9.3.1	需求分析	275
9.3.2	概念数据模型	277
9.3.3	逻辑数据模型	280
9.3.4	物理数据模型	282
9.4	地理空间数据库设计技巧	284
第10章	基础地理空间数据库建立	288
10.1	基础地理空间数据库建设流程	288
10.1.1	建设方法选取	288
10.1.2	地形图数字化方法	289
10.1.3	遥感影像数字化方法	289
10.1.4	数字高程模型库建立过程	289
10.2	资料收集与处理	291
10.2.1	资料收集与分析	291
10.2.2	资料处理	292
10.3	基础地理空间数据获取	294
10.3.1	空间数据获取的一般原则	294
10.3.2	空间数据获取方法	295
10.3.3	元数据获取	299
10.4	国家基础地理空间数据库介绍	300
10.5	地理空间数据质量	307
10.5.1	地理空间数据质量概念	307
10.5.2	空间数据质量评价	308
10.5.3	空间数据质量问题的来源	309
10.5.4	常见空间数据源的误差分析	312
10.5.5	空间数据质量控制	313
第11章	地理空间数据仓库与互操作	316
11.1	空间数据互操作	316

11.1.1	多源空间数据	316
11.1.2	空间数据互操作的概念	317
11.1.3	空间数据互操作相关标准	320
11.1.4	空间数据互操作的实现方法	323
11.1.5	组件技术实现 GIS 互操作	328
11.1.6	基于 XML 的空间数据互操作实现技术	329
11.2	空间数据仓库	330
11.2.1	空间数据仓库的起源	330
11.2.2	空间数据仓库的基本特征	331
11.2.3	空间数据仓库体系结构	333
11.2.4	空间数据仓库功能组成	334
11.2.5	空间数据仓库硬件及网络结构	342
11.3	空间数据的元数据	343
11.3.1	元数据概念与分类	343
11.3.2	空间数据元数据的概念和标准	345
11.3.3	空间数据元数据的获取与管理	349
11.3.4	空间数据元数据的应用	350
主要参考文献		352

第1章 地理空间数据库导论

地理信息是描述地表形态及其所附的自然和人文地物特征和属性的总称。地理信息是人们认知世界、利用自然不可缺少的媒介，是经济社会发展的基础性、战略性资源。人类认识、利用乃至依赖地理信息和知识有着几千年的历史，最直接的证明便是地图。地图的出现甚至要早于文字。地图是运用一定的数学法则与地图语言，经过制图综合，将客观世界表现在平面上，实质上是公式化、符号化、抽象化地再现客观世界。通过地图，人类对自身生存的环境有了完整的认识。借助于地图这一简单却又有效的工具，我们可以认识未知的客观世界。地图传递信息和载负信息的能力是无法用其他任何手段替代的。因此，无论过去、现在还是将来，人们仍将大量使用地图，甚至依赖于地图。

20世纪计算机的产生和发展，几乎冲击了社会的各个领域，给许多行业带来了巨大的变化和深远的影响。计算机技术在信息处理领域的巨大功能和绝对优势，标志着社会进入了信息时代。

古老的地图学在这个变革时期也正发生着巨大的变化，新技术的引用，不但改变着传统的地图制作技术，也促使地图学理论与方法的研究不断深入。20世纪60年代把计算机引入地图学产生了计算机地图制图技术，人们用计算机表示地图要素及其相互联系，将连续的以模拟方式存在于纸质地图的空间物体离散化，以便计算机能够识别、存储和处理。

早期的计算机制图（地图制图自动化）只是把计算机作为工具来完成地图制图的任务，把人们从繁重的手工地图制图劳动中解脱出来，并由此带来了巨大的经济和社会效益。国家、军事部门和企业根据各自对地图数据的需要，投入了大量的人力、物力进行各种比例尺的地图数字化，产生了大量的地图数据。这些数据成为国家和军队的重要资源财富。它与其他数据相比，地图数据特殊的数学基础、非结构化数据结构和动态变化的时间特征，给数据获取、处理和存储带来很大难度，如何妥善保存和科学管理这些地图数据是人们长期以来十分关注的课题（毋河海，1990）。伴随着计算机数据组织存储技术的发展，地图数据的维护、更新和管理经历了从低级向高级的发展过程。最初采用文件系统的形式，后来逐步发展为地图数据库系统（map data base system, MDDBS）。该系统由地图数据、地图数据管理系统、计算机硬件设备和地图数据库管理人员等组成。

