目　　录

第一章　概论 …… 1
第 1—1 节　信息技术与地理信息系统 …… 1
一、信息、信息技术与信息社会 …… 1
二、地理信息、信息系统与地理信息系统 …… 2
第 1—2 节　GIS 的发展 …… 3
一、GIS 的发展简史 …… 3
二、中国 GIS 的发展概况 …… 5
第 1—3 节　GIS 的硬软件环境 …… 7
一、GIS 的硬件配置 …… 7
二、GIS 软件的特点、基本要求和组成 …… 10
第 1—4 节　GIS 的功能与应用领域 …… 15
一、GIS 的基本功能 …… 15
二、GIS 技术的主要应用领域 …… 18
第 1—5 节　GIS 的相关学科与技术 …… 19
第二章　GIS 的数据结构 …… 21
第 2—1 节　概述 …… 21
一、GIS 中的空间信息及其特征 …… 21
二、GIS 中的数据类型 …… 23
三、空间数据的数据结构 …… 23
四、空间数据的抽象表示 …… 23
五、空间数据的拓扑关系 …… 24
第 2—2 节　矢量数据结构 …… 25
一、矢量数据结构的图形表示 …… 25
二、矢量结构的几种编码方式 …… 26
第 2—3 节　栅格数据结构 …… 28
一、栅格数据结构的图形表示 …… 28
二、栅格结构的建立 …… 29
三、栅格结构编码的方法 …… 30
第 2—4 节　四叉树与八叉树数据结构 …… 32
一、四叉树概述 …… 32
二、基于四进制的线性四叉树编码 …… 34
三、基于十进制的线性四叉树编码 …… 35
四、二维行程编码 …… 36
五、八叉树结构与十六叉树结构 …… 36
第 2—5 节　数据结构的转换 …… 41
一、矢量与栅格结构的比较 …… 41
二、由矢量向栅格的转换 …… 41

三、由栅格向矢量的转换 …… 43
第 2—6 节 栅格-矢量一体化数据结构 …… 46
一、栅矢一体化结构的基本概念 …… 46
二、栅矢一体化数据结构设计 …… 47
三、Voronoi 动态空间数据结构 …… 50
第三章 GIS 的数据模型 …… 52
第 3—1 节 空间目标数据模型化 …… 52
第 3—2 节 关系模型、层次模型与网络模型 …… 54
一、层次数据模型 …… 55
二、网络数据模型 …… 55
三、关系数据模型 …… 56
第 3—3 节 综合数据模型 …… 57
一、双元模型 …… 57
二、分层模型 …… 58
三、综合模型 …… 58
第 3—4 节 超图的概念与多媒体空间信息系统 …… 59
一、超图数据结构的理论框架 …… 59
二、超图数据库设计基本原理 …… 61
第 3—5 节 时空 GIS 数据库系统的概念 …… 63
一、数据模型 …… 64
二、目标数据处理和数据结构 …… 66
三、四维时空数据的组织和时空复合分析 …… 68
第四章 整体空间数据库 …… 71
第 4—1 节 从传统数据模型到面向对象模型 …… 71
一、传统数据模型的弱点 …… 71
二、语义数据模型 …… 72
三、面向对象的数据模型 …… 72
第 4—2 节 面向对象技术的原理与方法 …… 73
一、对象、消息和封装 …… 73
二、类和实例 …… 75
三、继承及类之间的层次关系 …… 76
四、概括、聚集、联合与传播 …… 80
第 4—3 节 面向对象数据库的分析与设计 …… 81
一、概念模式设计 …… 81
二、逻辑模式设计 …… 84
三、物理数据库设计 …… 86
四、面向对象的属性数据库实现方法 …… 86
第 4—4 节 人工智能与专家系统 …… 87
一、知识表示 …… 87
二、知识推理 …… 89
三、知识获取 …… 90
四、专家系统 …… 91
第 4—5 节 空间知识的表示 …… 91

一、谓词逻辑用于空间知识表示 …… 91
二、语义网络的应用 …… 93
三、规则在拓扑关系中的应用 …… 95
四、空间知识的框架表示 …… 95
五、空间知识的过程表示 …… 96
六、不精确的空间知识表示 …… 97
七、空间知识工程示例 …… 97
第五章 商品化 GIS 软件系统介绍与评述 …… 99
第 5—1 节 概述 …… 99
第 5—2 节 ARC/INFO GIS 软件 …… 100
一、PC ARC/INFO 系统的软件结构 …… 100
二、PC ARC/INFO 系统的文件结构 …… 101
三、实现汉字注记的技术方案 …… 103
四、PC ARC/INFO 3.4.2 的特点和功能 …… 104
五、ARC/INFO7.0.2 版本的特色 …… 105
第 5—3 节 桌面式 GIS 软件 ARCVIEW2.0 …… 106
第 5—4 节 GENAMAP GIS 软件 …… 108
第 5—5 节 Intergraph MGE GIS 软件 …… 110
第 5—6 节 GRASS GIS 软件 …… 112
第 5—7 节 城市之星（Citystar）GIS 软件 …… 114
第 5—8 节 吉奥之星（Geostar）GIS 软件 …… 115
第六章 空间数据的处理与管理 …… 118
第 6—1 节 空间数据的输入与编辑 …… 118
一、图形数据的输入 …… 118
二、属性数据的输入 …… 119
三、GIS 数据的编辑 …… 119
第 6—2 节 遥感数据与 GIS 的综合 …… 121
一、GIS 与遥感图像处理系统的结合 …… 121
二、遥感图像信息的解译 …… 123
三、遥感图像信息与其它空间信息的综合 …… 125
第 6—3 节 空间数据管理 …… 126
一、图幅数据的处理 …… 126
二、数据格式的转换 …… 128
三、空间数据与属性数据的连接 …… 129
四、空间数据开窗处理 …… 130
五、空间数据的压缩处理 …… 133
六、图幅接边与分割 …… 135
七、空间数据的统计综合 …… 137
八、数据更新 …… 137
九、空间数据的索引与检索 …… 139
第 6—4 节 空间数据的质量与表示 …… 147
一、源误差 …… 148
二、操作误差 …… 150

三、GIS 空间操作中误差的传播 …… 151
第七章 GIS 中的数据分析与输出 …… 153
第 7—1 节 空间数据分析 …… 153
一、综合属性数据分析 …… 154
二、缓冲区分析 …… 156
三、空间合成叠置分析 …… 160
四、网络分析 …… 166
第 7—2 节 数字地面模型（DTM）与地形分析 …… 171
一、DTM 的概念与作用 …… 171
二、DTM 的数据采集与表示 …… 172
三、DTM 的空间内插方法 …… 174
四、DTM 的地形分析 …… 179
第 7—3 节 数据输出与自动制图 …… 185
一、数据输出设备 …… 185
二、图形设备驱动软件 …… 188
三、制图符号与汉字注记 …… 189
四、制图数据的自动处理 …… 193
五、GIS 数据制图与输出的动画和多媒体技术 …… 196
第 7—4 节 数据共享和标准化 …… 199
一、数据共享 …… 199
二、数据标准化和规范化 …… 201
第八章 地理信息系统的应用 …… 207
第 8—1 节 概述 …… 207
第 8—2 节 GIS 的工程建设 …… 207
一、技术力量的组织管理 …… 208
二、GIS 工程建设 …… 208
第 8—3 节 GIS 的应用模型 …… 209
一、应用系统的数学模型 …… 209
二、模型的开发方法 …… 210
三、应用模型开发举例 …… 212
第 8—4 节 国家基础地理信息系统 …… 215
第 8—5 节 土地资源管理信息系统 …… 216
一、土地信息系统的必要性和可能性 …… 217
二、LIS 的基本硬软件配置要求 …… 217
三、建立 LIS 需要解决的一些主要技术问题 …… 218
第 8—6 节 城市规划管理信息系统 …… 221
第 8—7 节 区域资源与环境信息系统 …… 222
一、自然资源的信息分类 …… 222
二、全球资源与环境数据库系统 …… 224
三、水资源动态分析预测系统 …… 224
四、多元地学信息找矿系统 …… 225
第 8—8 节 矿区资源与环境信息系统 …… 226
一、矿区资源开发及其环境问题 …… 226

二、矿区资源环境信息系统的基本内容和关键技术 …… 227
三、应用模型研究及举例 …… 236
附录 世界主要GIS商品软件特征一览表 …… 240
参考文献 …… 246
后语 …… 248

第一章 概　　论

第1—1节　信息技术与地理信息系统

一、信息、信息技术与信息社会

在席卷全球的新技术革命浪潮中，人类迎来了信息时代（社会）。面临着信息时代的强大活力，预示着未来社会的文明和人类智力劳动的现代化。当今，信息已成为一种重要的社会资源和社会动力源，引起了社会职能、结构、价值、产业及生活方式的变革，也给人们的观念、方法和行为带来了一场革命。我国正处在社会主义市场经济迅速发展的新时期，信息和信息交换的重要性越来越显示出来，越来越被人们所认识和重视。

信息（Information）是用文字、数字、符号、语言、图像等介质来表示事件、事物、现象等的内容、数量或特征，从而向人们（或系统）提供关于现实世界新的事实和知识，作为生产、建设、经营、管理、分析和决策的依据。信息具有客观性、适用性、可传输性和共享性等特征。信息来源于数据（Data）。数据是一种未经加工的原始资料。数字、文字、符号、图像都是数据。数据是客观对象的表示，而信息则是数据内涵的意义，是数据的内容和解释。例如，从实地或社会调查数据中可获取到各种专门信息；从测量数据中可以抽取出地面目标或物体的形状、大小和位置等信息；从遥感图像数据中可以提取出各种地物的图形大小和专题信息。

信息流是指在空间和时间上向同一方向运动过程中的一组信息。它也是物质流、能量流、价值流的外化形式。人类调控生产、经营活动是通过掌握物质流、能量流、价值流发出的信息流来实现的。信息流的特点之一是以物质流和能量流为载体，进行双向性传递，即有从输入到输出的信息传递，也有从输出向输入的信息反馈。人们可以按照这些反馈信息来改变输入的内容或数量，以便对被控制对象产生新的影响。信息流的畅通是保证各种生产、经营和社会活动正常运行的必要条件。因此，加强对信息的组织管理是提高经济建设和社会发展水平的重要保证。本世纪以来，世界劳动人口从物质生产领域向包括信息产业在内的服务性行业转移。今后，第四产业——信息产业在世界经济和社会发展中的作用将越来越显著。信息是未来的“核心”。信息技术是当今世界高技术群的代表，可以帮助人们收集情报、揭示问题和预测未来，极大地扩大和增强人类征服自然的能力，为人类社会带来更大的经济效益和福利。

我国政府正在实施一项跨世纪、全国性的“金桥工程”，即“中国经济信息国道”和“中国信息高速公路”。这是一项庞大的系统工程，是国家经济信息化的基础建设。“金桥工程”与“金关”、“金卡”工程合称“三金工程”。“金桥工程”是为加强国家宏观经济调控需要，为实现“金关”、“金卡”工程以及各地区、各部门、各单位的信息业务系统提供一个技术上先进、经济上合理的公用信息通道。它以卫星综合数字业务网作为基干网，与邮电分组交换网、数字数据网互联互通，互为备用，形成空地一体的网络体系，传输数据、话音、图像。

二、地理信息、信息系统与地理信息系统

地理信息是指与研究对象的空间地理分布有关的信息，它表示地球表层物体及环境所固有的数量、质量、分布特征，相互联系和变化规律。从地理实体到地理数据，再到地理信息的发展，反映了人类认识的巨大飞跃。地理信息属于空间信息。地理数据的种类、特征是与其地理位置联系在一起的，因此具有地域性。地理信息又具有多重结构的特征，即在同一经纬度位置上可以有多种专题和属性的信息结构。例如，在一个地面点位上，有其相应的高程值，重力、岩层特征、污染等多种信息。此外，地理信息还有明显的时序特征，即动态变化特征。这就要求及时采集和更新地理信息，并根据多时相的数据或信息来寻求随时间的分布和变化规律，进而对未来作出预测或预报。

为了有效地对信息流进行控制、组织管理，实现双向传递，需要通过某种信息系统。它能对数据和信息进行采集、存储、加工和再现，并能回答用户的一系列问题。信息系统有四大基本功能：数据的采集、管理、分析和表达。从计算机技术在信息科学中的应用角度看，信息系统是由计算机硬件、软件、数据和用户四大要素组成的问答系统。智能化的信息系统还包括知识和经验。计算机硬件包括各类计算机处理机及终端和外部设备；软件是支撑数据和信息的采集、存储、加工、再现及回答问题的程序系统；数据是系统中的重要组成部分，包括定量和定性数据；用户是信息系统的服务对象或使用者，是系统的主人。有一般用户和从事系统的设计、建设、维护、管理和更新的高级用户。

信息系统通常包括企业、事业管理信息系统、财务管理信息系统、交通运输信息系统和空间信息系统等。其中的空间信息系统（Spatial Information System，SIS）是一种十分重要而特殊的信息系统，它采集、管理、处理、输出和更新的信息要包括空间信息和数据。

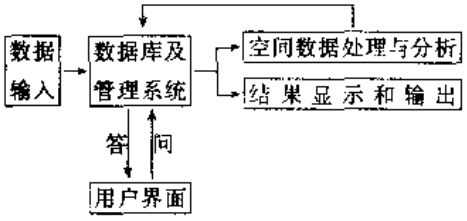


图 1 -1　GIS 的概念框架构成

地理信息系统（Geographic Information System 或 Geo-information System，GIS）有时又称为“地学信息系统”或“资源与环境信息系统”。它是一种特定的十分重要的空间信息系统。它是在计算机硬、软件系统支持下，对整个或部分地球表层（包括大气层）空间中的有关地理分布数据进行采集、储存、管理、运算、分析、显示和描述的技术系统。这里，“地理”二字并非指地理学，而是广义地指地理坐标参照系统，也即按地理坐标来组织空间数据。GIS 处理、管理的对象是多种地理空间实体数据及其关系，包括空间定位数据、图形数据、遥感图像数据、属性数据等，用于分析和处理在一定地理区域内分布的各种现象和过程，解决复杂的规划、决策和管理问题。GIS 的概念框架如图 1—1 所示。

通过上述的分析和定义可得出 GIS 的如下基本概念：

（1）GIS 的物理外壳是计算机化的技术系统，它又由若干个相互关联的子系统构成，如数据采集子系统、数据管理子系统、数据处理和分析子系统、图像处理子系统、数据产品输出子系统等，这些子系统的优劣、结构直接影响着 GIS 的硬件平台、功能、效率、数据处理的方式和产品输出的类型。

（2）GIS 的操作对象是空间数据，即点、线、面、体这类有三维要素的地理实体。空间数据的最根本特点是每一个数据都按统一的地理坐标进行编码，实现对其定位、定性和定

量的描述。这是GIS区别于其它类型信息系统的根本标志，也是其技术难点之所在。

(3) GIS的技术优势在于它的数据综合、模拟与分析评价能力，可以产生常规方法或普通信息系统难以得到的重要信息，实现地理空间过程演化的模拟和预测。

(4) GIS与测绘学和地理学有着密切的关系。大地测量、工程测量、矿山测量、地籍测量、航空摄影测量和遥感技术为GIS中的空间实体提供各种不同比例尺和精度的定位数据；电子速测仪、GPS全球定位技术、解析或数字摄影测量工作站、遥感图像处理系统等现代测绘技术的使用，可直接、快速和自动地获取空间目标的数字信息产品，为GIS提供丰富和更为实时的信息源，并促使GIS向更高层次发展。地理学是GIS的理论依托。有的学者断言，"地理信息系统和信息地理学是地理科学第二次革命的主要工具和手段。如果说GIS的兴起和发展是地理科学信息革命的一把钥匙，那么，信息地理学的兴起和发展将是打开地理科学信息革命的一扇大门，必将为地理科学的发展和提高开辟一个崭新的天地"①。GIS被誉为地学的第三代语言——用数字形式来描述空间实体。

GIS按研究的范围大小可分为全球性的、区域性的和局部性的；按研究内容的不同可分为综合性的与专题性的。它们间的关系如图1—2所示。同级的各种专业应用系统集中起来，可以构成相应地域同级的区域综合系统。在规划、建立应用系统时应统一规划这两种系统的发展，以减小重复浪费，提高数据共享程度和实用性。

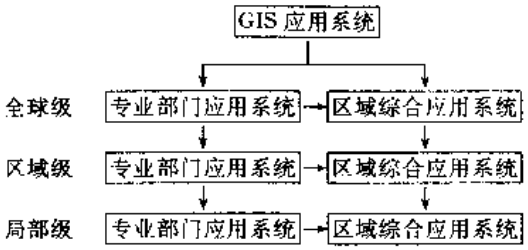


图1—2 GIS应用系统的类型与关系

第1·2节 GIS的发展

过去是现在发展的起点，回顾过去是为了未来的发展。当1960年加拿大测量学家诺基尔·汤姆林逊（Roger F·Tomlinson）博士提出，要把地图变成数字形式的地图，便于计算机处理和分析这一个新颖的思想时，极少有人能看到计算机在GIS这个新的领域中的应用前景。之后于1963年，他又首先提出了地理信息系统这一术语，并建立了世界上第一个GIS——加拿大地理信息系统（CGIS），用于自然资源的管理和规划。同样，当时也很少有人能预见到这会成为一个新兴产业的名称标志。然而，在过去短短的三十余年中，地理信息系统由一个诱人的思想发展成为一个新兴的产业和学科——地理信息产业和地理信息学或地学信息学（Geomatics或Geo-informatics）。先驱者的足迹，开创了地理信息系统产业化的道路。下面简单回顾一下GIS的发展历程和它在中国的发展概况。

一、GIS的发展简史

随着50年代电子计算机科学的兴起及其在航空摄影测量与地图制图学中的应用，以及政府部门对土地利用规划与资源管理的需求，逐渐产生利用计算机来收集、存储和处理与空间地理分布有关的图形和属性数据，并借助计算机的分析和输出为管理和规划决策服务，这就导致了计算机化地理信息系统的问世。1956年奥地利测绘部门首先利用计算机建立了地籍数据库，随后各国的土地管理和测绘部门也逐步发展土地信息系统（LIS）用于地籍管

