

GIS^① 空间数据组织管理

龚 强 编著



哈尔滨地图出版社

ZUZHIGUANLI

GIS 及

空间数据组织管理

龚 强 编著

哈尔滨地图出版社
· 哈尔滨 ·

图书在版编目 (GIP) 数据

GIS 及空间数据组织管理/龚强编著. —哈尔滨: 哈尔滨地图出版社, 2004. 4

ISBN 7 - 80529 - 744 - 4

I. G... II. 龚... III. 地理信息系统 - 应用软件,
GIS VI. P208

中国版本图书馆 GIP 数据核字 (2004) 第 017803 号

哈尔滨地图出版社、发行

(地址: 哈尔滨市南岗区测绘路 2 号 邮政编码: 150086)

黑龙江省教育厅印刷厂印刷

开本: 850 mm × 1 168 mm 1/32 印张: 8.5 字数: 230 千字

2004 年 4 月第 1 版 2004 年 4 月第 1 次印刷

印数: 1 ~ 2 000 定价: 26.00 元

序

地理信息系统(GIS)是空间信息科学的重要组成部分。作为对地理信息自动处理与分析的系统, GIS可以实现数据采集—分析—决策应用的全部过程,并能够回答和解决位置、条件、趋势、模式、模拟等五类问题。正是由于GIS具有获取、处理、分析、显示、访问以及在不同用户、不同区域、不同地点、不同系统之间传输空间信息的特殊功能,它已经广泛应用于全球化问题的研究、国家安全保障、国民经济的各个领域以及社会生活的诸多方面。由于在全世界的信息集合中,有70%~80%的信息具有空间属性,随着人类获取空间信息数据能力的不断提高、计算机技术及空间信息技术的迅速发展,地理信息系统的应用将提升到一个新的高度,拓展到一个新的广度。

作为一门新兴的交叉科学,地理信息系统正处在一个快速发展的阶段。对此,国家科技部部长徐冠华院士提出:今后要实现地理信息系统的“四个转变”,而其中的第一个转变就是“由系统为中心向数据为中心的转变”。应该说,这是地理信息系统发展的一个方向。

龚强博士的又一部著作——《GIS及空间数据组织管理》出版了。这是他在为地理信息系统专业的学生讲授“GIS及空间数据组织管理”专业课而写的讲

义基础上,充分考虑不同读者的需求,经过反复修改才成稿的。作为这本书的第一个读者,我以为它有以下一些主要特点:

集中阐述、重点突出。据我所知,这可能是第一本集中、系统、专门介绍空间数据组织、处理、结构、转换、挖掘、融合、管理的书。许多 GIS 的著作、讲义多是在介绍 GIS 其他理论方法的同时,用部分章节介绍空间数据。而本书全面介绍了这方面的理论、技术与方法。

注重热点、立足前沿。空间信息技术发展非常快,我国这项事业发展也非常快。鉴于此,本书介绍了诸如我国基础地理信息建设的进展状况、当前 GIS 研究的热点问题以及空间数据集成问题等;还有许多内容属于空间地理信息乃至计算机专业相对比较前沿的问题,如数据挖掘、空间数据仓库、空间数据挖掘与数据融合等,特别是网格理论、网格的关键技术、基于网格对 GIS 问题的研究等。网格是上个世纪末期出现的新技术,目前我国测绘界对这方面的研究还不是太多,因此,专题探讨网格也是本书的一个特点。

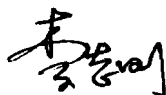
完整介绍、实用性强。除了对前沿理论的阐述,本书还把一些新技术、新方法的来龙去脉简要进行了介绍,使读者了解其发展轨迹。目前,许多生产单位的设备尚不能马上更新,仍然是用原来的手段或方法作业,全面介绍新技术、新方法,如空间数据的数字化问题、数据集成问题等是很有必要的。

图文并茂、通俗易懂。把现代前沿理论、专业化很强的学术问题写得通俗易懂、简明好懂并不是一件很容易的事。不仅这本书,龚强发表的百余篇专业论文以及他的专著——《测绘运筹学》、《城市综合功能地理信息系统》都有这个特点。本书行文流畅,语言朴实,深入浅出,图文并茂,融汇创新,难能可贵。