地图数据的主要来源是普通地图，反之，生产地图也是早期地图数据库建设的主要目的。因此地图数据有以下几个特点：

第一，地图比例尺影响。地图数据是某一特定比例尺的地图经数字化而产生的。地理物体表示的详细程度，不可避免地受地图综合的影响。经过了人为制图综合，地理物体的几何精度（形状）和质量特征已经不是现实世界中的真实反映。为了满足地图应用的需要，不同比例尺地图建立不同的地理数据库。

第二, 强调数据可视化, 忽略了实体的空间关系。地图数据主要是为地图生产服务的, 强调数据的可视化特征。采取的方式主要用“图形表现属性”, 地理物体的数量特征和质量特征用大量的辅助符号表示, 包括线型、粗细、颜色、纹理、文字注记、大小等数十种。地图数据是以相应的图式、规范为标准的, 依然保留着地图的各项特征。各种地理现象之间的空间位置关系, 例如, 道路两旁的植被或农田、与之相邻的居民地等, 是通过读图者的形象思维从地图上获取的。地理物体(如道路、居民地和河流)在空间关系上是相互联系的有机整体, 但在地图数据表示中是相互孤立的。因此, 地图数据不强调实体的关系表示。

第三, 按地图印刷色彩分层管理。为满足地图印刷的需求, 依据地图制图覆盖理论, 对地图数据按色彩分层管理, 不是按照地理物体的自然属性进行分类分级。这种分层不仅割裂了地理物体之间的有机联系, 也导致了同一个地物在不同层内重复存储, 如河流两岸的加固陡坎隐含着河流的水涯线信息, 道路与绿化带平行接壤使道路边沿线隐含着绿化带的边沿, 河流、道路和铁路等线状地物可能隐含着区划界限。

第四, 地图图幅限制了数据范围。受印刷机械、纸张和制图设备的限制, 传统的地图用图幅限制地图的大小, 地图数据用图幅来组织和管理。地图图幅割裂了地理物体的完整性和连续性, 比如, 一条境界线因为地图的分幅而断作几条记录存储在不同的图幅内。

随着科学技术的发展和地图数据应用的深入, 特别是计算机技术、数据库理论、信息系统理论的发展和实践的成功, 使人们对地理信息的应用已不再局限于地图这一单一产品上, 人们研究和解决空间问题需要综合地利用各种数据, 包括资源、环境、经济和社会等领域的一切带有地理坐标的数据。与地图数据相比, 这种数据主要通过属性数据描述地理实体的定性特征, 用数字表示空间实体的数量特征、质量特征和时间特征。

在数据获取手段上, 不再局限于地图的数字化, 获取空间物体信息的手段越来越多样化, 特别是随着传感器技术、航空和航天平台技术、数据通信技术的发展, 现代遥感技术已经进入一个能够动态、快速、准确、多手段提供多种对地观测数据的新阶段。新型传感器不断出现, 已从过去的单一传感器发展到现在的多种类型的传感器, 并能在不同的航天、航空遥感平台上获得不同空间分辨率、时间分辨率和光谱分辨率的遥感影像。遥感影像的空间分辨率已达到米级; 光谱分辨率已达到纳米级, 波段数已增加到数十甚至数百个; 回归周期可达几天甚至十余小时; 微波遥感已逐渐采用多极化技术、多波段技术及多种工作模式。全球定位系统(GPS)和惯性导航系统(INS)等高技术系统相结合的智能型实时地理信息获取系统步入了实用阶段, 为地理数据的实时更新提供了一个实用、简便、低廉的工具。