① 陈洪经、冯惠琳：信息地理学初探，资源与环境信息系统国家重点实验室年报，1992。

理。从1963年开始投入人力和财力开发的世界上第一个GIS（CGIS），至1971年完成，历时八年。稍后，美国哈佛大学研究出SYMAP系统软件。但由于当时计算机技术水平不高、存储量小，使得GIS带有较多的机助制图色彩，地学分析功能少且极为简单。在这一期间出现了基于栅格数据格式和矢量数据格式的GIS，但图形交互功能缺乏。此间，许多与GIS有关的组织和机构纷纷建立。例如，1966年美国城市和区域信息系统协会（URISA）成立，1969年又建立起州信息系统全国协会（NASIS），国际地理联合会（IGU）于1968年设立了地理数据收集和处理委员会（CGDSP）。这些组织机构的建立为传播GIS知识，发展GIS技术起了重要的推动作用。到60年代末期，GIS的应用前景已经引起许多政府部门、研究机构和大学中不少思想敏锐人士的关注。主要的关注点集中在什么是GIS，它能够干什么等？一些学术会议专门就GIS问题展开了探讨。因此，可以说60年代是GIS思想和技术方法的探索时期。

70年代是GIS的发展时期。在这期间计算机发展到第三代，大容量随机存取设备——磁盘的使用，为空间数据的录入、存贮、检索和输出提供了强有力的手段。用户屏幕和图形、图像卡的发展增强了人机对话和高质量图形显示功能，促使GIS朝着实用化方向迅速发展。一些发达国家先后建立了许多专业性的土地信息系统和地理信息系统。例如，从1970年到1976年，美国地质调查局建成了50多个专业GIS，为地理、地质和水资源等的空间信息管理服务。加拿大、联邦德国、瑞典、日本等国也相继发展了自己的专业GIS。同时，一些商业公司开始活跃起来，软件在市场上受到欢迎。据统计70年代约有300多个系统投入使用。1980年美国地质调查局出版了《空间数据处理计算机软件》的报告，基本总结了1979年以前世界各国GIS的发展概况。此外，D. F. Marble等拟订了处理空间数据的计算机软件登录的标准格式，对全部软件作了系统的分类，提出GIS发展研究的重点是，空间数据处理的算法，数据结构和数据库管理系统这三个方面。在这期间，由于西方国家环境浪潮的兴起，导致GIS在自然资源开发、环境保护和土地利用规划等领域的应用。同时，许多大学和研究机构开始重视GIS软件设计和应用研究。例如，美国纽约州立大学布法罗校区创建了GIS实验室，后来在1988年发展成为包括加州大学和缅因州大学在内的由美国国家科学基金会支持的国家地理信息和分析中心（NCGIA）。说明GIS技术已经受到政府部门、商业公司和大学的普遍重视，成为一个引人注目的领域。

80年代是GIS在理论、方法和技术上取得突破与趋向成熟的阶段。此期间越来越多的专业杂志创立，并发表GIS的应用和理论、方法方面的文章；一些计算机公司开始向用户介绍和展示GIS软件样版；对GIS有所了解和认识的人日渐增多。此外，计算机硬件技术大为发展，图形工作站推出，微型PC机的性能价格比不断提高，软件开发工具的广泛应用和数据库技术的推广，都对GIS的成熟起着推动作用。这期间，GIS的数据处理能力，空间分析功能，人机交互对话，地图的输入、编辑和输出技术均有了较大发展，并且在许多领域中得到应用。80年代后期，以工作站为平台的GIS软件日益发展。商品化的GIS软件除了具有上述功能之外，还能为用户提供良好的用户界面，多种数据格式转换的接口，计算机网络通讯，有的甚至还提供一种“宏语言”作二次开发之用。在实用方面，从解决基础设施的规划（如道路、输电线）转向更复杂的区域开发问题，例如土地的多目标规划，城市发展战略研究及投资环境决策研究等。在这些研究工作中，地理因素成为不可缺少的依据。GIS与卫星遥感技术相结合，开始用于解决全球性的难题，如全球沙漠化、全球可居住

区的评价、厄尔尼诺现象及酸雨、核扩散及核废料，以及全球海面变化与监测等。在此期间推出了不少商品化的GIS软件，代表性的有ARC/INFO，TIGRIS，GENAMAP，IGDS/MRS，Microstation，SYSTEM9等。有的可在工作站和微机两种版本上运行。

进入90年代以来，由于高性能低价格的工作站和微机充斥市场，计算机网络技术的推广应用，以及UNIX操作系统，Window或X-Window开发工具和并行处理机技术的不断发展，使GIS发展成为具有开放性、信息资源共享和支持多种硬件平台（多层次）的高技术产品，使用环境更为方便。为了满足GIS高速增长和广泛使用的需要，人才培养成为关键。讲授GIS的课程或培训班大受欢迎。据不完全统计，到1991年底约有两千多所高等学校开设了或者即将开设GIS的课程。出版的教材或专著已有数十种。以GIS为主题的国际学术活动十分活跃。如加拿大每年召开两次GIS学术会议，在内容上分别侧重于理论研究和应用研究，美国每年有多次综合性的和行业性学术会议都涉及到GIS的内容。特别引人注目并有重大影响的国际GIS学术会议是"International Symposiums on Spatial Data Handling"，从1984年起每两年举行一次。在过去五次大会上作主题报告的专家和题目如下：第一次1984年，Roger Tomlinson，"地理信息系统——一门新的前沿学科"(Geographical Information Systems——A New Frontier)；第二次1986年，Jack Dangermond，"GIS的发展趋向与经历"(GIS Trends and Experiences)；第三次1988年，Michael W. Dobson，"国际商业出版界的数字地图学"(Digital Cartography in the World of Commercial publishing)；第四次1990年，Michael Goodchild，"空间信息科学"(Spatial Information Science)；第五次1992年，David Rhind，"GIS的信息基础结构"(The Information Infrastructure of GIS)。第六次1994年的会议在英国爱登堡大学召开，论文集共收录了71篇论文，但没有主题报告，3维和4维GIS，数据质量，多媒体技术和面向对象模型等是讨论的热点。由此可略知GIS发展的主线。在1990年北京第二届国际GIS会议上，一位日本科学家在谈到GIS的特殊意义时风趣地说"彬彬有礼、智能、微笑"(Gentlemen like • Intelligence • Smile)，这三个英文词第一个字母的缩写也是GIS，这也说明GIS工作者应具有的素质。

90年代世界各国对GIS科技的开发和研究，主要集中在如下一些方面：空间信息分析的新模式和新方法；空间关系及数据模型，包括与时间有关的四维数据结构；人工智能和专家系统的引入；显示与输出技术；GIS的效益评价；保密性等。总之，国际上的GIS已从具备储存、分析评价、查询检索、自动制图等一般功能，向信息采样自动化、多样化、多功能的综合分析评价模型，智能化的专家系统，充分与遥感遥测技术直接连接进行信息更新，以及规范化、标准化和信息充分共享等方面发展。

二、中国GIS的发展概况

我国GIS的起步较晚，但发展较快、势头良好。大体上可分为下列几个阶段。

第一个阶段是准备阶段。70年代初期我国开始在测量、制图和遥感领域应用计算机技术。随着世界遥感技术的发展，我国于1974年开始引进美国陆地卫星图像资料，开展了遥感图像处理和解译工作。1976年召开了第一次遥感技术规划会议，形成了遥感技术试验和应用研究的蓬勃发展新局面。此后，开展了各种区域性航空遥感试验研究，获得了大量多层次、多时相、多应用目标的综合地学信息。没有基础理论指导的技术有可能是盲目的技术，没有新技术武装的理论有可能是空洞、落后的理论。遥感在地学各个领域中的推广应用，地学实验技术和理论研究的需要，以及地球科学、环境科学研究向宏观和微观的战略

发展，为地理信息系统的发展开辟了道路。1978 年国家计委在安徽黄山召开了全国第一届数据库学术讨论会。同年 10 月在浙江杭州召开第一届环境遥感学术讨论会，在会上我国著名的地理学家、地图学家陈述彭院士发出了在我国开展地理信息系统研究的倡议。从此，地理信息系统研究的舆论准备、组织建设和可行性实验，在全国由点到面逐步地开展起来。

第二个阶段是试验起步阶段。进入 80 年代之后，我国在 GIS 的研究和应用方面取得了实质性的进展。在理论探索、规范探讨、实验技术、软件开发、系统建立、人才培养和区域性、专题性试验等方面都积累了经验，取得了突破和进展。例如，在二滩、渡口地区建立了我国第一个信息系统模型以及全国范围的空间数据库试验方案，建成了 1∶100 万国土基础信息系统和全国土地信息系统，1∶400 万全国资源和环境信息系统，1∶250 万水土保持信息系统。在专题性信息系统试验研究方面，如黄土高原信息系统，洪水灾情预报与分析系统、全国树木资源与环境信息系统，矿产资源数据库，煤种资源数据库，农业资源数据库等。用于辅助城市规划的各种小型信息系统在城市建设和规划部门获得了认可。

在学术交流和人才培养方面，此期间也得到较大发展。在国内召开了多次关于 GIS 的国际学术讨论会。1980 年 1 月，中国科学院遥感应用研究所成立了第一个地理信息系统研究室。1985 年 2 月，在中国科学院和国家计委领导下建立了“资源与环境信息系统国家重点实验室”。南京大学、北京大学、武汉测绘科技大学等高校在计算机辅助制图和地理信息系统的教育、科研和人才培养方面先行一步，并取得了可喜的成绩。

同时，信息系统规范化标准化的研究工作也有计划地开展。1983 年 11 月，国家科学技术委员会组建了资源与环境信息系统国家规范研究组，在充分调查分析国内外历史和现状的基础上，于 1984 年 9 月提出了“资源与环境信息系统国家规范研究报告”。上述各个方面卓有成效的工作，使我国逐步形成了自己的建立和发展 GIS 的技术路线、原则和政策，避免了一些盲目分散或过度集中的通病；使具有实际效能的区域性 GIS 和专业（题）性 GIS 得到并行发展，还注意发展微机 GIS，以便于推广应用。

第三个阶段是发展阶段。从 1986 年到 1995 年前后，我国的 GIS 事业随着社会主义市场经济的发展走上了全面、迅速发展的阶段。在此期间，全国许多行业部门和部分省区积极发展各有特色的 GIS 系统，在理论和应用研究、人才培养等方面取得了丰硕成果。由于沿海、沿江经济开发区的发展，土地的有偿使用和外资的引进，急需 GIS 为之服务，推动了 GIS 的发展。上海、北京、天津、深圳、海口、三亚、常州、沙市等大中城市都在紧张地建设城市 GIS 系统。北京农业大学等积极开展土地和农业信息系统的研究。用于农作物估产和灾害监测的遥感与 GIS 运行系统在建设之中。矿区是一种特殊的地理区域，其地理空间要素和社会经济要素的内容广泛、综合、复杂，而且变化迅速。矿区地理信息系统(MGIS)或称矿区资源与环境信息系统(MREIS)所面向的地理空间包括地面、地下以及大气层的多层次三维空间，甚至要求建立动态的时空一体化的四维 GIS。中国矿业大学和西安煤田航测遥感局等在其理论、技术方法和应用研究，以及人才培养方面做了许多工作，取得了可喜的成果。

总之，我国的 GIS 技术在研究基地和队伍建设、技术研究和成果应用、前沿领域开拓以及参与重大的全球性研究计划，如国际空间年、国际减灾十年等方面服务的进展迅速，已大大缩短了同世界先进国家的差距，并已跻身于国际先进行列。

1992 年 10 月，联合国经济发展部 (UNDESD) 在北京召开了城市 GIS 学术讨论会，并建议筹建中国 GIS 协会。经过一年多的筹备工作，于 1994 年 4 月 25～27 日在北京召开中

国地理信息系统协会（简称 CAGIS）成立大会暨技术报告会，以指导、协调和推动全国 GIS 事业的发展。

此外，1992 年 8 月在美国纽约州的布法罗市召开了首届 GIS 中国学者大会，海内外 120 多名中国学者与会，共同研讨 GIS 的现状和未来的发展，显示了中国学者对国际 GIS 学科的关心和在理论与应用研究中作出的贡献。

第四个阶段是持续发展、形成行业和走向产业化的阶段（1996 开始）。经过前面十多年的发展，中国 GIS 已在研究和应用上逐步形成行业，开始走向产业化的道路，将成为国民经济建设和社会生活的一种共同需要和普遍使用的工具，并在城市建设，农业、能源等基础建设、环境保护、灾害防治和海洋开发等方面发挥重大作用。其中包含着许多理论、技术方法和应用问题需要研究和解决。

第 1—3 节　GIS 的硬软件环境

GIS 一般由硬件、软件、数据和用户构成，对于一个实用的 GIS 系统缺一不可。反之，它们对 GIS 的发展既有贡献也有约束作用。

一、GIS 的硬件配置

GIS 的硬件配置一般如图 1—3 所示。计算机与一些外围设备联接构成 GIS 的硬件环境。计算机是硬件系统的核心，用作数据的处理与管理。GIS 输出和贮存用的外部设备通常是标准的计算机外围设备，包括打印机、高分辨率显示装置、绘图仪、磁带机等。地图制图或 GIS 系统的输入专用设备，包括数字化仪和/或扫描仪，全站型测量仪器以及数字摄影测量工作站和遥感图像处理系统等。从发展的观点看，附加的硬件配置还应有远程通讯装置。

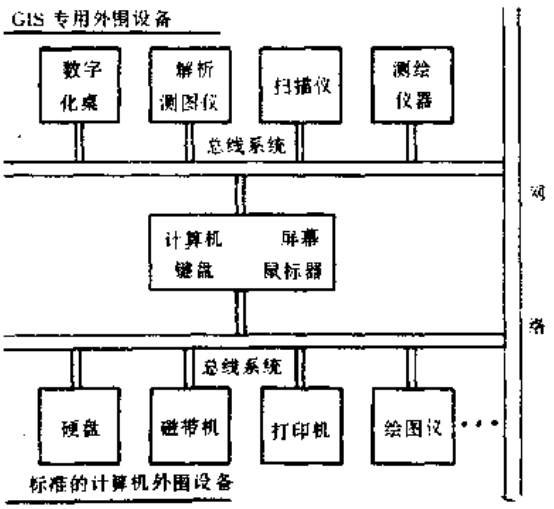


图 1—3　GIS 的主要硬件配置

（一）数据处理与管理设备

GIS 的数据处理、分析和管理设备是电子计算机系统，可以是单独一套计算机也可以是多台机联网使用。计算机及其相关技术的发展日新月异。目前运行 GIS 的计算机，包括大型、中型、小型机，工作站和微型计算机，其中以工作站和微型机为主流。现在，微机和工作站的处理速度相当于十年前的巨型机。1992 年推出的“奔腾”（P5）微机的性能已接近亿次机。微机的内存已进入 GB（10^9）级。计算机处理速度的极大提高，使得它处理大量地理数据并提高空间分辨率成为可能；允许选择更加复杂、精确的应用模型，以进行复杂的分析评价和科学决策；还可能使用更智能化、自动化的算法，并使图形图像的显示，成果的目视化更加逼真、丰富多彩。最近，3—D 立体镜已商品化，SGI、SUN 等公司开发出实时三维图形卡，从而给三维仿真、虚拟现实的实用化开辟了新路。这在采矿工程、地球物理、水文学、景观生态学及城市景观模拟等领域起的作用更大。

工作站的发展时间虽短，但前景很好，有可能成为 GIS 的主流用机。配有大屏幕、高

分辨率图形终端（通常是19"1152×900或20"1280×1024），以及专门的图形加速卡的工作站系统特别适用于GIS中的图形图像处理。工作站在GIS中的应用将促使多用户网络化的硬件结构模式流行起来。这样，便于将计算机服务器、文件服务器、工作站和分享式终端集线起来，任一用户可随机存取和使用网络上的各种资源，可以达到每秒兆字节以上的数据传输速度。还便于实现数据远程通讯。

存贮装置是计算机系统的重要部分，这些年来发展很快。磁盘和磁带一直是计算机的主要存贮装置。磁盘分硬盘和软盘两种，硬盘的存取速度和存贮容量比软盘大得多。微机用的硬盘容量已达200MB以上，工作站的硬盘已超过1GB的容量。高密软盘的容量，5吋盘已达1.2MB，3.5吋盘可达2MB。光盘存贮技术近年来已商品化，其存贮容量可达几GB，如可读光盘WORM（写一次，读多次）CD-ROM有2GB的容量，可存贮约100张地图的内容，但它不能擦和重写。而可擦写光盘也已投放市场，如NEXT计算机，它能以很高的数据密度存取数据。

平行处理和RISC（减少指令表计算）处理技术将在GIS硬件系统中发挥重要作用。能模拟人脑思维的神经网络计算机即将实用化。人们正在研究超导和光路计算机，以代替现在的“电子”计算机，如获成功，将使计算技术和信息技术进入一个新时代。

总之，信息技术当今发展的几个明显趋势是：不同种类机型的界线逐渐模糊，顾客/服务环境的流行，多媒体技术的发展和应用以及计算机与通讯技术的融合。这些发展趋势将促使GIS向不依赖于平台的、适应网络环境的、能处理多种类型数据的方向发展，对空间数据的表达更加全面、直观和生动逼真，从而使GIS更容易普及，被社会广泛接纳。

（二）数据采集与输入设备

GIS有大量图形图像数据以及各种文字、数字数据需要经常地输入和不断地更新，与之相适应的数据采集与输入设备也是多种多样的。当今使用较多的是数字化仪、扫描仪、解析和数字摄影测量设备以及电子速测仪、GPS接收机等测量仪器。

1. 图形数字化仪

目前，图形数字化仪是图形输入的主要设备。市场上出售的数字化仪的规格有A_0、A_1、A_2、A_3等幅面，最大的是A_{00}。典型使用的是A_0规格，幅面为1m×1.5m。对于多数使用而言，A_1的也能满足要求，因为它能包容一幅0.5m×0.5m大小的地形图或矿图。而A_3幅面的数字化仪多用于试验或培训工作，因为数字化一幅地图往往需要移动、定向几次，作业麻烦。