恢复高考后,龚强先后就读于武汉测绘学院大地系,哈尔滨教育学院外语专业,哈尔滨工业大学经济管理专业,毕业后考取

哈尔滨工业大学研究生，成为我省第一代 MBA(工商管理硕士)，接着又获哈工大工学硕士和博士学位。他还具有翻译职称。作为教授级高级工程师，他有外业测绘、GIS 教学、主持项目、系统集成等较多的实际工作经历，曾五次获国家、省部级科技成果奖。他基本功扎实且治学严谨，而他多年如一日锲而不舍地学习、探索，潜心于科学研究，一直拼搏奋斗在我们这个专业的科技前沿也是令人钦佩的。

我郑重地将这本内容新颖、特点突出的《GIS 及空间数据组织管理》推荐给大家，相信对地理信息系统专业的学生来说，它是一本好教材；对从事这一专业教学、生产和研究的同仁来说，它也一定是一本有价值的参考书，具有很强的借鉴和启迪思维的作用。



2004 年 3 月 1 日

前言

地理信息系统 (GIS) 是一门新兴的科学技术, 自问世以来, 无论是理论还是应用都处在一个迅速发展的阶段。地理信息系统将计算机技术、空间信息技术融合在一起, 更加丰富了空间地理信息的管理与分析方法。目前, 在地学研究、区域规划、城市管理、环境整治、企业经营等许多领域都大量地应用 GIS。正如陈述彭院士所说: 从 GIS 学科的本质及其自身的发展规律来看, 地理信息系统不仅仅是技术, 而且是科学, 地理信息是现代信息的重要组成部分, 是发展地球系统科学不可缺少的部分。

陈述彭院士曾把地理信息系统比做水利设施, 将数据视为水利设施中的“水”, 足见数据的重要。而实际也是这样: 地理数据 (包括空间数据和属性数据) 的采集、处理、存储、转换、管理和应用, 一直是 GIS 研究的重点领域之一。

本书共分八个部分:

第一章地理信息系统概论部分。首先, 概略地介绍了地理信息系统及其发展轨迹, 分析了 GIS 发展的原动力, 并介绍了我国基础地理信息建设的进展情况以及当前 GIS 研究的若干重点问题; 其次, 对空间数据的有关概念、信息系统及其类型等分别

进行了讨论。

第二章论述空间数据的组织与处理。在这一章里较系统地阐述了空间数据的组织,包括空间数据采集、空间数据检核、空间数据质量、空间数据的初步处理、空间数据处理和使用中的误差、空间数据质量控制以及数据存储与组织;还概略地介绍了空间查询与分析,在此基础上对地理信息系统的核心功能进行了论析,并介绍了空间数据的处理(包括地图数字化,利用数字化仪和扫描矢量化,以及空间数据录入后的处理等)。

第三章的内容是空间数据库的管理。在较全面地介绍了空间数据库之后,分析了地理信息系统的内部数据结构;从实际工作需要出发,介绍了传统的空间数据与属性数据连接;最后介绍了现代空间数据与属性数据的集成。

第四章重点介绍栅格数据结构、矢量数据结构;两种数据的获取方法;栅格数据结构和矢量数据结构的编码;在对两种数据结构进行比较分析的基础上,论述了两种数据互换的必要性,并较详细地介绍了栅格数据结构和矢量数据结构的互换。

第五章介绍了空间索引机制和空间信息查询。

第六章是空间数据的挖掘与融合。在这一章里,介绍了数据挖掘、数据仓库、数据挖掘的方法以及数据挖掘的步骤与模式;介绍了空间数据挖掘的概念、空间数据仓库;论析了空间数据挖掘的方法和空间数据挖掘的应用前景;最后介绍了数据的融合。

第七章集中分析、论述了 GIS 与相关学科、相关技术的关系以及内在的联系,而后论述了 GIS 与可持续发展问题。

第八章是网格技术与 GIS。在这一章里,概述了网格的概念、网格的关键技术、网格的研究与发展现状;分析介绍了网格技术与 GIS 发展的关系;并基于网格对地理信息系统理论问题进行了探讨。