在地理信息表示方面, 广大科学工作者开始思索如何利用它来反映自然和社会现象的分布、组合、联系及其时空发展和变化, 研究在计算机存储介质上如何科学、真实地描述、表达和模拟现实世界中地理实体或现象、相互关系以及分布特征。初期的地图数据仅仅把各种空间实体简单地抽象成点、线和面, 这远远不能满足实际需要, 要想进一步拓宽应用前景, 必须进一步研究它们之间的关系(空间关系)。空间关系是通过一定的数据结构来描述与表达具有一定位置、属性和形态的空间实体之间的相互关系。当我们用数字形式描述空间物体, 并使系统具有特殊的空间查询、空间分析等功能时, 就必须把空间关系映射成适合计算机处理的数据结构, 这时必须考虑数据的表示方法。

在数据组织上,为了满足地理分析需求,不受传统图幅划分的限制组织数据,在人们认识世界和改造世界的一定区域内(即现实世界地理空间)不管逻辑上还是物理上均为连续的整体。

从理论上讲,地物在地理空间只有唯一的地理数据表示,空间物体本身没有比例尺的含义,应尽可能详细、真实地描述物体形状、几何精度和属性。但人们对地理环境的认识往往需要一个从总体到局部,从局部到总体反复认识过程。为了满足人们对地理空间这种认识需求,必须考虑空间物体的多尺度性,以满足不同的社会部门或学科领域的人群对空间信息选择需求。

综上所述,从数据内容、获取手段、表示方法和数据组织上这些数据已经超出了地图数据表示范畴,为了与地图数据相区分,人们称之为地理信息数据(geo-information data)。

地理信息数据的获取、处理、管理和分析及其在地学领域的应用导致了地理信息系统(geographic information system, GIS)的产生和发展。它是利用计算机及其外部设备,采集、存储、分析和描述整个或部分地球表面的空间信息系统。它的研究对象是整个地理空间,为人们采用数字形式和分析现实空间世界提供了一系列空间操作和分析方法。该系统的核心是地理信息数据库(geographic information data base, GIDB)。它是在一定的地域内,将地理空间信息和一些与该地理信息相关的属性信息结合起来,实现对地理几何特征和属性信息的采集、更新和综合管理。

地图数据和地理信息数据都是带有地理坐标的数据,是地理空间信息两种不同的表示方法,统称为地理空间数据(geospatial data)。地理空间数据库(geospatial data base, GDB)的主要任务是研究地理空间物体的数据表示、数据模型以及在计算机内数据存储结构和建立空间索引方法,如何以最小的代价高效巧妙地存储和处理空间数据,正确维护空间数据的现实性,为用户提供现实性好、准确性高、完备、开放和易用的地理空间数据。

地理空间数据库系统(geospatial data base system, GDBS)的核心软件是空间数据库管理系统(geospatial data base management system, GDBMS)。它是为了满足日益发展的空间数据管理的需要,在文件的基础上发展起来的一种空间数据管理技术。它按一定的方式组织和存储、管理地理空间数据,具有较高的程序和数据独立性,能以较少的重复为多个用户或应用程序提供数据服务。

把计算机技术与数据库技术应用于地理空间数据的管理,需要解决地理空间数据的管理和更新一系列复杂的问题,研究解决这些问题的理论方法推动了地理空间数据库技术的产生和发展。

1.1 地理空间数据

1. 地理空间

“空间”(space)的概念不同的学科有不同的解释。从物理学的角度看,空间是指宇宙在3个互相垂直的方向上所具有的广延性。从天文学的角度看,空间是指时空连续体系的一部分。地理学是研究地球表层空间分布规律的科学,因此地理学的空间是一个定义在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,物理距离只是这

些关系中的一种度量；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。也许正因为如此，今天大多数的学者都强调空间位置和拓扑关系。也就是说，地理空间 Geospace 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。这样，通过地理空间和欧氏空间的统一，将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，宏观特性主要体现在地理对象之间的拓扑关系与非拓扑关系上（通过数据模型体现），其载体则是具有精确位置、起着联结结点作用的那些单个地理空间对象（通过单对象的数据结构体现）。

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，这些事物和现象的一个典型特征，是与一定的地理空间位置有关，都具有一定的几何形态，这些事物或现象称为地理空间物体（geospatial object）。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。地理空间实体（geospatial entity）是根据分析应用的需要对空间物体进行的抽象表示。在本书中，常用地理空间物体和地理空间实体区分客观存在的物体和数字表示的物体，前者如长江，后者如表示长江的曲线（坐标串数据）。