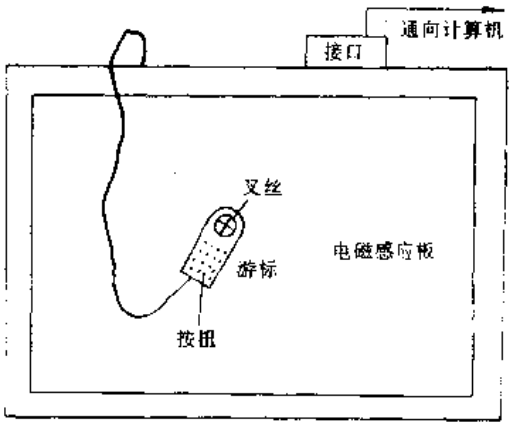


图1—4 图形数字化仪基本结构图

图形数字化仪由电磁感应板、游标和相应的电子电路组成，如图1—4所示。当使用者在电磁感应板上移动游标到图件的指定位置，并将十字叉丝的交点对准数字化的点位时，按动相应的按钮，数字化仪则将此时对应的命令符号和该点的位置坐标（x，y）值排列成有序的一组信息，然后通过接口（多用串行接口）电路传送到主计算机。有些数字化仪还能将原始数据进行一些预处理，再送到主计算机。地图数字化的精度受数字化仪本身的精度、

数字化方式、操作者的经验和心理状态等因素的影响。手扶数字化仪的跟踪、对准精度通常为 0.2mm 左右，而多数数字化仪的标称分辨率为 0.001″或 0.025mm，一般能满足要求。

2. 图形扫描仪

手扶跟踪数字化仪的速度慢、精度不稳定，工作效率低。应用图形扫描仪配合相应的软件能进行图形、图像的自动扫描输入。按工作原理分，扫描仪有光电式、摄像式和 CCD 器件等多种。按辐射分辨方式分，则有二值扫描仪、灰度值扫描仪和彩色扫描仪。二值扫描仪每个像素占 1 比特，取值 0 或 1，用于线划图和文字的扫描数字化。灰度值扫描仪每个像素占 8 比特，分成 256 个灰度级，可扫描图形、影像和文字等。彩色扫描仪通过滤色片将彩色图件或像片分解成红绿蓝三个波段，每个波段占 1 比特，每个像素占 3 比特，加上黑白图像每个像素共占 4 比特。按结构方式分，扫描仪可分为滚筒式扫描仪、平台式扫描仪和 CCD 摄像扫描仪。还有透光和反光扫描之分。CCD 摄像扫描仪是在摄像架上对图件进行中心投影摄影来取得数字影像的。

扫描仪按照扫描方式又可分为栅格式扫描仪和直接沿线划扫描的矢量式扫描仪。栅格扫描仪扫描得到的图像信息，需要经目标识别和由栅格到矢量的转换。目前，全自动的栅格扫描仪还处于研究阶段，而人机交互的半自动扫描仪已商品化，如 HP9190A（平台式）、Tek/4991（滚筒式）型等。矢量扫描仪是直接跟踪图件上的线划，并直接产生矢量数据。目前多采用激光扫描。

扫描仪的分辨率有两种表示：像素大小（10～100μm）和每吋的网点数（dpi）。例如 400dpi、1000dpi 或 64μm×64μm，25μm×25μm。

3. 解析和数字摄影测量设备

解析测图仪也可以作为 GIS 的图形数据采集的外围设备，而不直接与绘图机联机作业。如蔡司厂的 P3，威特厂和 Prime 公司生产的 System9P 工作站（解析测图仪是 BC3）。模拟立体测图仪经过改造后，利用一定的接口和软件也可与计算机联接而成为机控测图的输入系统。新一代数字摄影测量工作站和各种遥感图像处理系统，则更便于作为 GIS 的数据采集和更新设备，或者与 GIS 系统融为一体。

4. 其它测量仪器

各种光电测距经纬仪、电子速测仪（俗称全站仪）、GPS 接收机等测量仪器，通过各种电子数据记录器、数据终端能以数字形式自动记录野外测量数据，并转输到主计算机中，实现 GIS 数据采集。

（三）输出设备

GIS 是一种交互式图形图像分析处理系统，图形的显示和输出是重要的组成部分。图形显示终端用于图形的交互式输入、编辑、分析、处理和输出。各种绘图仪和打印机用于图形和文字、表格的输出。

1. 图形显示终端

GIS 要求图形显示终端有丰富的软件支持，如国际图形软件标准 GKS、PHIGS、X 等软件包，还要有一定的显示幅面、分辨率、色彩和图形平面存贮器等。如微机上常用的 VGA 卡或 TVGA 卡及其适配器可用于数字测图系统或 GIS 系统，同时要求配有高分辨率的图形显示终端。在微机和工作站系统中，图形终端和相应的图形卡构成了一个丰富的可选资源。如 Sun 工作站配置的 GX 卡或 GS 卡能起到图形加速作用，减轻主机的负荷。

立体显示终端是最近的技术成果，它的出现为三维仿真、虚拟现实的实用化开创了新局面。它在制图和摄影测量中能起重要作用，用它可以进行人机交互编辑、修正。目前有多种系列、型号的显示终端，如 Tektronix 公司生产的 4128、4335 型，与 IBMPC 机兼容的 SGS—430 三维转换器等。应指出，在选用显示终端时要注意是否被 GIS 软件支持或兼容。

2. 绘图仪

矢量绘图仪是目前我国使用最多的图形输出设备，并有滚筒式和平台式两种方式。这类绘图仪的输出质量主要取决于控制马达的步进量。绘制字母和符号等复杂图形的功能主要由绘图软件决定。

栅格式绘图设备目前主要是行式打印机，主要供输出草图之用。激光打印机是一种阵列式打印机，分辨率可达 600～1200 点/吋。它解决了行式打印机精度差的问题，具有速度快、精度高、图形美观等优点。某些阵列打印机带有三色色带，可打印出多色彩的图件。目前它未能用作主要图形输出设备的原因是幅面偏小。

喷墨绘图仪是一种精度高质量优的绘图设备。它按栅格数据的像元值来控制喷射到纸张上的墨滴大小，控制功能取决于静电电子数目。彩色喷墨绘图仪能产生 400 多种颜色，分辨率可达每吋 300～400 点（10～12 点/毫米），但目前价格昂贵。价格适中且分辨率较低的喷墨绘图仪为 3～5 点/毫米，绘图质量能满足一般要求。

3. 打印机

GIS 还有多种表格、文字的数据需要输出，可利用多种打印机完成。打印机的类型有针式打印机、激光打印机、液晶打印机等。如上所述，有些打印机还能打印图形，或具有彩色功能。

二、GIS 软件的特点、基本要求和组成

GIS 软件是系统的核心，关系到其性能的优劣、功能的多少。GIS 的外观表现为计算机的硬软件系统，而其内涵则是由计算机程序和数字化的地理数据组成的地理实体空间信息模型。目前，GIS 软件开发面临的主要问题是成本高，开发周期长，远远跟不上计算机硬件的迅速发展。现在世界上已开发的 GIS 软件有几百种，使用比较广泛的也有几十种之多，使广大用户有眼花缭乱之感。通过本节的描述，拟使读者对 GIS 软件的特点和基本要求有一较综合、系统的理解。

（一）GIS 软件类型

从用户的角度看，GIS 软件系统可分为如下的两大类：

（1）实用型地理信息系统或称面向项目的信息系统（Project—Oriented Information System）。这是一类有具体应用目标，为特定区域或专业（题）的管理决策服务的实用 GIS。有些实用 GIS 是介于二者的区域性专题信息系统。

（2）工具型地理信息系统（GIS—Tools）或称面向管理的地理信息系统（Management—Oriented GIS）。这类系统一般没有具体的应用目标，基本上是一种系统支撑软件供建立实用的 GIS 之用。它具有对计算机硬件适应性强、数据管理和操作分析效率高、功能较完善，且容易扩展和商品化等优点。其最大特点是在建立、维护和更新数据等工作时都遵循一定的规范和标准，系统的数据字典和用户手册等文档完备，因此这种系统能够与处于同一层次或不同层次上的其它系统构成体系，并与它们实现数据及功能的共享和交流，是 GIS 的发展方向。它一般包括系统的公用部分和专门应用部分。公用部分的内容是其下面各层

次的信息服务对象所共同涉及的方面如基础性的空间数据和属性数据（如基础地形图数据），一般性的查询、检索、制图和分析工具等。专门应用部分是指各个应用部门、单位所需的特定数据和工具。可见，面向管理的GIS的优点在于：使用频率最高的数据和工具只在较高层次的系统中存储和开发一次，便于公用部分的统一维护和数据更新，能节省开发力量和系统资源；各个子系统之间的相同部分因遵循统一的规范标准，可通过移植得到，节省了开发力量，也便于实现系统之间的通讯和数据共享。

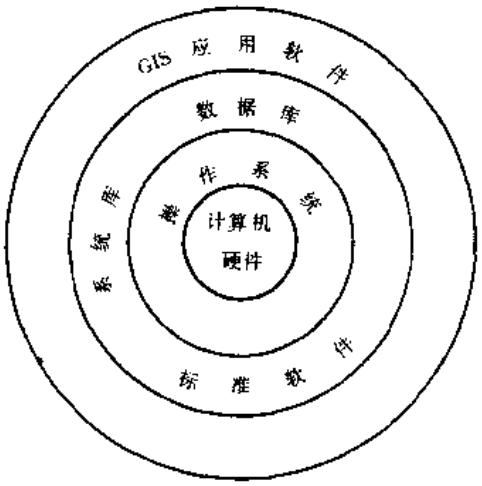


图 1—5　计算机系统中的GIS软件层次

应该指出，如果面向项目的GIS确实遵循一定的规范和标准，则能够在项目完成之后通过一定的机制演变成为面向管理的工具型GIS。

无论属何种GIS软件，它们在计算机软件层次中属于外层（见图1—5）。外层通过内软件层来完成用户需求的特定任务。GIS应用软件还可再划分为基本功能软件包、应用软件包和一些用户接口、通讯软件。

（二）GIS软件开发的基本原则

1. 实用性是软件开发的核心和出发点

由于计算机技术发展水平的限制，人们对GIS应用系统模型重要性的认识不足，地理空间实体的复杂性，以及用户的广泛性、多样性等原因，使至今的GIS软件大都是以计算机数据库为其核心，而应用模型在整个系统中仅处于从属地位。或者说，软件开发者对软件功能、数据结构的关注大于可用性。这样，就使现有的软件实用性差，主要表现在：使用软件系统需要专业技术人员，并经过培训，因而影响到GIS技术的普及；大多数软件系统有各自的系统环境和结构；使用户利益未被保护，功能得不到充分发挥，难以融进部门的信息流中去；多数软件系统都使用了不同的加密方法，难以适应网络环境下的实时在线运行服务。

众所周知，人类对客观世界的认识、利用和改造过程，经历了由现实世界到观念世界的转化。观念世界的建立是通过人类对现实世界的认识与抽象，即各种专业研究和系统分析。产生和发展出GIS软件工具和应用系统所需的概念模型，属GIS调查研究的范畴；而由观念世界到信息世界的转化，要在计算机硬件的支持下，靠适当的软件工具予以实现。类似地，它相应于根据概念模型利用软件工具建立GIS应用系统的过程；由信息世界返回客观世界的过程，实质上是GIS系统的应用过程，即用以指导、帮助人类利用和改造客观世界的实践活动，并受这些活动及其效果的检验，进而加深对客观世界的认识，使上述转化过程达到一个新水平，如图1—6所示。因此，某个GIS应用系统是客观世界中相应系统经由观念世界到信息世界的映射。这种系统的应用反映了人类对客观世界利用、改造的能动作用，是GIS技术产生社会、经济效益的关键所在。为此，加速以应用模型或模型库为中心的GIS软件开发研究，是应该努力实现的重要方向。

2. 充分认识和掌握地理实体的空间和动态特征

地理信息系统与其它描述性信息系统相区别的基本点在于，GIS处理的信息流具有鲜

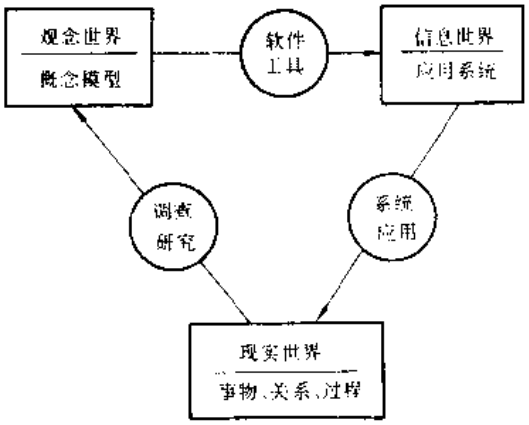


图 1-6 应用系统的形成和目的

明的空间和动态变化特征。具体来说，地理空间数据有如下基本特征：

(1) 空间位置数据，即在某一坐标系统下标识地理实体空间位置的数据。按地理实体的空间分布特征，可区分为点状、线状、面状和区域四类，相应地可用点、弧、多边形和数据平面四种空间数据类型来表示。

(2) 拓扑关系数据，表示各种空间地理实体或现象相互关系的数据。其关系可分为：包含、邻接和联通三类。显示或隐含地存储这些关系数据，是分析空间格局和过程的基础。但是至今，人们对空间关系的理解和表达形式还没有一个完整的、确定性的框架，因而限制了GIS 功能的发挥和不同类型系统的出现。

(3) 属性数据，或称非空间数据。它与空间位置数据在逻辑上紧密相连，是表示空间地理实体特征和地理意义的数据，从而使空间数据具有空间关系和对象描述这两重性。属性数据可分为定性和定量两种。前者包括事物的名称、类型、特征等；后者一般是表示事物的数量和等级。任何地理实体至少有一个属性数据。GIS 的分析、检索和表示主要通过属性的操作运算来实现。在一些实用系统中，属性的数据量远大于空间位置的数据量。可见，属性的分类、评定指标对系统的功能有很大影响。在工具型 GIS 中，更必须有效地逻辑联接空间和非空间数据。

(4) 生命周期。空间数据的生命周期长短差别很大，如土壤、地形等数据的生命周期一般很大；而有些数据，如土地利用、庄稼的生长、井下巷道的进展等的变化往往很快，需要及时更新。当今，遥感和 GPS 技术是更新地理数据源的主要技术手段，因而 GIS 与遥感和 GPS 技术相结合是发展方向，以获取多光谱、多时相、多分辨率的地球表层数据，便于从不同角度去分析地球环境要素的物理、化学和电磁特性，进而更深刻地理解其信息机理。

与此相关的是地理空间数据的动态性，即空间数据和属性数据随时间相应独立地变化。据此，合理地设计数据结构、数据模型和数据管理系统十分重要。

(5) 数据质量和精度。质量是一切产品的生命。进入 GIS 数据库的空间数据来源不同、时间不同，地图的比例尺不同。数据的质量和精度过高或偏低都不妥。GIS 中数据的质量特征一般用如下指标来描述或评定：数据渊源，精度（包括位置精度和属性精度），现势性，逻辑一致性，完整性，可靠性和安全等级。在开发 GIS 软件时应研究确定：进入数据库的各种数据质量和精度的方法；各种应用模型对数据误差的传播规律及其控制机制；根据对输出结果的要求，求定输入数据的误差限等。

3. 良好的用户界面

对 GIS 运行化的要求包括四个要素：数据质量、用户界面、应用模型及成本费用分析。其中，用户界面是关键。对用户来说，用户界面就好象系统本身，它决定了系统使用的效率和方式。用户关心的是系统能解决什么问题，怎样解决，而对系统中的一些技术问题却往往不太关心。

（三）GIS 软件的基本组成

通用 GIS 软件系统既独立于各种具体的应用系统，又为这些应用系统进一步扩展提供有力的技术支持，因此它是加速 GIS 产业发展和应用的关键。通用 GIS 软件工具的功能越齐全、性能越可靠，它的应用范围就越广泛，产生的效益就越显著。

不同规模和水平的 GIS 通用软件工具，其内容和技术水平可有很大的差别。就 GIS 的计算机硬件规模而言，目前大体上可分为微机型和工作站或中小型机型两类。在发展 GIS 软件系统时要注意不同机型和水平的配套，为此，要尽可能采用一致的数据结构或格式，使不同规模的软件工具能联网运行，以便充分有效地实现系统资源的共享和利用，减少不必要的重复浪费。

在设计和建造 GIS 软件系统时，应注意遵循如下原则：

（1）系统地应用计算机科学中已发展起来的技术和方法，如软件工程、数据库理论、算法、计算机图形学、人工智能以及自然语言处理。

（2）综合使用相关学科中较成熟的科技方法，包括遥感、数字制图、影像理解、计算机视觉等。GIS 和遥感的有效结合是 GIS 发展的一个必然趋势和重要内容，特别是随着遥感技术的不断发展，人们可以动态地从遥感图像中提取大量有用和现势性好的信息，这是快速更新 GIS 数据的主要途径。因此，在 GIS 软件中一般应提供这方面的功能。

（3）在大型数据库中，对复杂空间目标的查询是一件很困难的事情。查询的效率在很大程度上取决于存贮的数据量、数据编码方法和文件结构设计。对这个问题需要认真研究解决，尽可能避免使用低效率的搜索技术。

（4）方便用户使用，易于掌握，并且尽可能满足多用户的需要。

GIS 软件工具一般包括如下主要部分或模块（图 1—7）：

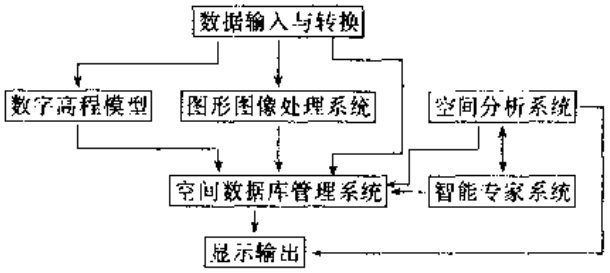


图 1—7　GIS 软件系统的主要功能模块组成

1. 空间数据库管理系统

它是 GIS 软件工具的核心部分，统一管理属性和空间数据，具有初始化、输入、更新、删除、检索、变换、量测、维护等功能，并为其它模块提供基本图形图像支持工具和接口。