本书是专门为地理信息系统专业学生准备的教材。从课程性质讲,“GIS及空间数据组织管理”是一门专业技术课。其教学的基本要求是:全面了解地理信息系统及其发展轨迹、空间数据及各类信息系统;重点掌握对空间数据的组织、处理;掌握空间数据库的管理;掌握栅格数据结构和矢量数据结构的编码与互换,以及空间索引机制与空间信息查询方法;认识数据挖掘、空间数据挖掘与融合的一般理论与方法;了解当前最新的信息技术——网格。如果说这门课的直接目的是使学生全面掌握上述理论与方法,那么,另外一个目的就是培养学生的空间数据抽象能力,拓展学生的知识领域,尤其是信息和空间信息的理论与方法。在此基础上使其深入理解GIS与相关学科、相关技术以及可持续发展的内在关系,开拓学生的视野,逐步培养学生向更深层次和更广范围探索、研究的能力。

本书的编写得到了黑龙江工程学院测绘工程系主任周秋生教授、测绘工程系副主任马俊海教授、王延亮副教授、朱春晓高级工程师、黄明讲师的支持和帮助;版式设计、校对、印刷等方面的工作得到了黑龙江省地方志办公室胡玉芬处长的帮助;哈尔滨师范大学计算机双学位班的冀天卓同学为本书提了很多有益的意见和建议,并为本书制作了插图,在此一并表示衷心感谢。

由于我的水平有限,书中错误之处在所难免,敬请广大读者指正。

龚 强

目 录

1	地理信息系统概论	(1)
1.1	地理信息系统及其发展	(3)
1.1.1	GIS 的基本概念及内涵	(3)
1.1.2	GIS 的产生与发展轨迹	(5)
1.1.3	GIS 发展的原动力	(10)
1.1.4	我国基础地理信息建设的进展	(16)
1.1.5	当前 GIS 研究的几个重点问题	(22)
1.2	空间数据的有关概念	(28)
1.2.1	信息	(28)
1.2.2	数据	(30)
1.2.3	地理信息和地理数据	(32)
1.2.4	空间数据	(34)
1.3	信息系统及其类型	(35)
1.3.1	信息系统概述	(35)
1.3.2	信息系统的类型	(38)
1.3.3	数字城市的管理信息技术	(39)
	思考题	(43)
2	空间数据的组织与处理	(45)
2.1	空间数据采集、检核与组织	(45)
2.1.1	空间数据采集	(45)
2.1.2	空间数据检核	(46)
2.1.3	空间数据质量	(47)

2.1.4	空间数据初步处理	(52)
2.1.5	空间数据处理中的误差	(52)
2.1.6	空间数据使用中的误差	(54)
2.1.7	空间数据的质量控制	(55)
2.1.8	数据存储与组织	(56)
2.2	空间查询与空间分析概述	(57)
2.2.1	空间数据查询与分析	(57)
2.2.2	图形与交互显示	(61)
2.3	空间数据的特征与 GIS 核心功能论析	(62)
2.3.1	空间数据的一般特征	(62)
2.3.2	地理信息系统核心功能论析	(65)
2.4	空间数据的处理	(69)
2.4.1	地图数字化	(69)
2.4.2	地图数据的两种类型	(72)
2.4.3	利用数字化仪实现数字化	(74)
2.4.4	扫描矢量化及处理流程	(75)
2.4.5	空间数据录入后的处理	(76)
	思考题	(77)
3	空间数据库的管理	(79)
3.1	空间数据库	(79)
3.1.1	地理信息系统与一般管理信息系统的对比分析	(79)
3.1.2	空间数据库	(81)
3.1.3	数据文件及其组织	(85)
3.2	GIS 的内部数据结构	(88)
3.2.1	GIS 空间数据建模	(90)
3.2.2	矢量模型	(92)
3.2.3	栅格模型	(93)