地理空间的数学描述可以表示为：设 $E_1, E_2, E_3, \dots, E_n$ 为 n 个不同类的地理空间实体； R 表示地理空间实体值的相互联系，相互制约关系； $\Omega = \{E_1, E_2, E_3, \dots, E_n\}$ 表示地理空间中各个组成部分（实体）的集合，那么地理空间可以表示为 $S = \{\Omega, R\}$ （王家耀，2000）。

2. 地理空间数据

空间数据（spatial data）是数据的一种特殊类型。它是指凡是带有空间坐标的数据，如建筑的设计图、机械设计图和各种地图等表示成计算机能够接受的数字形式。

地理空间数据（geospatial data）是空间数据的一种特殊类型。它是指带有地理坐标的数据，包括资源、环境、经济和社会等领域的一切带有地理坐标的数据。地理空间数据区别于计算机辅助设计制造中的空间数据。在本书中，在不至于混淆的情况下，我们不加区分，统称为空间数据。

地理空间数据是地理实体的空间特征和属性特征的数字描述。地理实体的空间特征表现为地理实体的几何（定位）特征（地理实体的位置、形状、大小及其分布特征）和实体间的空间关系。地理实体的属性（定性）特征表现为实体的数量特征、质量特征和时间特征。定位是指一个已知的坐标系里空间实体都具有唯一的空间位置。定性是指有关空间实体的自然属性，它伴随着空间实体地理位置。时间特征是指空间实体随时间的变化而变化。

从概念上分，地理空间数据可分为两大类。一类是空间对象数据，它是指具有几何特征和离散特点的地理要素，如点对象、线对象、面对象、体对象等。另一类是场对象数据，它是指在一定空间范围内连续变化的地理对象，如覆盖某一地理空间的格网数字高程模型、不规则三角网、栅格影像数据等。每个离散的空间对象有一个唯一的对象标

识或相应的属性描述信息。一个场对象通常作为一个整体，场内的局部特征已经由构造该数据场的节点特征表达，如一个格网点的高程表现了该点的高度。由于离散的空间对象与场对象的特征不同，所以需要采用不同的方法进行处理和管理。根据地理实体数字描述方式的不同，空间数据可分为矢量数据和栅格数据。

空间数据适用于描述所有呈二维、三维甚至多维分布的关于域的现象，空间数据不仅能够表示实体本身的空间位置及形态信息，而且还有表示实体属性和空间关系（如拓扑关系）的信息。在空间数据中不可再分的最小单元现象称为空间实体，空间实体是对存在于这个自然世界中地理实体的抽象，主要包括点、线、面以及实体等基本类型。在空间对象建立后，还可以进一步定义其相互之间的关系，这种相互关系被称为“空间关系”，又称为“拓扑关系”。因此可以说空间数据是一种用点、线、面以及实体等基本空间数据结构来表示人们赖以生存的自然世界的数据库。

空间数据比一般信息处理中的统计数据更复杂。其复杂性体现在，一是数据类型多，有几何数据，还有表示地图要素间相互联系的关系数据，以及便于图化处理的辅助数据等，而且数据还随时间的变化各自独立地发生变化。二是数据操纵复杂，空间数据的操纵不但需要一般数据的检索、增加、删除、修改等功能，而且还需要一些特有的检索方式，如定位检索、拓扑关系检索以及一些特有的操作方式，如图形编辑。三是数据输出多样，有数据、报表，还有图形。四是数据量大，加之空间数据来源多样，不但有测量、统计数据、文字资料，还有地图，遥感图像等图形图像数据。

3. 空间数据非结构化特征

从数据组织和管理角度看，空间数据与一般的事务数据相比具有非结构化特征。

在事务数据库中，数据记录一般是结构化的，即每一个记录有相同的结构和固定的长度，记录中每个字段表达的只能是原子数据（不可再分的数据），内部无结构，不允许嵌套记录。而这种结构化不能满足空间数据表示的要求。这是因为，地理实体都具有空间坐标，空间坐标不仅指示了地理实体的位置、大小和形状，还记录了拓扑信息来表达地理实体之间的关联、邻接、包含等空间关系，拓扑数据一方面方便了空间数据的空间查询和空间分析，另一方面也给空间数据的一致性和完整性维护增加了复杂性。