2. 图形图像处理系统

该系统不仅要包括通用图形处理功能，如图形数据输入、编辑、构建拓扑关系、地图整饰、图幅接边等，而且还应具有图像处理功能，如几何纠正、滤波、边缘提取、图像分类等。图形图像处理系统应作为一个整体，处于同一界面之下，以实现 GIS 和遥感的完全结合。

3. 数字高程（地面）模型模块

数字高程（地面）模型是一种特殊的数据模型，在地球科学和区域规划设计中有广泛的用途，一般将它设计成一个单独的模块。

4. 空间数据分析系统

空间数据的处理、分析是GIS软件的又一重要内容和特色所在。关于空间数据分析可分为三个不同的层次：一是简单的空间搜索、空间叠加，如缓冲区分析、网络通路、资源分配、多边形叠加等，可称为传统手工操作的计算机化；二是空间格局的关系及其描述，包括对空间格局在数量上的描述和定性的趋势分析，如空间目标的聚散度分析；三是空间模拟，如空间过程机理、空间动态模型、预测空间格局的发展变化等。

自60年代以来，空间统计分析的技术方法取得了很大发展，如破碎度、离散度、优势度、分数维等特征值表述，自相关、遥相关等相关方法，profile、趋势面等空间过程，以及点模式、空间统计插值、空间推理等方法。然而，由于GIS技术应用目标（任务）的复杂性，计算机技术发展水平的限制，以及人们对空间统计分析理论和技术方法的掌握还不够，因而这些方法还基本上没有被地理信息系统所接纳，在空间分析系统的研建方面任务艰巨，内容丰富。

5. 智能专家系统工具

该系统工具以人工智能为基础，是通用的GIS软件工具，它具有组织和使用知识及不充分、不准确数据，模仿专家的思维、推理，进行分析和解决问题的能力是一个有待发展和令人鼓舞的领域。

6. 输入输出支持系统

该系统应能实现常用GIS数据格式间的转换，能够支持多种形式的数据输入，如文本、数字、矢量和网格图形数据的输入。在输出功能方面，应具有文本、表格、图形和图像等多种形式数据的功能，包括点阵打印、矢量绘图仪，甚至栅格绘图仪，以及自动分色排版等。

总而言之，计算机技术的不断发展和进步对地理信息系统的发展有深刻的影响。现在计算机硬件发展十分迅速，性能价格比逐年提高，各种型号的工作站和超微机已成为测绘、遥感和GIS的标准硬件。到本世纪末，个人用计算机工作站可以达到如下性能：500～1000MIPS速度的CPU，500MB内存，5GB硬盘和50GB以上的光盘，2000×2000像素的图形显示屏幕，每秒100兆比特的传输速率。

但是，GIS软件的发展却十分缓慢，既费时又费钱。硬件和软件发展的这种不协调造成了所谓"软件危机"。产生这种境况的主要原因是：

(1)编程语言变化太快，给软件研制和用户带来不少困难。计算机高级语言变化、更新很快。早期的ALGOL算法语言不再使用。COBOL，FORTRAN以及近20年来广为应用的新语言Pascal和C也有被取代的趋势。全新的面向对象的C++和Object C语言受到重视和引用，特别是在地理信息系统开发时的引用。而熟悉和运用新语言需要付出大量精力和时间。

(2) 计算机操作系统的发展变化给软件开发带来困难。当今通用的MS—DOS、UNIX和VMS等操作系统已延用了一二十年。而今后的发展潮流如何仍不明朗，主流是UNIX还是WINDOWS NT尚难以断定。研制GIS软件必须建立在一定的操作系统基础上，选择的失误将导致难以挽回的损失。

(3)适用于GIS的数据库管理系统尚未问世。目前，服务于以非空间属性数据为主的数据库系统主要采用关系数据模型，而描述空间信息的数据模型主要有矢量数据模型和栅格数据模型。这些传统的数据模型难以表达复杂的空间地理实体，实现快速查询和动态空间分析。对于图形、图像和属性数据高度集成的地理信息系统数据库，采用什么样的数据库管理

系统？目前人们都期望于以面向对象模型(Object—Oriented Model，OOM)和面向对象程序设计(Object—Oriented Programming，OOP)为基础的面向对象数据模型和数据结构。但是它们能否有效地处理各种类型的空间数据及其复杂关系，管理海量的地理信息，支持带智能的空间分析与决策，还有待实践检验。

第1—4节　GIS的功能与应用领域

GIS为什么这样热门，用途如此广泛？这与它丰富多彩的功能是分不开的。

一、GIS的基本功能

GIS的功能强弱、多少和它的硬软件环境有密切关系。随着计算机科技以及GIS相关学科的不断发展，GIS的功能也将日益增强。

(一)数据采集与输入

有多种多样的数据源可供GIS使用，例如野外测量数据，运用现代定位技术(全球定位系统GPS、惯性测量系统ISS等)所获得的数据，摄影测量与遥感图像数据，现有的图形资料，以及统计调查的文字、数字数据等。不论地理数据信息的形式怎样多样化，它大体上可分为两类：一是地理基础数据或空间数据，如地形、地物的位置，地面和井下测量点的位置，面积和体积数据等；二是属性数据或描述数据、表格数据，如对地形、地物、土地的分类、特征表述，生产统计数据，矿产资源状况数据，社会和生态环境数据等。对上述各种数据的采集和输入GIS中的方式，可归纳为如图1—8所示。

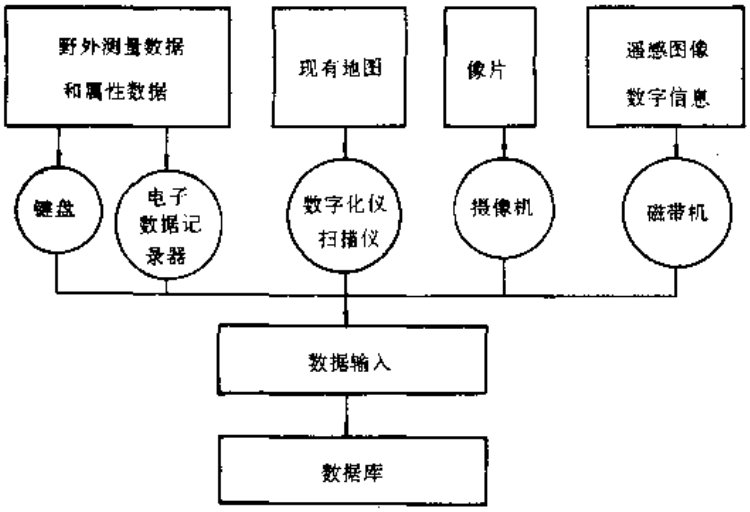


图1—8　GIS数据采集与输入

GIS的数据库不仅可以提供对各种空间数据和属性数据的编辑、处理、统计、分析和评价，还可以按一定的要求建立报表，绘制图形。

(二)空间数据的分析与处理

GIS软件系统具有较强的空间数据分析、管理与处理功能。至于属性数据则更无问题，况且在GIS中属性数据与空间数据往往难以彼此截然分开分析处理。不同的软件系统所具有的功能不尽相同，但通常应具备如下主要功能：

1. 空间数据管理

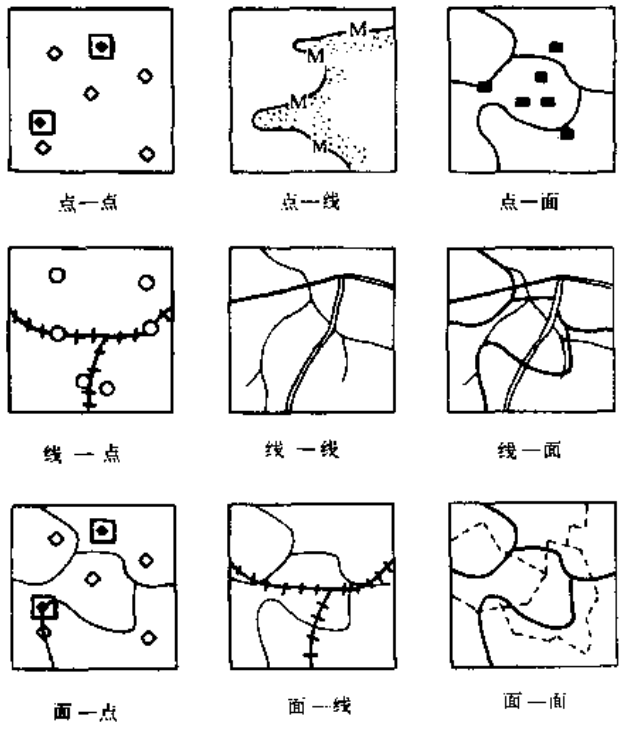


图 1—9 空间单元的双元组合

除了与属性数据有关的 DBMS 功能之外，对空间数据的管理技术主要包括：基础文件的建立、更新、检索、开窗、列表与显示，文件的接边、统计综合，将 x、y 坐标转换成格网或将格网转换成 x、y 坐标等。

2. 拓扑关系与制图综合

空间概念能用几何关系和拓扑关系来度量。基本空间单元及其相关的特征是：点、线（长度、弯曲、方向）、面积（范围、周长、凹凸、连结、重叠）和体积。空间实体可定义为多个空间单元的组合，这种组合可以是相同类型实体的组合，也可以是不同类型实体的组合。以点、线、面实体的二元组合为例，空间单元具有九种不同的二元组合：点—点（点是否相邻？）；点—线（点是否在线上，它们是否相邻接？）；点—面（点是否在面域内？）；线—点（线是否通往某城镇？）；线—线（道路是否穿越河流？）；线—面（河流是否在面状地物内，河流是否与区域边界重合？）；面—点（邮区是否包含某学校？）；面—线（区域是否包含铁路？）以及面—面（区域之间是否相接或重叠？），如图 1—9 所示。

地图综合的内容很多，如线坐标压缩、边界消除、接边、多边形简（细）化等。

3. 地图提取

地图提取与制图综合有密切关联，目前主要的技术方式包括：计算中心点，由不规则分布的数据点生成等值线，生成近似多边形（如梯森多边形），多边形再分类，将 x、y 坐标转换成规则的格网等。图 1—10 表示若干例子。

4. 图幅管理及空间数据处理

在同一地区不同单位、不同时间的数据源难免在地图的比例尺、坐标系、投影方式等存在差别，因此往往需要对一些原始数据作预处理。常见的处理内容和方法如图 1—11 所示。

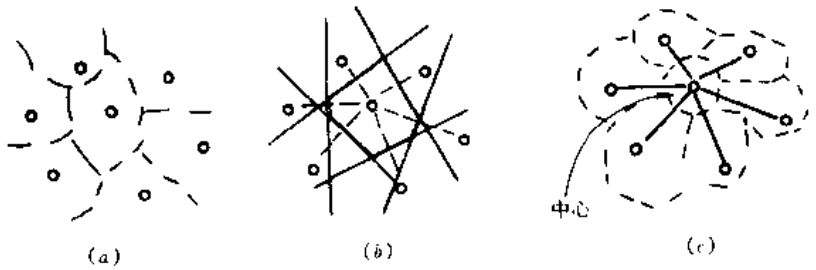


图 1—10 空间目标的置换

a—多边形的中心；b—点至最邻近的面域；c—形成格网

5. 缓冲区生成，多边形重叠和消除

缓冲区的生成与分析是根据数据库中的点、线、面实体，自动建立其周围一定宽度范围的缓冲区多边形（图 1—12），它们是GIS的基本空间分析功能。例如，建立一个电视塔，其有效辐射范围有多大；在铁路线下进行煤炭开采，需要留设多宽的煤柱才能保证运输安全；煤矿开采引起的地表下沉区域或积水面积等。

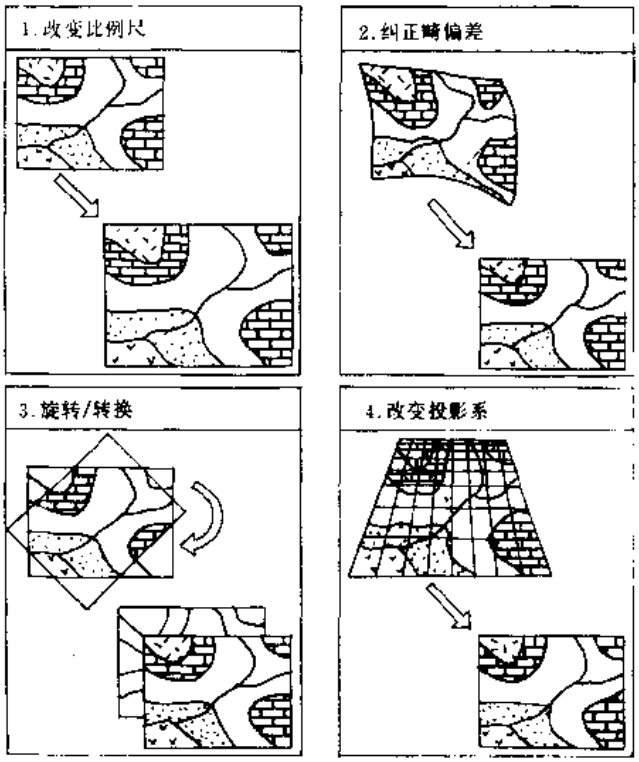


图 1—11　原始数据处理与图幅管理

多边形重叠和分离（消除）是将同一地区、同一比例尺的几种多边形要素的数据文件进行重叠处理，从而产生许多新的多边形。例如，地质地形图就是同一地区、相同比例尺的地形地物数据文件与地质资料的数据文件叠加的结果。地学信息的重叠和分离是经常需要做的操作，如图 1—13 所示。

6. 数字地形分析

众所周知，等高（值）线法很适合于表示不规则的连续变化的表面，但它不适用于进行数字地形分析。数字地形模型（DTM）或数字高程模型（DEM）是以数字方式表示空间起伏变化的连续表面。许多 GIS 都提供了构建 DTM 的软件包及进行地形分析的功能。数字地形分析的主要内容有：等高线的生成与分析，地形要素的计算与分析，断面图分析，三维立体显示和计算等。

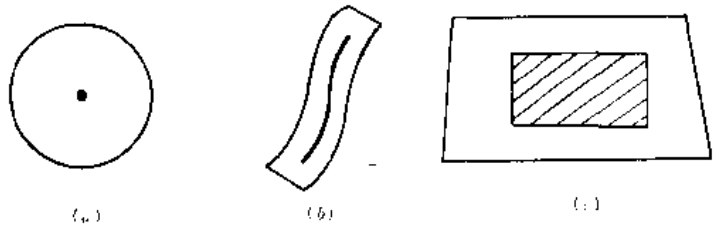


图 1—12　缓冲区的生成

a　点缓冲区；*b*　线缓冲区；*c*　面缓冲区

7. 格网分析与量测计算

例如，格网叠加与计算，半径查找，距离计算，路径优选，多边形面积量算及容积计算等。

8. 自动制图

自动制图是 GIS 的主要功能之一。在 GIS 数据库中存储有各种空间定位数据和属性数据，如各种地形、地理、资源和环境的数据。对这些数据可及时更新、增删。制图功能模块或软件包通过图形编辑，可根据用户的需要对数字地图进行整饰，按照给定的符号、注记和颜色进行图形显示或绘图仪输出。

（三）数据输出

经 GIS 系统数据处理的结果能够以多种形式表示，有的形式是用户可以阅读或理解的，有的则需要转换到计算机系统的其它设备中。前者如各种图表、像片；后者主要是 CCT

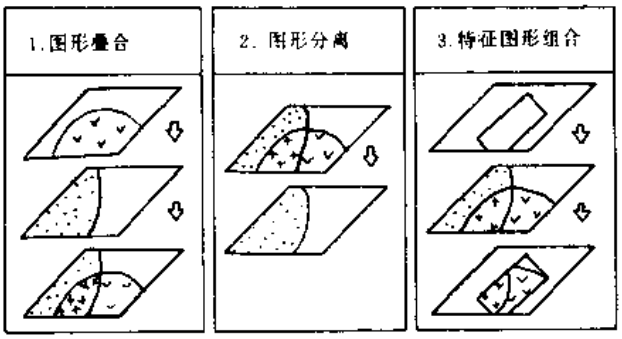


图 1—13　图形(像)的重叠和分离

磁带。图 1—14 表示目前常用的各种输出设备和输出形式。

二、GIS 技术的主要应用领域

科学技术的总体发展趋势是既不断细分，又不断综合、交叉，以解决各种复杂、困难的宏观和微观问题。多年来，许多独立发展的学科都试图找出一种类同的程式——为多种目标或任务发展一套强有力的工具，来收集、存储、变换及显示从自然界和人类社会获取的数据，这一套工具就是地理信息系统，有时又称资源与环境信息系统，或者土地信息系统等。此外，由于 GIS 具有许多功能和优点，因此它的应用领域相当广泛，并且还在向深度和广度发展。现介绍几个主要领域。

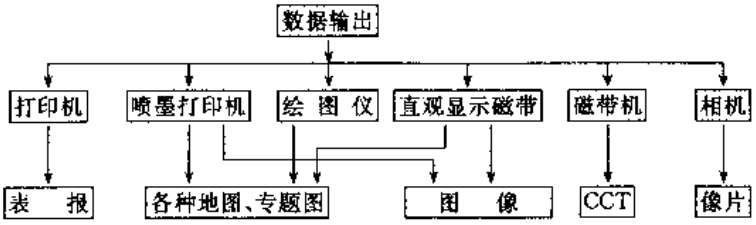


图 1—14　GIS 的主要输出设备与输出形式

1. 资源清查与管理

资源的清查、管理和分析是 GIS 应用中趋于成熟而重要的领域，包括土地资源、森林资源和矿产资源的清查、管理，土地利用规划，野生动植物的保护等。GIS 系统的主要任务是将各种来源的数据和信息有机地汇集在一起，并通过系统的统计、叠置分析等功能，按照多种边界和属性条件，提供区域多种条件组合形式的资源统计和资源状况分析，为资源的合理开发、利用和科学管理提供依据。