3.3	传统的空间数据与属性数据连接	(95)
3.3.1	属性数据的组织	(95)
3.3.2	空间数据与专题属性数据的连接	(99)
3.4	现代空间数据与属性数据的集成	(103)
3.4.1	几种典型的数据模型	(104)
3.4.2	空间数据与属性数据的集成	(107)
	思考题	(113)
4	栅格矢量结构、编码及互换	(115)
4.1	栅格数据结构	(115)
4.1.1	定义及特点	(115)
4.1.2	栅格数据的获取	(117)
4.1.3	栅格单元代码的决定方式	(119)
4.1.4	编码方式	(122)
4.2	矢量数据结构	(130)
4.2.1	定义及特点	(130)
4.2.2	矢量数据的获取	(131)
4.2.3	编码方法	(132)
4.3	栅格结构与矢量结构的比较及互换	(140)
4.3.1	两种结构的比较	(140)
4.3.2	两种结构相互转换的必要性	(143)
4.3.3	矢量格式向栅格格式的转换	(144)
4.3.4	栅格格式向矢量格式的转换	(150)
	思考题	(156)
5	空间索引机制与空间信息查询	(157)
5.1	空间索引概述	(157)
5.2	空间索引的类型	(159)
5.2.1	格网型空间索引	(159)
5.2.2	BSP 树空间索引	(160)

5.2.3	KDB 树空间索引	(160)
5.2.4	R 树和 R+ 树空间索引	(162)
5.2.5	CELL 树空间索引	(163)
5.3	空间信息查询	(165)
5.3.1	基于属性特征的查询	(168)
5.3.2	基于空间关系和属性特征的查询(SQL)	(169)
5.3.3	一种空间扩展 SQL 查询语言——GeoSQL	(169)
	思考题	(171)
6	空间数据的挖掘与融合	(172)
6.1	数据挖掘	(172)
6.1.1	数据挖掘概述	(172)
6.1.2	数据仓库	(174)
6.1.3	数据挖掘方法	(175)
6.1.4	数据挖掘步骤与模式	(177)
6.2	空间数据挖掘	(178)
6.2.1	空间数据挖掘的概念性框架	(180)
6.2.2	空间数据仓库	(183)
6.2.3	空间数据挖掘方法论析	(184)
6.2.4	应用前景	(191)
6.3	数据融合	(193)
	思考题	(196)
7	GIS 与相关学科、技术	(197)
7.1	GIS 与相关学科	(197)
7.1.1	GIS 与地理学	(197)
7.1.2	GIS 与测绘学	(199)
7.1.3	GIS 与地图学	(200)
7.1.4	GIS 与计算机科学	(202)
7.1.5	GIS 与测绘运筹学	(204)

7.2	GIS 与相关技术	(208)
7.2.1	GIS 与遥感	(208)
7.2.2	GIS 与全球定位系统	(209)
7.2.3	GIS 与互联网	(210)
7.2.4	GIS 与数字地球	(212)
7.2.5	GIS 与数字城市	(216)
7.3	GIS 与可持续发展	(219)
	思考题	(223)
8	网格技术与 GIS	(224)
8.1	网格概述	(224)
8.1.1	网格的概念与内涵	(225)
8.1.2	网格的关键技术	(227)
8.1.3	网格研究与发展概况	(233)
8.2	网格技术与 GIS 发展	(239)
8.2.1	我国学者关于空间信息多级网格理论方法体系的初步设想	(239)
8.2.2	网格技术对 GIS 的支持	(240)
8.3	基于网格对 GIS 的思考	(242)
8.3.1	网格技术是高科技发展的必然	(242)
8.3.2	基于网格对 GIS 理论问题的思考	(243)
	思考题	(245)
	参考文献	(247)
	后记	(250)

1

地理信息系统概论

在席卷全球的知识经济和新技术革命大潮中,人类社会步入了 21 世纪,迎来了信息时代 (Information Era)。同时,人类社会也迅速地演化为信息社会 (Information Society)。信息时代强大的生命力和信息革命巨大的社会推动力,预示着未来社会的进一步高度文明和人类智能劳动的现代化。这一点已为最近几年的社会发展所证实 (随着时间的推移,将进一步得到证实)。今天,信息已经成为一种人类社会不可缺少的,对人类进步有着巨大影响的社会资源和社会动力源。而这种社会资源和社会动力源引发了社会职能、结构、价值、产业及生产、生活方式的变革,同时也给人们的观念、意识、行为方式等带来了一场革命。毫无疑问,在信息社会中,信息、知识和技术将成为社会发展的动力和经济发展的基础。