因此，空间数据的组织和管理不同于一般的事务性数据，要根据它的空间分布特征，建立相应的空间索引，从而实现空间数据高效快速的存储和提取。

如果一个关系表表示一类空间对象，一条记录表示其中的一个空间对象，则它的数据项可能是变长的，例如，公路的长度是变化的，可能用两对坐标表示，也可能要用几十上百对坐标表示。另外，一个空间对象可能包含一个或多个其他空间对象，例如，若用一个记录表示一条弧段，另一个记录表示一个多边形，则当一个多边形由多条弧段组成时，一条多边形记录就要嵌套多条弧段的记录。这就是空间数据难以直接采用通用的数据库管理系统的主要原因。

相对于一般的事务数据而言，空间数据量大，一幅标准的地形图矢量数据可达几兆，一幅标准地图分幅的数字影像数据可达上百兆，一个区域地理信息系统的空间数据量可能达几十千兆，或数百千兆。

用以描述事物或现象随时间的变化。这种变化表现为三种可能的形式，一是属性变

化，其空间坐标或位置不变；二是空间坐标或位置变化，而属性不变，这里空间坐标或位置变化既可以是单一实体的位置、方向、尺寸、形状等发生变化，也可以是两个或两个以上的空间实体之间的关系发生变化；三是空间实体或现象的坐标和属性都发生变化，例如，土地权属变更、海岸线变化、土地城市化、道路改线、环境变化等，需要保存并有效地管理历史变化数据，以便将来重建历史状态、跟踪变化、预测未来。如何有效地表达、记录和管理现实世界的实体及其相互关系随时间不断发生的变化，增加空间数据的时间维，使空间数据的表示非常复杂，大大增加了组织、管理、操作时间和空间数据的难度。

这些特征对地理空间数据的管理有着重要的影响：用商业数据库管理系统存储和管理这些非结构化的空间数据十分困难；另一方面地理对象之间存在着位置空间关系，地理空间数据查询时必须考虑地理对象之间的联系，即空间拓扑关系，提供空间查询语言是空间数据库的一个重要特征，使用关系数据中的“select-from-where”模式很难完成构建空间数据查询。通过扩充 SQL 语言，使其支持空间对象类型、空间关系和空间操作，为空间查询语言的设计和开发提供了一个框架。空间地理数据的时间特征使空间地理数据模型必须具有时间维，来保存地理要素随时间变化的历史性数据。

由于空间实体间的相互关系及其时空变化的描述与表达、数据组织、空间查询分析等方面均有较大的复杂性和特殊性，一般的商业数据库管理系统难以满足要求。因此，人们不得不跟随计算机软件和硬件技术发展，特别是数据库技术的前沿，研究空间数据管理的理论、技术和方法，研制空间数据库管理软件或在商业数据库管理系统的基础上开发空间数据库管理系统。

1.2 空间数据管理演变过程

空间数据的管理技术与计算机硬件和软件技术是密不可分的。特别是随着面向对象、组件技术、分布式计算技术以及网络技术和计算机存储技术的发展，空间数据管理能力也不断发展，空间数据的管理技术大体经历了 6 个阶段：人工管理阶段、文件系统阶段、文件与数据库系统混合管理系统阶段、全关系型空间数据库管理系统阶段、对象关系数据库管理系统阶段和面向对象的数据库系统阶段。

1. 人工管理阶段（20 世纪 50 年代中期）

从首次使用计算机管理与地球相关的数据（地理空间数据）以来，我们已经走过一段漫长的路程。这个阶段计算机主要应用于科技计算，硬件背景是外存只有磁带、卡片、纸带等，没有磁盘等直接存取存储设备；计算机没有操作系统，没有管理数据的软件，数据处理的方式是批处理。

人工管理阶段数据管理的特点是：

1) 数据不保存

因为计算机主要应用于科技计算，一般不需要将数据长期保存，且数据量远远小于程序量，只在计算时数据输入，用完撤走。