2. 区域和城镇规划

区域或城镇规划具有高度的综合性，涉及大量的组成要素，如资源、环境、人口、交通、经济、教育、文化、通讯和金融等。GIS 的数据库管理系统有利于对这些复杂的因素进行统一的处理和分析。GIS 的空间搜索算法、多元信息的叠置处理、空间分析方法和网络分析功能，能够完成道路交通的规划、公共设施的配置、城市建设用地适宜性评价、商业布局、区位分析和地址选择等。我国大、中城市很多，很必要加强城市规划建设、管理决策的科学化、现代化。对此，GIS 是强有力的工具。

3. 灾害监测

利用 GIS 方法和多时相遥感数据，可以有效地用于森林火灾的预测预报、洪水监测和灾情损失的估算，为防洪决策和救灾抢险及时提供准确的信息。例如黄河三角洲地区的防洪减灾研究表明，在 GIS 的支持下，通过建立大比例尺数字高程模型以及获取的其它有关信息，包括土地利用、水系、居民点、油井、工厂、工程设施及有关社会经济统计数据等，利用 GIS 的叠置操作和空间分析功能，可计算出若干个泄洪区域内的土地利用及其面积，比较不同泄洪区内的房屋、财产损失等，进而确定最佳的泄洪区域，制定出泄洪区内人员撤退、财产

转移、救灾物资供应的最佳路线。

4. 环境保护和评价

GIS 技术也是进行环境评价，环境规划管理的有力工具。其内容包括：环境监测和数据收集，建立基础数据库和环境动态数据库，建立环境污染的有关模型，提供环境管理的统计数据和报表输出，环境作用分析和环境质量评价，环境信息传输和制图等。

5. 土地管理和房地产管理

这里，涉及的数据包括土地的位置、房地界、名称、面积、类型、等级、权属、质量、地价、税收、地理要素及有关设施等内容。借助 GIS 进行地籍数据的管理，开展房地产质量评价和经济评价，输出地籍图，便于为用户提供有关信息，进行土地的科学管理和合理利用。

6. 宏观决策

GIS 利用拥有的数据库，通过一系列决策模型的构建和比较分析，可为国家或区域的宏观决策提供科学依据。例如，在我国三峡工程的决策研究中，利用 GIS 的多种功能建立了环境监测系统，为这项宏大工程提供了建库前后环境变化的数量、速度和演变趋势等可靠数据，作了三峡工程的宏观决策，以避免发生重大失误和损失。又如，通过对矿区地理信息系统（MGIS）数据库中的地质、采矿、地形等数据的分析研究，利用图形叠置等功能和开采沉陷规律模型，可以进行沉陷区的预计，为保护地面建筑物和开采设计提供科技手段，可取得良好的效益。

总之，建立在系统论、信息论和控制论等现代科学理论，以及计算机技术、空间技术等基础上的 GIS，通过充分发挥其在理论、技术与应用三结合的优势，已跻身于世界高新技术领域；并且随着 GIS 应用领域的扩大和深入，又必然要推动 GIS 理论和技术方法的不断发展和完善，进而形成独立的学科——地理信息学（Geo-informatics 或 Geomatics），以及新兴的产业——地理信息产业。目前，主要的问题不是要不要 GIS，而是如何应用它、发展它，使它发挥最佳的效益，为人类造福。

第 1—5 节　GIS 的相关学科与技术

GIS 是现代科学技术发展和社会需求的产物。人口、资源、环境、灾害这是影响人类生存与发展的四大基本问题。为了解决这些问题必须要自然科学、工程技术、社会科学等多学科、多手段联合攻关。于是，许多不同的学科，包括地理学、测量学、地图制图学、摄影测量与遥感学、计算机科学、数学、统计学以及一切与处理和分析空间数据有关的学科，都在寻找一种能采集、存储、检索、变换、处理和显示输出从自然界和人类社会获取的各式各样数据、信息的强有力工具，其归宿就是地理信息系统，或称空间信息系统，资源与环境信息系统。因此，GIS 明显地具有多学科交叉的特征，它既要吸取诸多相关学科的精华和营养，并逐步形成独立的边缘学科，又将被多个相关学科所运用，并推动它们的发展。

地理学和测绘学是以地域单元研究人类居住的地球及其部分区域，研究人类环境的结构、功能、演化以及人地关系。空间分析是 GIS 的核心，地理学作为 GIS 的分析理论基础，可为 GIS 提供引导空间分析的方法和观点。测绘学和遥感技术不但为 GIS 提供快速、可靠、多时相和廉价的多种信息源，而且它们中的许多理论和算法可直接用于空间数据的变换、处理。

遥感是一门 60 年代以后发展起来的新兴学科。由于遥感信息所具有的多源性，弥补了

常规野外测量所获取数据的不足和缺陷，以及在遥感图像处理技术上的巨大成就，使人们能够从宏观到微观的范围内，快速而有效地获取和利用多时相、多波段的地球资源与环境的影像信息，进而为改造自然，造福人类服务。GPS 卫星全球定位系统是新一代卫星导航和定位系统。美国已于 1993 年完成了整个系统的部署，达到全效能服务的阶段。它在测量和勘察领域可以取代常规大地测量来完成各种等级的定位工作，在航空摄影和遥感领域 GPS 遥感对地定位系统很有发展前途，在舰船、飞机、汽车的导航定位，导弹的精确制导方面应用更为广泛，在地球动力学、重力场、磁场等的研究中也能发挥很大作用。

此外，GIS 最初是从机助制图起步的，早期的 GIS 往往受到地图制图中在内容表达、处理和应用方面的习惯影响。但是建立在计算机技术和空间信息技术基础上的 GIS 数据库和空间分析方法，并不受传统地图纸平面的限制。GIS 不应当只是存取和绘制地图的工具，而应当是存取和处理空间实体的有效工具和手段，存取和绘制地图只是其功能之一。

再者，GIS 与计算机科技、数学、运筹学、统计学、认知学等学科也密切相关。计算机辅助设计（CAD）为 GIS 提供了数据输入和图形显示的基础软件；数据库管理系统（DBMS）更是 GIS 的核心；数学的许多分支，尤其是几何学、图论、拓扑学、统计学、决策优化方法等被广泛应用于 GIS 空间数据的分析。

总之，遥感技术可以为资源开发和环境监测提供丰富、实时的宏观信息，并为计算机制图系统和 GIS 的数据更新提供可靠、快速的数据源。但遥感对浩如烟海的社会经济统计数据，人类活动的大量信息却无力获取；计算机地图制图技术可为地理信息的时空分布和产品输出提供先进的手段，但它本身无区域综合、分析和决策的功能；GPS 技术、数字摄影测量和遥感技术可成为 GIS 数据采集和及时更新的主要技术手段和有力支撑；而 GIS 既能提供信息查询、检索服务，又能提供综合分析评价，它在资源和技术方面的博才与运筹帷幄的优势，是遥感、GPS 和自动制图技术所不及的。因此，只有它们的有机结合，才能使遥感和 GPS 技术所获取的瞬时信息经过积累和延伸，才具有反映自然历史发展过程和人为影响的能力，并达到实时处理的功能，为科学管理、规划决策服务。这样，逐步形成了 GIS 与诸多学科之间互相有联系，也有挑战，彼此推动、共同发展的关系和局面。图 1—15 示意表示这样的关系。

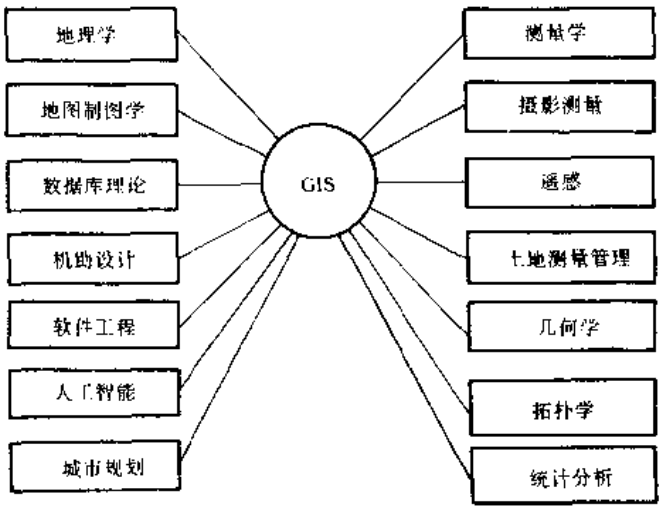


图 1—15　GIS 的相关学科技术

第二章　GIS 的数据结构

第 2—1 节　概　　述

一、GIS 中的空间信息及其特征

地理区域系统是由人文和自然要素组成的复杂的动态区域系统。它包括资源、环境、社会和经济信息四个子系统，这四个子系统在内容上、表达方式上、时空尺度上都有着各自不同的特点。随着地理信息系统的发展和日益被广泛应用，加上区域持续发展的要求，对资源、环境、社会和经济综合信息及其数理应用模型的研究发展迅速，因为地理信息系统存储、处理的信息是综合信息、具有空间分析、决策的能力。

由于具体研究区域的大小、规模不同，因而对地理空间对象的空间分辨率的要求也不同，也即是变化的，不确定的。例如，在一张小比例尺的全国地图中，一个城市一般用一个点或点状符号来表示，而在一张大比例尺的城市地图中，则要详细反映街区和建筑物。于是，在地理信息系统中以“空间实体”作为描述、反映空间对象的单体。所谓“实体”(entity) 是指自然界、自然现象和社会经济事件中不能再分割的单元。例如，在小比例尺地图中北京市是一个实体；在大比例尺图中一座房子可以是一个实体，如果它不能（或不必要）再被细分成多座房子的话。但在必要时，一座房子又可分为许多房间。因此，“实体”是一个概括性的、复杂的、具有相对意义的概念或术语。

我们所生活的地球“空间”包含有许许多多复杂的实体。如果它们是可以定义的、可以定界的、可以加以说明的话，那么对于某一特定实体现象就可以有如表 2—1 所示的五种特征：标识符 (identifier)，如名称、编码；空间位置；属性特征；作用、性能或功能；空间性质。

表 2—1　实　体　特　征

实体	标识符	位　置	空间特征	特征属性	功能属性
学　校	名　字	坐　标	与另一实体的间距	学生数	教　育
铁　路	编码号	某车站坐标	连接两个城市	运输量	运营中
钻　孔	标识号	空间坐标	深　度	孔　径	目　的
蓄水池	名　字	地面坐标	容　积	水　质	水位变化

可见，实体的特性可分为空间特性，也称几何特性和非空间特性两类。非空间特征也称为属性，其特征数依实体的不同而多少不等，但一般可大致归结为名称和属性特征两类。

实体的空间特征可以从以下三个方面来考察：空间维数；空间特征类型；实体之间的空间关系和结合。实体的空间特征可以有四种不同类型的表示，即点、线、面和体。它们的整数维数分别是 0 维、1 维、2 维和 3 维。

1）点或节点 (node)　是对点状的地形、地物要素的几何描述。有时对点实体的描述可附有高程 (Z)，甚至还需要一个方向。

2）线或折线　是对线状空间实体的部分或全部进行的几何描述，它由两个或两个以上的一系列按顺序相连结的坐标对组成。就其形式有曲线和折线两种。折线点可以附有高程，线也可以有方向，如河流的流水方向。

3）面或多边形　是由一条线或一系列线界定的几何表示，是对面状空间实体的几何描述。面可能是简单的，也可能是复杂的，可以是由外轮廓线和内轮廓线组成的，也可以是由两个或两个以上的面相邻或叠加而成的，例如，中间有岛的湖泊等。

4）体　是对空间三维实体或多面体的几何表示。其形状有简单的、复杂的和带空洞的实体。实体空间对地质、采矿、海洋、建筑、土工、气象等领域的研究是很重要的几何表示方式。

对于点、线和面实体几何表示的规定值，由于自然界和人类社会存在着各式各样的确定性和不确定性或模糊性，因而，对点来说有“确定的”和“位置近似的”这两种情况。对于线和面来说有“确定的”、“不确定的”、“位置近似的”和“虚拟的”或者“导出的”等情况。例如，国境线有确定性的和不确定性的，湖泊的岸线因季节不同而异，矿体的边界、储量等。

空间实体在许多情况下是由点、线、面和体中的两种或两种以上的实体构成的混合实体。

分形几何（Fractal geometry）能够比欧氏几何更充分更准确地表示现实世界的实体或现象。分形几何在科学研究和生产实践中有着重要意义和作用。分数维数（Fractal dimensionality）是分形几何的一个重要的定量指标和概念。曲线的维数随着其弯曲程度的不同可以是大于1、小于2的分数，曲面的维数随着其起伏、变化程度的不同可以是大于2、小于3的分数。

要完整地描述空间实体，一般需要同时有空间（几何）特征数据和属性数据。在信息系统或者任何数据集中，特别是地理信息系统数据库的数据集中，对实体的描述除上述空间和非空间特性之外，还往往附带有一些用来定义或判别其准则或过程的资料，这些资料被称为数据（目录）文件，或数据元文件（Metainformation，或Metadata）。它们是关于数据的数据，可以包括：对实体空间和属性的定义，对属性或编码的解释，对基于遥感图像信息的假彩色编码的解释，对空间实体定界规则的描述，有关数据源、数据质量及其产生日期的说明，对漏失值或度量不适宜的解释，对数据存取方式、组织形式的说明等。

元文件数据主要有三种用途：一是作为数据的目录，提供数据集内容的摘要，类似于图书馆中的一种图书卡片；二是有助于数据共享，提供数据集或数据库转换和使用所需要的数据内容、形式、质量方面的信息；三是内部文件记录，用以记录数据集或数据库的内容、组织形式、维护和更新等情况。

目前国际上还没有元数据的标准。据悉美国联邦地理数据委员会（FGDC）以及国家标准和技术研究所（NIST）正着手元数据标准的制定，以作为它的空间数据转换标准（SDTS）的补充。加拿大遥感中心（CCRS）也在着手这方面的工作。

基于上述论述和分析不难理解，在GIS中的实体数据具有以下特点：

（1）以统一的坐标系统（全球的、全国的或局部的）进行空间定位。

（2）空间数据与非空间数据是相互结合的整体。空间几何关系、空间特性不准确、不可靠不行，要有一定的分辨率，但没有非空间属性数据，或者不可靠、不准确，也会影响

实际应用，甚至引出谬误。

（3）图形数据与图像数据，或者说矢量数据方式与栅格数据方式是并存或叠合的。

（4）实体数据间存在着复杂的关系。

（5）实体数据具有时效性，即周期性和时间性，因此GIS的4维时空数据结构和数据模型是实体数据的重要内容。过时的信息不具备现势的意义，但不一定就没有用，可以作为历史过程保存。

二、GIS中的数据类型

地理信息中的数据来源和数据类型很多，概括起来主要有以下四种数据类型：

1）几何图形数据　来源于各种类型的地图和实测几何数据。几何图形数据不仅反映空间实体的地理位置，还要反映实体间的空间关系。

2）影象数据　主要来源于卫星遥感和航空遥感等。

3）属性数据　来源于实测数据，文字报告，或地图中的各类符号说明，以及从遥感数据中通过解释得到的信息等。

4）地形数据　来源于地形等高线图的数字化，已建立的格网状的数字化高程模型（DTM），或其他形式表示的地形表面（如TIN）等。

在具有智能化的GIS中还应有规则和知识数据。

空间数据的特征可以概括为空间特征、属性特征和时间特征。空间特征表示地理实体或现象的空间位置和相互关系；属性特征表示其名称、分类、数量等；时间特征指实体或现象随时间的变化。实体空间位置及属性常随时间相互独立的变化。因此，GIS是一个动态的复杂的空间数据管理系统。

三、空间数据的数据结构

数据结构(Data Structure)即数据组织的形式，是适合于计算机存贮、管理和处理的数据逻辑结构。对空间数据则是地理实体的空间排列方式和相互关系的抽象描述。数据如果不按一定的规律贮存在计算机中，不仅用户无法理解，计算机也不能正确处理，将成为一堆无用的垃圾。目前尚无一种统一的数据结构能同时存贮上述四种类型的数据，而是将它们分别以矢量数据结构，栅格数据结构、二维关系表、及其他类型的数据结构方式存贮。有些GIS软件包含矢量和栅格两种数据；并能实现它们之间的转换。目前正在研制矢量栅格一体化的数据结构。有的GIS是将空间几何数据和属性数据分别管理的，也有采用统一的数据管理模型。本章主要介绍空间几何数据的数据结构，有关数据模型问题将在下一章讨论。

所谓空间数据编码是空间数据实现中，将空间实体按一定的数据结构转换为适合于计算机存贮和处理的数据过程。

四、空间数据的抽象表示

地理数据可抽象为点、线、面三类元素，以便表示它们的位置、大小、形状、高低等。

点（Point），又称为元素（Element）或像元（Pixel）。它以一对坐标（x，y）表示其位置，并至少有一个属性，逻辑上不能再分。这里所指的点是抽象的点，它可以具体指某一个独立点，例如某个油井或某个钻孔；在小比例尺图或影像中可以表示某个村落，某一城市。对于线的起点，终点或交点，则称为结点（Node），它同样以一对坐标表示。

线（Line），是具有相同属性点的轨迹，用一个坐标序列表示。道路、河流、地形线、区域边界等均属于线状地物。线上各点具有相同的公共属性并至少存在一个属性。线有时也

称为弧段（Arc），特别是指两个结点之间的线段。

面（Area），是线包围的面域，或称为多边形（Polygon），也是具有相同属性点的轨迹，并以一坐标系列表示。

点、线、面都是一个空间实体，实体是基本的地理信息单元。空间的点、线、面可以按一定的地理意义组成区域（Region），有时称为一个覆盖（Overlay），或数据平面（Data Plane）。各种专题图在 GIS 中都可以表示为一个数据平面。

五、空间数据的拓扑关系

在 GIS 中，为了真实地反映地理实体，不仅要包括实体的位置、形状、大小和属性、还必须反映实体之间的相互关系。这些关系就是指它们之间的邻接关系，关联关系和包含关系。

地图上的拓扑关系是指图形保持连续状态下变形，但图形关系不变的性质。地图上各种图形的形状、大小会随图形的变形而改变，但是上述三种关系则不会改变。俗称拓扑关系是绘在橡皮上的图形关系，或者说拓扑空间中不考虑距离函数。