目前,尽管对信息的概念和定义还没有公认的定义,但信息是客观世界中继物质、能量之后的第三个现代科学的基本概念,已成为人们的共识。正如有人形容的,信息如同落地水银,无孔不入,它渗透到大千世界的所有领域,方方面面。

纵观人类发展的历史,我们可以清楚地看到:人类认识物质以后,带来的是 8 000 ~ 6 000 年以来的农业革命;人类认识能量之后,带来的是 300 年以来的

工业革命；而人类认识信息以后，带来的是 50 年以来的信息革命。农业革命主要涉及的是物质，工业革命主要涉及的是物质和能量的结合（体现为蒸汽机、内燃机、电动机等），而信息革命则把物质、能量和信息紧密地结合在一起（体现为计算机）。

为了推动信息社会的迅速发展，1993 年 2 月，当时的美国总统克林顿提出了信息高速公路的概念，并签署法令，要求尽快建立覆盖整个美国的信息高速公路，即“国家信息基础设施”（National Information Infrastructure, 简称 NII）。为了在信息高速公路上表达地理参考，使之与地理和地球有关的空间信息得以在因特网上准确地表达、描述和查询，需要建立全国的空间数据框架。为此，克林顿于 1994 年 4 月 13 日签署了“建立国家空间数据基础设施”（National Spatial Data Infrastructure, NSDI）的 12906 号总统令。

NSDI 这项总计投资 4 000 亿美元的跨世纪信息基础工程，目标是为 21 世纪信息文明打好物质基础，使公众拥有良好的信息环境。目前，美国的信息高速公路和空间数据基础设施已初见端倪。为了将信息技术推进到人们的日常生活、工作、娱乐之中，当时的美国副总统戈尔又于 1998 年 1 月 31 日提出了数字地球（Digital Global）的概念：“我们需要一个数字地球，一个可以嵌入海量地理数据的，多分辨率的，真实地球的三维表示。”他认为数字地球所需要的技术涉及以建模和与数字模拟为特征的计算机科学，海量储存技术，高分辨率的卫星图像技术，每秒传送一百万兆比特数据的宽带网络，互操作规范，元数据标准，以及卫星图像的自动解译、多元数据的融合和智能代理等。

数字地球的概念提出以后，国际社会反响强烈，并引起了科技界的高度重视。数字地球是一个向公众开放的系统，它可以虚拟现实技术，进一步互联和共享，真正实现“地球村”的设想。而更主要的是利用数字地球这一工具，来支持和改善人类活动以及生活质量。随着社会的发展和科学技术的进步，以及数字地球技术的普

及,人们将进一步体会到,数字地球系统将逐渐改变人类社会的生产方式和生活方式。例如:数字农业(也称为信息农业)、数字交通、数字旅游、数字海洋捕捞等;数字经济:包括电子商务、电子金融(银行、保险、股票、期货、税务)等;数字社会:包括电子政务、电子公安、电子法院等;此外,还有数字城市、数字社区等。这些都会极大地促进产业的生产规模扩大,推动社会经济的发展。由数字地球技术系统产生的数字军事、数字战场、数字外交、数字政治等也将改变我们的思维模式和操作方式。数字地球把遥感技术(RS)、卫星定位技术(GPS)、地理信息系统(GIS)、网络技术乃至网格(Grid)技术等与社会需求联系在一起,为全球信息化提供一个基础框架。

由此可见,GIS是数字地球的重要组成部分。数字城市是数字地球的局部,GIS则是这一局部最重要的数字化组成之一。

1.1 地理信息系统及其发展

1.1.1 GIS的基本概念及内涵

地理信息系统(Geographical Information System,GIS)是一种决策支持系统,具有信息系统的各种特点。GIS区别于其他信息系统的主要特征是其存储和处理的信息经过了地理编码,地理位置及与该位置有关的地物属性信息成为信息检索的重要部分。

关于GIS的概念,我国学者根据理解和实践,进行了精辟的总结和定义,在许多文献中都有论述,虽然表述方法不尽相同,但基本思想和内涵大体是一致的。

通俗地讲,地理信息系统(GIS)是整个地球或部分区域的资源、环境在计算机中的缩影。严格地讲,地理信息系统是反映人们