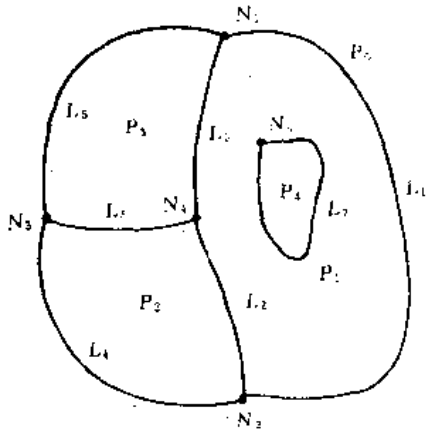


图 2—1　空间数据的拓扑关系

如图 2—1 所示：N_1，N_2……为结点；L_1，L_2……为线段（弧段）；P_1，P_2……为面（多边形）。

邻接关系：空间图形中同类元素之间的拓扑关系。例如多边形之间的邻接关系，P_1 与 P_2，P_1 与 P_3；又如结点之间的邻接关系 N_1 与 N_2，N_1 与 N_4 等。

关联关系：空间图形中不同元素之间的拓扑关系。例如结点与弧段的关联关系 N_1 与 L_1、L_3、L_5、N_2 与 L_1、L_2、L_4；多边形与弧段的关联关系 P_1 与 L_1、L_2、L_3，P_2 与 L_2、L_4、L_6、P_4 与 L_7 等。

包含关系：空间图形中同类但不同级元素之间的拓扑关系。例如多边形 P_1 中包含 P_4。

如果要将结点、弧段、面块相互之间所有的拓扑关系表达出来，可以组成四个关系表，即表 2—2，表 2—3，表 2—4 和表 2—5。

上述关系也可以只用其中的部分表来表示，而其余关系则隐含其中，需要时再建立临时关系表。例如表 2—4 对于网络分析非常重要，而对于主要以面状目标的管理系统来说则可以省略。

表 2—2 中，弧段前的负号表示面域中含有岛。表 2—3 中每一弧段的左、右结点分别

表 2—2　面域与弧段的拓扑关系

面域	弧　段
P_1	L_1，L_2，L_3，$-L_7$
P_2	L_2，L_4，L_6
P_3	L_3，L_6，L_5
P_4	L_7

表 2—3　弧段与结点的拓扑关系

弧　段	结　点
L_1	N_1，N_2
L_2	N_2，N_4
L_3	N_4，N_1
L_4	N_2，N_3
L_5	N_3，N_1
L_6	N_3，N_4
L_7	N_5，N_5

表 2—4　结点与弧段的拓扑关系

结　点	弧　段
N_1	L_1, L_3, L_5
N_2	L_1, L_4, L_2
N_3	L_4, L_5, L_6
N_4	L_2, L_6, L_3
N_5	L_7

表 2—5　弧段与面域的拓扑关系

弧　段	左邻面	右邻面
L_1	P_0	P_1
L_2	P_2	P_1
L_3	P_2	P_1
L_4	P_0	P_2
L_5	P_0	P_3
L_6	P_3	P_2
L_7	P_1	

称为始结点和终结点。

空间数据的拓扑关系，对数据处理和空间分析具有重要的意义，因为：

(1) 拓扑关系能清楚地反映实体之间的逻辑结构关系，它比几何关系有更大的稳定性，不随地图投影而变化。

(2) 利用拓扑关系有利于空间要素的查询，例如某条铁路通过哪些地区，某县与哪些县邻接。又如分析某河流能为哪些地区的居民提供水源，某湖泊周围的土地类型及对生物栖息环境作出评价等。

(3) 可以根据拓扑关系重建地理实体。例如根据弧段构建多边形，实现道路的选取，进行最佳路径的选择等。

第 2—2 节　矢量数据结构

一、矢量数据结构的图形表示

矢量数据结构是最常见的图形数据结构。它通过记录坐标的方式精确地表示点、线、面实体的位置。这里 x、y 可以对应于地面点的经纬度，也可以对应于平面直角坐标系的 x 与 y 坐标。它们通常是通过跟踪式数字化仪数字化求得的。图 2—2 为一个例子，图 2—3 为其矢量结构表示。

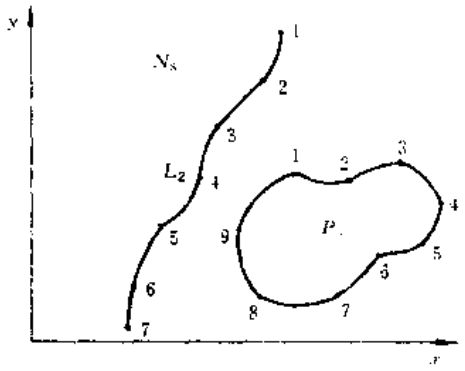


图 2—2　矢量数据结构

	特征码	位置坐标
点	8	x, y
线	2	x_1, y_1; x_2, y_2 …… x_7, y_7
面	5	x_1, y_1; x_2, y_2 …… x_9, y_9; x_1, y_1

图 2—3　矢量结构表示

点实体：只需记录特定坐标系下的坐标和属性代码；

线实体：用两个点或一系列点的坐标表示，并加上属性代码；

面实体：用边界线上一系列点的坐标表示，由于多边形封闭，首尾点坐标必须相同。

线和多边形边界通常是由光滑曲线间隔采样而来，如取点越多，以后恢复时越接近原曲线，失真越小，但数据量也越大。反之取点过少，恢复时失真就大，甚致成为折线。因此数字化时采样取点应适当。这些坐标一般都以文件的方式贮存在计算机中。

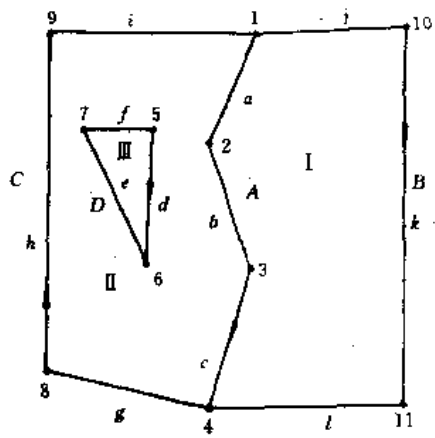


图 2—4　矢量结构图形

二、矢量结构的几种编码方式

在矢量结构中，点实体和线实体的编码比较简单，只要能将空间信息和属性信息记录完全就可以了。而多边形实体的编码要复杂得多，不但要表示其位置和属性，还要能表示其拓扑关系。因此矢量结构有时称为多边形结构。

下面介绍几种多边形矢量编码的方法。

（一）实体式

它以多边形为单位，由多边形边界坐标对集合及说明信息组成。例如图 2—4 所示的多边形，可以组成如图 2—5 的文件。

多边形	数　据　项
Ⅰ	x_1, y_1; x_2, y_2; x_3, y_3; x_4, y_4; x_{11}, y_{11}; x_{10}, y_{10}; x_1, y_1
Ⅱ	x_1, y_1; x_2, y_2; x_3, y_3; x_4, y_4; x_8, y_8; x_9, y_9; x_1, y_1
Ⅲ	x_5, y_5; x_6, y_6; x_7, y_7; x_8, y_5

图 2—5　实体式编码文件

这种文件的结构简单，容易实现以多边形为单位的运算和显示。但这种方法的缺点也很明显，它们是：

（1）多边形之间的公共边界被数字化两次和存贮两次，造成数据冗余和不一致。

（2）缺少多边形的邻域信息和图形的拓扑关系。

（3）岛只作为一个单个图形，没有与外界多边形的联系。因此这种方法只用在简单的制图系统中。

（二）索引式

这种方法对所有边界点的坐标按顺序建立坐标文件，再建立点与边界线，线与多边形的索引文件，如图 2—6 所示。

文件的结构如表 2—6、2—7、2—8 和 2—9 所示。

这种索引编码法，虽然多了两个文件，但是消除了多边形边界的数据冗余和不一致，邻域信息也可以通过多边形文件的线索引得到。但是这种方法在处理岛状信息及检索拓扑关系时会带来困难。

（三）双重独立式

它是有名的美国人口统计系统采用的一种编码方式，简称 DIME（Dual Independent Map Encoding）编码系统。它以城市街道为编码的主体。其特点是采用了拓扑编码结构。这种结构最适合于城市信息系统。其中街道与河流为线状要素，弯曲线段则由一系列直线段表示。仍以图 2—4 为例，可组成如表 2—9 所示的线文件。

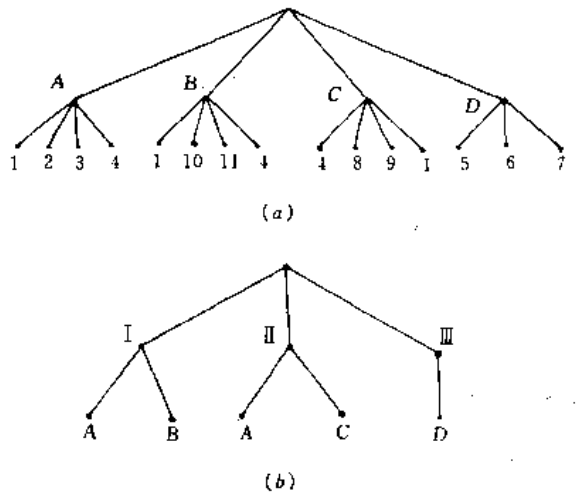


图 2—6　线与点，面与线的索引

表 2—6　点　文　件

点号	坐　标
1	x_1, y_1
2	x_2, y_2
3	x_3, y_3
4	x_4, y_4
5	x_5, y_5
6	x_6, y_6
7	x_7, y_7
8	x_8, y_8
9	x_9, y_9
10	x_{10}, y_{10}
11	x_{11}, y_{11}

表 2—7　弧　段　文　件

线号	起点	终点	点　号
A	1	4	1, 2, 3, 4
B	1	4	1, 10, 11, 4
C	4	1	4, 8, 9, 1
D	5	5	5, 6, 7, 5

表 2—8　多边形文件

多边形号	边界线号
I	A, B
II	A, C
III	D

表 2—9　线　文　件

线号	左多边形	右多边形	起点	终点	线号	左多边形	右多边形	起点	终点
a	I	II	1	2	g	0	II	4	8
b	I	III	2	3	h	0	III	8	9
c	I	III	3	4	i	0	III	9	1
d	III	III	5	6	j	0	III	1	10
e	I	III	6	7	k	0	III	10	11
f	III	III	7	5	l	0	III	11	4

这种结构的邻接关系和关联关系明确。通过线文件能自动寻找并形成面文件。此外，这种结构还需要有点文件和面文件，这里不予列出。

（四）链状双重独立式

这种方式是对上述 DIME 方法的一种改进。在 DIME 中，一条边只能用直线两端点及相邻的面域来表示。而这种结构可将若干线段合为一个弧段（或链段），每个弧段可以有许多中间点，其端点则为弧段的交点或起终点。这种方式的弧段坐标文件是在数字化过程中由一系列点的位置坐标组成。弧段（或链段）文件由弧记录组成，存贮弧段的起止结点号和弧段左右多边形号。结点文件由结点记录组成，存贮每个结点的结点号、结点坐标及与该结点连接的弧段等。结点文件一般是通过软件自动生成的，因为在数字化弧段的过程中，由于数字化操作的误差，各弧段在同一结点处的坐标不可能完全一致，需要进行匹配处理。

当其偏差在允许范围内时，可取同名结点的坐标平均值。如果偏差过大，则弧段需要重新数字化。

多边形文件由多边形记录组成，数据项包括多边形号，多边形组成的有关弧段号，以及周长、面积、中心点坐标及有关岛状区域信息等。多边形文件也可以通过软件自动检索各有关弧段生成，并同时计算出多边形的周长和面积以及中心点的坐标。多边形中包含岛状区则面积为负，在总面积中减去。其组成的弧段号前也冠以负号。

图 2—4 所示图形的弧段坐标文件、弧段文件及多边形文件分别见表 2—10、2—11 和 2—12。

表 2—10 弧段坐标文件

弧段号	坐　　标	弧段号	坐　　标
A	x_1, y_1; x_2, y_2; x_3, y_3; x_4, y_4	C	x_4, y_4; x_8, y_8; x_9, y_9; x_1, y_1
B	x_1, y_1; x_{10}, y_{10}; x_{11}, y_{11}; x_4, y_4	D	x_5, y_5; x_6, y_6; x_7, y_7; x_5, y_5

表 2—11 弧段文件

弧段号	起结点	终结点	左多边形	右多边形
A	1	4	I	I
B	1	4	0	I
C	4	1	0	I
D	5	5	I	I

表 2—12 多边形文件

多边形号	弧段号	周长	面积	中心点坐标
I	A, B			
I	A, C, −D			
I	D			

这种矢量结构的编码方式被一些商品化的 GIS 软件所采用，例如美国的 ARC/INFO GIS 软件。利用这种软件对线划图形数字化时，只需沿每条弧段用跟踪数字化仪数字化，其软件能自动生成各有关的文件，因而十分方便。

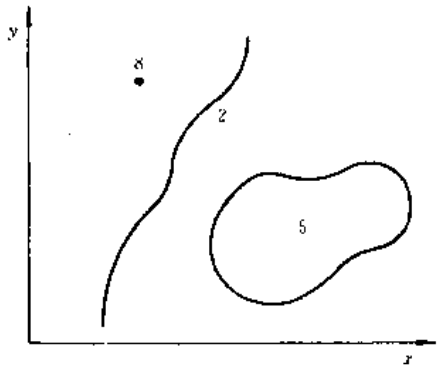
第 2—3 节　栅 格 数 据 结 构

一、栅格数据结构的图形表示

栅格结构或称为网格结构（Raster 或 Grid cell），又称为像元结构（Pixel）。它是将地表表面划分为均匀紧密相邻的网格阵列。如图 2—7*a* 所示的点、线、面图形，在栅格结构中就成为图 *b* 表示的网格阵列。每个网格（或像元）的位置由行列号定义。它包含一个代码，以表示该网格的属性（如灰度）或指向属性记录的指针。图 2—7*b* 阵列中的 8 是属性代码为 8 的一个点；2 为属性代码为 2 的一条线，5 为属性代码为 5 的面域。

遥感影像属于典型的栅格结构，其特点是属性明显，而位置隐含。阵列中的数字直接记录属性或属性指针，而所在的位置需根据行列号才能转换为相应的坐标。每个存贮单元的行列位置要根据它在文件中记录的位置得到。

栅格结构表示的地理实体是不连续的、经量化的近似数据。例如在栅格结构中表示的地表面，地表被分成相互邻接、规则排列的矩形方块，每个地块与一个栅格单元相对应。栅格数据的比例尺就是栅格大小与地表相应单元大小之比。在栅格数据处理时，一般认为，表示的量化表面是连续的，以便能使用某些函数计算。在计算面积、长度、距离等空间指标



(a)

0	0	0	2	0	0	0	0
0	8	0	2	0	0	0	0
0	0	2	0	0	0	0	0
0	0	2	0	0	0	5	5
0	2	0	5	5	5	5	5
0	2	0	5	5	5	5	5
2	0	0	0	5	5	0	0
2	0	0	0	0	0	0	0

(b)

图 2—7　图形的栅格数据结构表示

时，如果栅格较大，会造成较大的误差。因为在一个栅格的地表范围内，可能多于一种地物，而表示在栅格单元中只能有一个代码。这类似于遥感影像的像元，其记录的光谱数据只能是该像元代表的区域地物类型光谱的综合效果。

二、栅格结构的建立

（一）建立栅格结构的途径

（1）在专题图上均匀划分网格，逐个网格决定其代码。通常可按网格的尺寸准备一张聚脂薄膜的透明格网，将此网叠置于原图上，从图上读取每个格网的数据。

（2）用扫描数字化仪，按行和列逐点扫描专题地图，将扫描数据重采样和再编码得到网格数据。

（3）从矢量数据转换得到。

（4）将经过分类解释的遥感影像数据直接输入或重采样后输入系统中。

（二）栅格代码的确定

在决定栅格代码时应尽量保持其真实性。例如图 2—8 所示为某栅格范围的一块矩形的地形区域，内部含有 A，B，C 三种地物类型，O 为中心点。在决定其栅格代码时，可根据需要采取如下的方案。

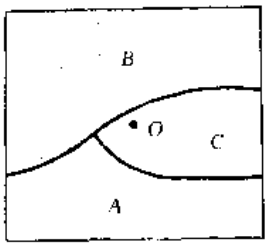


图 2—8　栅格代码的确定

1）中心点法　用位于栅格中心处的地物类型决定其代码。由于中心点位于代码为 C 的地物范围内，故其代码为 C。这种方法常用于有连续分布特性的地理现象。

2）面积占优法　以占矩形区域面积最大的地物类型决定栅格单元的代码。从图看出，B 类地物所占面积最大，故相应栅格单元代码为 B。

3）重要性法　根据栅格内不同地物的重要性，选取最重要的地物类型决定相应的栅格单元代码。设图中 A 类地物为最重要的地物类型，则栅格代码为 A。这种方法常用于有特殊意义而面积较小的地理要素，特别是点状和线状地理要素。如城镇、交通线、水系等。在栅格代码中应尽量表示这些重要地物。

（三）栅格系统的原点和尺寸

栅格系统的起始坐标应尽量和国家基本比例尺地形图公里网的交点相一致，并分别采

用公里网的纵横坐标轴作为栅格系统的坐标轴。这样也为坐标的隐含编码提供方便。

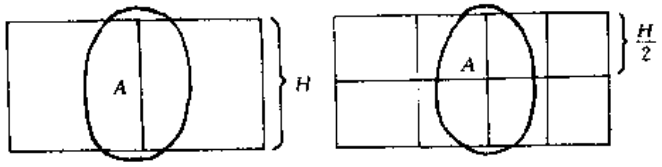


图 2—9　栅格尺寸的选择

栅格单元的合理尺寸应能有效地逼近空间对象的分布特征，以保证空间数据的精度。但是用栅格来逼近空间实体，不论采用多细小栅格，与原实体比较都会有误差。通常以保证最小图斑不丢失为原则来确定合理的栅格尺寸。如图 2—9 所示，设研究区域的最小图斑面积为 A，当栅格边长为 H 时，该图斑可能得不到反映而丢失。而当边长为 $H/2$ 时，该图斑就能得到反映，故其合理尺寸可以定为。

$$H=\frac{1}{2}\ (\min\ \{A_i\})^{\frac{1}{2}}$$

式中 i 为多边形号，i＝1，2……n。

对图 2—8 所示的区域，如果将它分为四个小区域，代码将分别为 B、B、A、C，则显示的信息将更丰富些。但是，增加栅格单元将使数据量大大增加。为解决这一矛盾和难题，发展了一系列压缩数据量的编码方法。对此将在后面介绍。

（四）栅格数据的组织

GIS 中空间实体的类型很多，不论是在栅格数据结构，还是在矢量数据结构中，都有如何组织这些不同类型实体的问题。通常的办法是将不同类型的地理实体分层编排，每层只具有单一的类型。例如点状实体、线状实体和面状实体分别处在不同的层上。同样类型的实体又可按不同的专题内容分层。例如同样是线状要素，可将道路作为一个层，水系又作为另一个层。这样，在一个层中可以只有一种变量反映其不同的属性，也就是说一个栅格单元只有一个属性代码。当然，也可以将多个变量存贮在同一个栅格单元中，其文件格式也将有所不同，但这样会给栅格结构的编码增加困难。

三、栅格结构编码的方法

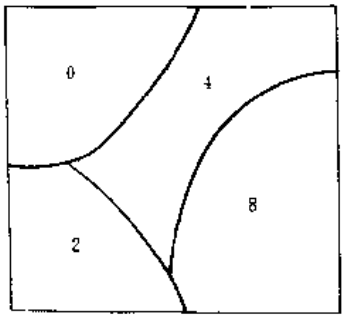
对栅格数据的编码方法，考虑到栅格数据的数据量非常大，冗余数据很多，必须采取一些压缩数据的编码方法。压缩编码有信息保持编码和信息不保持编码两种。信息保持编码指编码过程中没有信息损失，通过解码操作可以恢复原来的信息；信息不保持压缩编码是指为了最大限度地压缩数据，在编码过程中要损失一部分认为不太重要的信息，解码时这部分信息难以恢复。在 GIS 中多采用信息保持编码。而对于原始遥感影像进行压缩编码时，有时也采用信息不保持的压缩编码方法。

1. 直接栅格编码

直接栅格编码就是将栅格数据看作一个数据矩阵，逐行逐个记录代码数据。可以每行都从左到右，也可从奇数行从左到右而偶数行从右到左。或者为了特殊目的而采用特殊的方法。

如图 2—10a 所示的地块图，0、2、4、8 分别表示四种不同属性的面域代码，其栅格阵列见图 b，而直接栅格编码的数据记录见图 c。

上述数据可以采用顺序或随机文件记录，记录长度＝单字长×列数。有些系统为了显示和处理方便，也可将一块（若干行）作为一个记录。



(a)

0	0	0	0	0	4	4	4
0	0	0	4	4	4	4	4
0	0	4	4	4	4	8	8
0	0	4	4	4	8	8	8
2	2	4	4	8	8	8	8
2	2	2	4	8	8	8	8
2	2	2	2	8	8	8	8
2	2	2	2	8	8	8	8

(b)

记录 1	0	0	0	0	0	4	4	4
记录 2	0	0	0	4	4	4	4	4
记录 3	0	0	4	4	4	4	8	8
记录 4	2	2	4	4	4	8	8	8
记录 5	2	2	4	4	8	8	8	8
记录 6	2	2	2	4	8	8	8	8
记录 7	2	2	2	2	8	8	8	8
记录 8	2	2	2	2	8	8	8	8

(c)

图 2—10　面域的直接栅格编码

这是最简单最直接的一种栅格编码方法。通常称这种编码为栅格文件或格网文件。它不采用任何压缩数据的处理，因此是最直观最基本的格网存贮结构。

2. 行程编码

它又称为游程长度编码 (Run Length Encoding)，是栅格数据压缩的重要编码方法，也是图像编码中比较简单的方式之一。在行程编码中，将原图只表示属性的数据阵列变成数据对 (s_i, l_i) 的序列，其中 s_i 表示属性值，l_i 表示行程。对图 2—10 的格网数据，按行程编码可以表示成图 2—11*a*。

这里所记录的是属性及其行程。这种方式的缺点是位置不明显。还有一种行程编码是按行程终点的列数编码。从列数的变化也可以推断行数的变化。它的编码如图 2—11*b* 所示。可见，只要列数为 8，下一个属性便是下一行的内容。

0	5	4	3					0	5	4	8				
0	3	4	5					0	3	4	8				
0	2	4	4	8	2			0	2	4	6	8	8		
0	2	4	3	8	3			0	2	4	5	8	8		
2	2	4	2	8	4			2	2	4	4	8	8		
2	3	4	1	8	4			2	3	4	4	8	8		
2	4	8	4					2	4	8	8				
2	4	8	4					2	4	8	8				

(a)　　(b)

图 2—11　行程编码方式

从以上数据可以看出，属性的变化愈少，行程愈长，则压缩的比例越大。或者说，压缩比的大小与图的复杂程度成反比。因此这种编码方式最适合于类型区面积较大的专题图、遥感影像分区集中的分类图，而不适合于类型连续变化或类别区域分散的分类图。

这种编码在栅格加密时，数据量不会明显增加，压缩效率高，它最大限度地保留了原始栅格结构，编码解码运算简单，且易于检索、叠加、合并等操作，因而这种压缩编码方

法得到广泛应用。

3. 链式编码

链式编码又称为弗里曼(Freemon)链码或边界链码。它将线状地物或区域边界表示为：由某一起始点和在某些基本方向上的单位矢量链组成。单位矢量的长度为一个栅格单元，每个后续点可能位于其前继点的 8 个基本方向之一，如图 2—12*a*。于是，像图 2—7 所示的线实体和面实体可编码为如图 2—12*b* 和表 2—13 所示的方式。具体编码过程是：起始点的寻找一般是遵从从上到下、从左到右的原则。当发现没有记录过的点，而且数值也不为零时，就是一条线或边界线的起点。记下该地物的特征码及起点的行列数；然后按顺时针方向寻迹，找到相邻的等值点，并按八个方向编码。如遇不能闭合的线段，结束后可以返回到起始点再开始搜寻下一个线段。已经记录过的栅格单元，可将属性代码置零，以免重复编码。

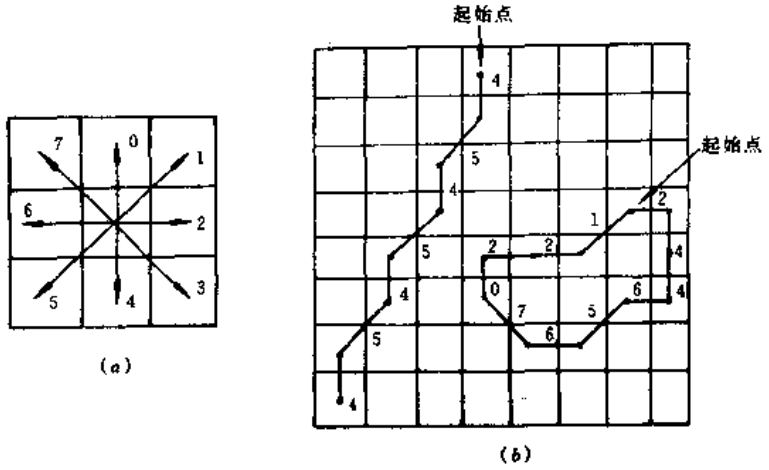


图 2—12　链式编码方式

表 2—13　链式编码表

特征码	起点行	起点列	链　码
2	1	4	4545454
5	4	7	24465570221

链码可以有效地压缩栅格数据，特别是对计算面积、长度、转折方向和凹凸度等运算十分方便。缺点是对边界做合并和插入等修改，编辑比较困难。这种编码有些类似矢量结构，不具有区域的性质，因此对区域空间分析运算比较困难。

第 2—4 节　四叉树与八叉树数据结构

一、四叉树概述

四叉树实际上是栅格数据结构的一种压缩数据的编码方法。近年来对这种结构进行了大量的研究，取得了许多新成果，应用前景广阔。

四叉树结构的基本思想是将一幅栅格地图或图像等分为四部分。逐块检查其格网属性值（或灰度）。如果某个子区的所有格网值都具有相同的值，则这个子区就不再继续分割，否则还要把这个子区再分割成四个子区。这样递次地分割，直到每个子块都只含有相同的属性值或灰度为止。

图 2—13 表示对图 2--10*a* 的分割过程及其关系。这四个等分区称为四个子象限，按顺序为左上（NW）、右上（NE）、左下（SW），右下（SE）。用一个树结构表示，如图 2—14。

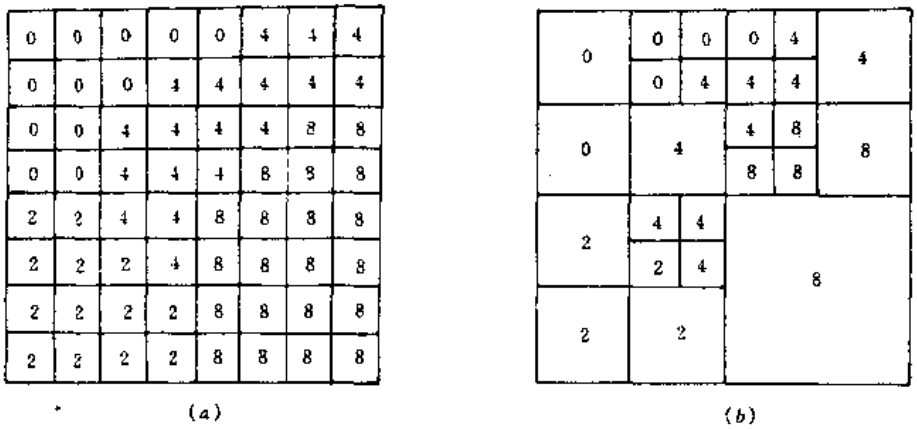


图 2—13 四叉树形成图

a—原栅格图；*b*—四叉树图

这种从上而下的分割需要大量的运算，因为大量数据需要重复检查才能确定划分。当 $n\times n$ 的矩阵比较大，且区域内容要素又比较复杂时，建立这种四叉树的速度比较慢。

另一种是采用从下而上的方法建立。对栅格数据按一定的顺序进行检测。如果每相邻四个网格值相同则进行合并，逐次往上递归合并，直到符合四叉树的原则为止。这种方法重复计算较少，运算速度较快。

四叉树结构按其编码的方法不同又分为常规四叉树和线性四叉树。

常规四叉树除了记录叶结点之外，还要记录中间结点。结点之间借助指针联系，每个结点需要用六个量表达：四个叶结点指针，一个父结点指针和一个结点的属性或灰度值。这些指针不仅增加了数据贮存量，而且增加了操作的复杂性。常规四叉树主要在数据索引和图幅索引等方面应用。

线性四叉树则只存贮最后叶结点的信息。包括叶结点的位置、深度和本结点的属性或

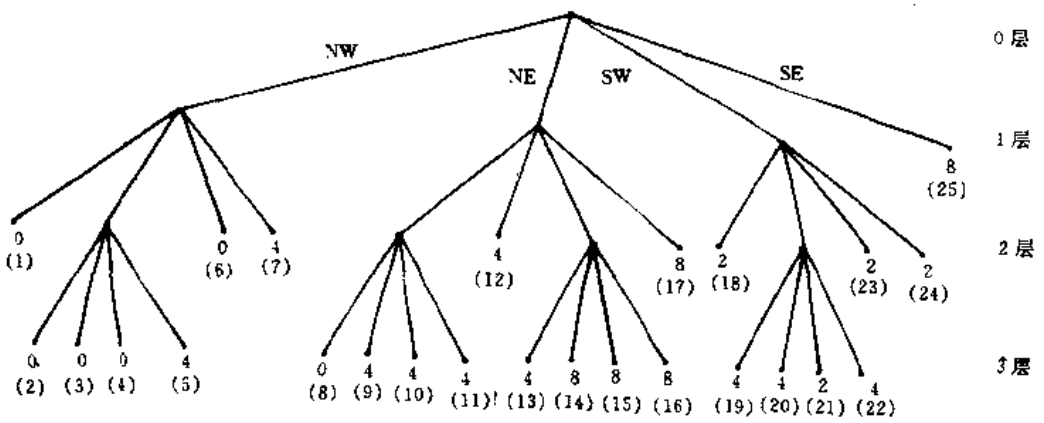


图 2—14 四叉树的树状表示

灰度值。所谓深度是指处于四叉树的第几层上。由深度可推知子区的大小。如图 2—13 和图 2—14 所示，原图 64 个栅格单元，所建立的四叉树有 25 个叶结点，它们分布在不同的层上。由于线性四叉树只贮存每个叶结点的三个量，数据量比常规四叉树大为减少，因而应用广泛。

线性四叉树叶结点的编号需要遵循一定的规则，这种编号称为地址码，它隐含了叶结点的位置和深度信息。最常用的地址码是四进制或十进制的 Morton 码。

二、基于四进制的线性四叉树编码

对于一个 $n\times n$ ($n=2^k$, $k>1$) 的栅格方阵组成的区域作四叉树编码，其中 k 称为分辨率。第一步分割成四个子象限，它们分别包括：

$$\left.\begin{aligned}
&P_0\supset\ P\ [i,\ j]\ (i=1,\ \tfrac{1}{2}n;\ j=1,\ \tfrac{1}{2}n)\\
&P_1\supset\ P\ [i,\ j]\ (i=1,\ \tfrac{1}{2}n;\ j=\tfrac{n}{2}+1,\ n)\\
&P_2\supset\ P\ [i,\ j]\ (i=\tfrac{n}{2}+1,\ n;\ j=1,\ \tfrac{1}{2}n)\\
&P_3\supset\ P\ [i,\ j]\ (i=\tfrac{n}{2}+1,\ n;\ j=\tfrac{n}{2}+1,\ n)
\end{aligned}\right\}\qquad(2\text{—}1)$$

如果要再分割下一层，其子象限分别为：

$$\left.\begin{aligned}
&P_{00}\supset\ P\ [i,\ j]\ (i=1,\ \tfrac{1}{4}n;\ j=1,\ \tfrac{1}{4}n)\\
&P_{01}\supset\ P\ [i,\ j]\ (i=1,\ \tfrac{1}{4}n;\ j=\tfrac{n}{4}+1,\ \tfrac{n}{2})\\
&\cdots\cdots\\
&P_{10}\supset\ P\ [i,\ j]\ (i=1,\ \tfrac{1}{4}n;\ j=\tfrac{n}{2}+1,\ \tfrac{3}{4}n)\\
&\cdots\cdots\\
&P_{33}\supset\ P\ [i,\ j]\ (i=\tfrac{3}{4}n+1,\ n;\ j=\tfrac{3}{4}n+1,\ n)
\end{aligned}\right\}\qquad(2\text{—}2)$$

式中“⊃”表示包含，标号 0、1、2、3 分别表示左上、右上、左下、右下四个子象限。根据上述公式可以求得任意一个小象限在全区的位置。在分割过程中，标号的位置不断增加，其标号即为 Morton 码，用 M_Q 表示。M_Q 的每一位都是不大于 3 的四进制数，并且每经过一次分割增加一位数字。分割的次数越多，所得到的子区就越小，相应的 M_Q 位数越多。最后叶结点的 Morton 码是所有各位上相应的象限值相加，即

$$M_Q=q_1q_2q_3\cdots\cdots q_k=q_1\times10^k+q_2\times10^{k-1}+\cdots\cdots q_k$$

正如前面已说过，这种自上而下分割的方法需要大量重复运算，因而应用得较少。对于自下而上合并的方法，则是将二维矩阵每个元素的下标转换成 Morton 地址码，并将元素按码的升序排列成线性表。其建立的过程如下：

先将十进制的行列号转换成二进制数表示，然后按下式计算每个栅格单元对应的 Morton 码：

$$M_Q=2I_b+J_b$$

式中 I_b 为二进制行号，J_b 为二进制列号，对于一个 $2^3\times2^3$ 的区域，其 Morton 码如图

M_Q 码 行号 / 列号		JJ	0	1	2	3	4	5	6	7
		J_b	0	1	10	11	100	101	110	111
II	I_b									
0	0		000	001	010	011	100	101	110	111
1	1		002	003	012	013	102	103	112	113
2	10		020	021	030	031	120	121	130	131
3	11		022	023	032	033	122	123	132	133
4	100		200	201	210	211	300	301	310	311
5	101		202	203	112	213	302	303	312	313
6	110		220	221	230	231	320	321	330	331
7	111		222	223	232	233	322	323	332	333

图 2—15　基于四叉树的 Morton 码

2—15 所示。

在排好序的线性表中，依次检查四个相邻的 M_Q 码对应的栅格值，如果相同则合并为一个大块，否则将四个格网值记盘，内容包括 M_Q 码、深度和格网值。这一轮检测完后依次检查四个大块的格网值，如相同就再合并，不同则分别记盘。如此循环，直到没有能够合并的子块为止。

三、基于十进制的线性四叉树编码

基于四进制的线性四叉树虽然直观上很切合于四叉树的分割，但是大多数语言不支持四进制变量，需要用十进制的长整型量表示 Morton 码，这是一种浪费；同时线性表的排序过程也要花费较多的时间。因此逐渐采用十进制的 Morton 码作为线性四叉树的地址码，并且采用自下而上的合并方法建立四叉树。这种十进制的 Morton 码（简称 M_D 码）是从 0 到 M 的自然数，合并过程的扫描方法可直接按这种自然数码的顺序进行。前后两个 M_D 码之差即代表了叶结点的大小，因而也就省去了叶结点深度的贮存。

计算 M_D 码的方法很多，现只介绍一种按位操作的运算法。设十进制表示的行、列号在计算机内部的二进制数分别为

$$II=(i_n i_{n-1}\cdots\cdots i_3 i_2 i_1),\qquad JJ=(j_n j_{n-1}\cdots\cdots j_3 j_2 j_1)$$

十进制的 Morton 码实际上是 II、JJ 的二进制数字交叉结合的结果，即

$$II=i_n i_{n-1}\cdots\cdots i_3 i_2 i_1,\qquad JJ=j_n j_{n-1}\cdots\cdots j_3 j_2 j_1$$

$$M_D=i_n j_n i_{n-1} j_{n-1}\cdots\cdots i_3 j_3 i_2 j_2 i_1 j_1$$

将得到的 M_D 由二进制数再转换为十进制数即可。用类似的方法，也可以由 M_D 码反求栅格单元的行列号。一个 $2^3\times 2^3$ 的区域得到的 M_D 码的结果见图 2—16。

M_b 码 列号 / 行号		JJ	0	1	2	3	4	5	6	7
		J_b	0	1	10	11	100	101	110	111
II	I_b									
0	0		0	1	4	5	16	17	20	21
1	1		2	3	6	7	18	19	22	23
2	10		8	9	12	13	24	25	28	29
3	11		10	11	14	15	26	27	30	31
4	100		32	33	36	37	48	49	52	53
5	101		34	35	38	39	50	51	54	55
6	110		40	41	44	45	56	57	60	61
7	111		42	43	46	47	58	59	62	63

图 2—16　基于十进制的 Morton 码

例如图中的第五行第六列，其二进制的行列号分别为。

$$II=0101,\ JJ=0110$$

得到的 M_D 码为

$$II=0101,\ JJ=0110$$

$$M_D=(00110110)_2=(54)_{10}$$

将图 2—13 按照十进制 Morton 码建立的线性四叉树如图 2—17 所示。

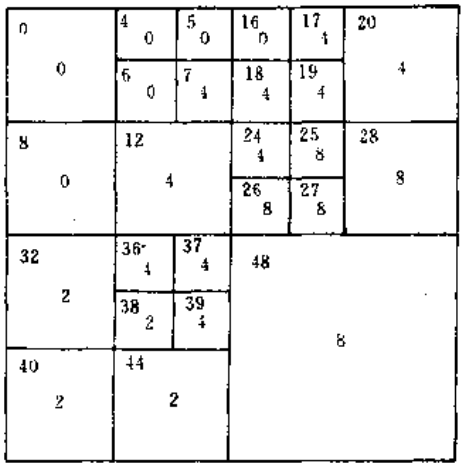


图 2—17　按十进制 Morton 码建立的线性四叉树

四、二维行程编码

在生成线性四叉树之后，仍然存在前后叶结点的值相同的情况，因而可以进一步压缩数据，将前后值相同的叶结点合并，形成一个新的线性表列。如图 2—18*a* 线性四叉树的线性表，是按 Morton 的大小顺序排列的。可以看出，在这个表中还有属性值相同而又相邻排列的情况，将值相同的叶结点合并后的编码表见图 2—18*b*。这种记录方式类似于传统的一维行程编码（也称游程编码），但是所合并的不是栅格单元，而是合并了代表范围大小不一的叶结点，所以称它为二维行程编码。比较两个表可以看出，二维行程编码又进一步压缩了数据。

二维行程编码采用了线性四叉树的地址码（Morton 码），并按照码的顺序完成编码，但却是没有结构规律的四叉树。二维行程编码比规则的四叉树更节省存贮空间，而且有利于以后的插入、删除和修改的操作。它与线性四叉树之间的相互转换也非常容易和快速，因此可将它们视为相同的结构概念。

五、八叉树结构与十六叉树结构

1. 三维与四维数据结构的提出

前面介绍的数据结构都是二维的，此时，对于空间目标的第三维坐标，例如高程，则

是将它作为属性值来对待。然而在有些信息系统中，需要有真三维的空间体系。例如矿山开采中的地下资源埋藏和采矿巷道的空间分布，如果用二维的坐标体系就根本无法很好表达。此外，矿山空间目标往往随时间不断变化着，这就提出了空间和时间信息系统的问题。

在时间信息系统中不考虑时间，是把时间看作不变的常数，即为当前的时间。而在时间和空间信息系统中，则把时间看作有过去、现在和将来的可变值。这种系统中，空间和时间是不可分割的信息，并起着同样重要的作用。

我们首先用三维来定义空间目标，在同一坐标系统下，用四维数据来定义时间和空间数据。根据这种方案，任何目标都可以由其坐标对（$\{s\}$，t）唯一确定。这里 $\{s\}=\{x, y, z\}$ 定义空间数据，而 t 定义时间数据。对每一个三维坐标数据 (x, y, z)，必定有而且仅有一个时间 t 值与之相对应，但反之则不然。

为了表示三维数据和四维数据，较好的数据结构方式是在四叉树基础上发展起来的八叉树和十六叉树结构。

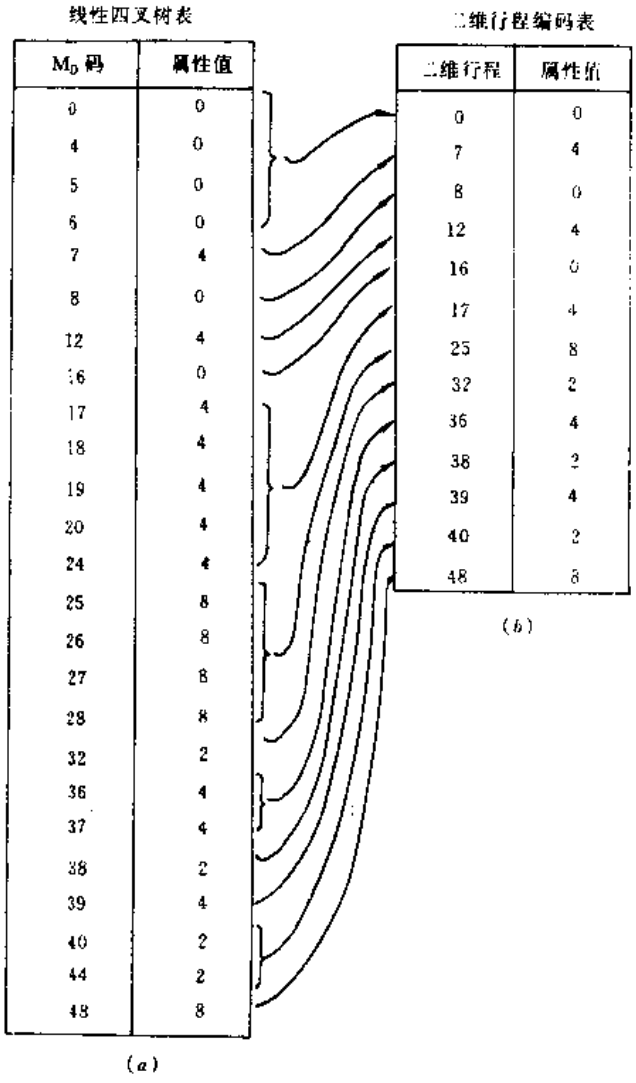


图 2—18　二维行程编码的建立过程

2. 八叉树结构及其编码

八叉树结构是从八叉树结构直接发展而来的，其原理就是将空间区域不断地分解为八个同样大小的子区域（即将一个六面的立方体再分解为八个相同大小的小立方体），分解的次数越多，子区域就越小，一直到同一区域的属性单一为止。按从下而上合并的方式来说，就是将研究区空间先按一定的分辨率将三维空间划分为三维栅格网，然后按规定的顺序每次比较 8 个相邻的栅格单元，如果其属性值相同则合并，否则就记盘。依次递归运算，直到每个子区域均为单值为止。

八叉树同样可分为常规八叉树和线性八叉树。常规八叉树的结点要记录十个值，即八个指向子结点的指针，一个指向父结点的指针和一个属性值（或标识号）。而线性八叉树则只需要记录叶结点的地址码和属性值。因此，它的主要优点是，其一节省存储空间，因为只需对叶结点编码，节省了大量中间结点的存储。每个结点的指针也免除了，而从根到某一特定结点的方向和路径的信息隐含在定位码之中，定位码数字的个位数显示分辨率的高低或分解程度；其次，线性八叉树可直接寻址，通过其坐标值则能计算出任何输入结点的定位码（称编码），而不必实际建立八叉树，并且定位码本身就是坐标的另一种形式，不必

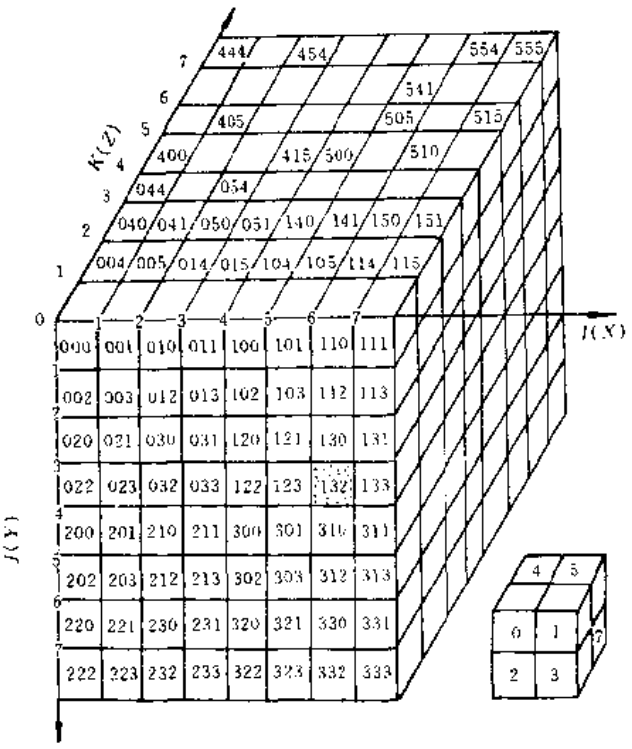


图 2—19　八叉树编码图解（$n=3$）

有意去存储坐标值。若需要的话还能从定位码中获取其坐标值(称解码)；其三，在操作方面，所产生的定位码容易存储和执行，容易实现象集合、相加等组合操作；此外，如果应用任务很大致使在核心存储器中不能容纳所有定位码时，也可以将定位码安排在 B 树中，以便划分成许多页面，并存储在外围设备中。

图 2—19 表示一种线性八叉树的生成。通过对八叉树及其编码方案的分析研究，可总结出下列一些自动计算定位码的有用规则：

(1)给定分辨率 n，也即确定了坐标系统的大小。每个坐标轴的取值范围从 0 到 2^n-1。在图 2—19 中的分辨率等于 3，则三个坐标轴的坐标范围为从 0 到 7。

(2) 每个八叉树点的编码采用的形式如图 2—20 所示。每个位置上的 q_i 是 (0，1，2，…，7) 八个数之一，q_i 的个数取决于分辨率 n。在本例中定位码以三个八进制数表示。

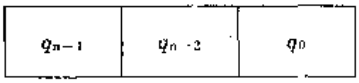


图 2—20　编码方式

(3) 对整个编码按 Z 字形的方式进行，其方向取决于如何选择第一、二、三和第四维坐标。

基于第一个规则，坐标 (X, Y, Z) 已知的八叉树点的编码可用下式表示：

$$\left.\begin{aligned} X&=c_{n-1}2^{n-1}+c_{n-2}2^{n-2}+\cdots+c_0\\ Y&=d_{n-1}2^{n-1}+d_{n-2}2^{n-2}+\cdots+d_0\\ Z&=e_{n-1}2^{n-1}+e_{n-2}2^{n-2}+\cdots+e_0 \end{aligned}\right\} \qquad (2—3)$$

式中系数 c_i、d_i 和 e_i $(i=0, \cdots, n-2, n-1)$ 的取值为 0 或 1，可通过比较方程式两边的值来确定。

根据第二、三规则的观点，q_i 值可由下式确定：

$$q_i=e_i2^2+d_i2^1+c_i \quad (i=n-1, n-2, \cdots, 0) \qquad (2—4)$$

取 $n=3$，$(X, Y, Z)=(6, 3, 1)$ 为例，则八叉树码的表达式为

$$6=1.2^{3-1}+1.2^{3-2}+0$$
$$3=0.2^{3-1}+1.2^{3-2}+1$$
$$1=0.2^{3-1}+0.2^{3-2}+1$$

因此

$$c_2=1,\ c_1=1,\ c_0=0$$

$$d_2=0,\ d_1=1,\ d_0=1$$
$$e_2=0,\ e_1=0,\ e_0=1$$

q_i 的值为

$$q_2=0.2^2+0.2^1+1=1$$
$$q_1=0.2^2+1.2^1+1=3$$
$$q_0=1.2^2+1.2^1+0=6$$

因此，按给定的八叉树点所产生的定位码为 136。

线性八叉树的地址码有八进制和十进制数两种。十进制地址码亦称 Morton 码。由于 Morton 码是自然数码，所以可以将二维数组转化成以 Morton 码为下标的一维数组。线性八叉树的自然数编码的表达式为

$$\begin{aligned}N=&2^0\ (c_0 8^0+c_1 8^1+\cdots+c_{n-1}8^{n-1})\\&+2^1\ (d_0 8^0+d_1 8^1+\cdots+d_{n-1}8^{n-1})\\&+2^2\ (e_0 8^0+e_1 8^1+\cdots+e_{n-1}8^{n-1})\end{aligned}\qquad(2—5)$$

式中，c_0，d_1，c_{n-1}，d_0，d_1，d_{n-1}和 e_0，e_1，e_{n-1}分别是行号和列号二进制化后由低位到高位的权。

当作图形、图像恢复时，可由 N 码经逆变换为下列十进制的行、列号：

$$I_1=\sum_{k=n-1}^{k=0}\mathrm{MOD}\ (T_k,\ 2)\ \cdot 2^k\qquad(2—6)$$

其中 $T_k=\mathrm{INT}\ (N_k/8^k)$；当 $k<n-1$ 时，$N_k=\mathrm{MOD}\ (N_{k+1},\ T_{k+1}\cdot 8^{k+1})$；当 $k=n-1$ 时，$N_k=N$。

$$I_2=\sum_{k=n-1}^{k=0}\mathrm{MOD}\ (T_k,\ 2)\ \cdot 2^k\qquad(2—7)$$

其中 $T_k=\mathrm{INT}\ (N_k/8^k\cdot 2)$；当 $k<n-1$ 时，$N_k=\mathrm{MOD}\ (N_{k+1},\ T_{k+1}\cdot 8^{k+1}\cdot 2)$；当 $k=n-1$ 时，$N_k=N$。

$$I_3=\sum_{k=n-1}^{k=0}\mathrm{MOD}\ (T_k,\ 2)\ \cdot 2^k\qquad(2—8)$$

其中 $T_k=\mathrm{INT}\ (N_k/8^k\cdot 2^2)$；当 $k<n-1$ 时，$N_k=\mathrm{MOD}\ (N_{k+1},\ T_{k+1}\cdot 8^{k+1}\cdot 2^2)$；当 $k=n-1$ 时，$N_k=N$。

3. 线性 2^m 叉树的自然数编码及其收敛性分析

基于线性四叉树和线性八叉树的数据结构，可以推想，对于一个四维的时空目标也可以用线性十六叉树来表示。因此推求线性 2^m 叉树的自然数编码的通式，不仅有理论意义，还便于指导实际应用。

根据以上的推导，可以得到线性 2^m 叉树自然数编码的一般公式为

$$N=\sum_{i=1}^{m}2^{i-1}\sum_{k=0}^{n-1}\mathrm{MOD}\ (I_{ik},\ 2)\ \cdot\ (2^m)^k\qquad(2—9)$$

其中

$$I_{ik}=\mathrm{INT}\ (I_{i(k-1)}/2)\qquad(k>0)$$

式中 m 为研究目标的维数，I_{ik}是 m 维欧氏空间中第 i 个坐标轴的十进制坐标。

我们知道图像分辨率 n 决定了栅格坐标大小，在式（2—9）中给出了二进制化的最大

取余迭代次数就是图像分辨率大小。而在同一图像分辨率下，m 维欧氏空间中不同的坐标取余迭代收敛次数为

$$n^{l}=\mathrm{INT}\ (\log_2 I_{i0})\ +1 \tag{2—10}$$

表明取余迭代速度是由坐标 I_{i0} 决定的，因此线性 2^m 叉树自然数编码的实用公式为

$$N=\sum_{i=1}^{m}2^{i-1}\sum_{k=0}^{\mathrm{INT}(\log_2 I_{i0})}\mathrm{MOD}\ (I_{ik},\ 2)\ \cdot 2^{mk} \tag{2—11}$$

其中

$$I_{ik}=\mathrm{INT}\ (I_{i(k-1)}/2)\qquad (k>0)$$

m 维图像恢复时，进行逆变换：

$$I_i=\sum_{k=n-1}^{k=0}\mathrm{MOD}\ (T_k,\ 2)\ \cdot 2^{k}\ (i=1,\ \cdots,\ m) \tag{2—12}$$

其中 $$T_k=\mathrm{INT}\ (N_k/(2^m)^k\cdot 2^{i-1})=\mathrm{INT}(N_k/2^{mk+i-1})$$

当 $k<n-1$ 时，$N_k=\mathrm{MOD}\ (N_{k+1},\ T_{k+1}\cdot\ (2^m)^{k+1}\cdot 2^{i-1})=\mathrm{MOD}\ (N_{k+1},\ T_{k+1},\ 2^{mk+m+i-1})$；

当 $k=n-1$ 时，$N_k=N_0$

同样，为了提高逆变换式（2—12）取余迭代收敛速度，应先确定 k 的取余迭代初始值，即满足条件

$$\frac{N}{2^{mk+i-1}}\geqslant 1 \tag{2—13}$$

有

$$k=\mathrm{INT}\left[\frac{1}{m}\ (\log_2 N+1-i)\right] \tag{2—14}$$

因此实用时，式（2—12）中 k 的初值用式（2—14）取代。

利用式（2—12）和式（2—14）进行图像恢复的逆变换时，取余迭代是否在 $k=0$ 处收敛呢？很明显要满足

$$\mathrm{MOD}\ (T_k,\ 2)\equiv 0 \tag{2—15}$$

恒成立，于是有 $mk+i-1\leqslant -1$。即

$$k\leqslant -\frac{i}{m}\qquad (1\leqslant i\leqslant m) \tag{2—16}$$

只有当 $i=m$ 时，$k\leqslant -1$。因此在取余迭代式（2—12）中，I_i 值一定会在 $k=0$ 处收敛。

作为上述公式应用的一个例子，图 2—21 为由一维时间和三维空间复合形成的四维时空数据结构的线性十六叉树自然数编码图。

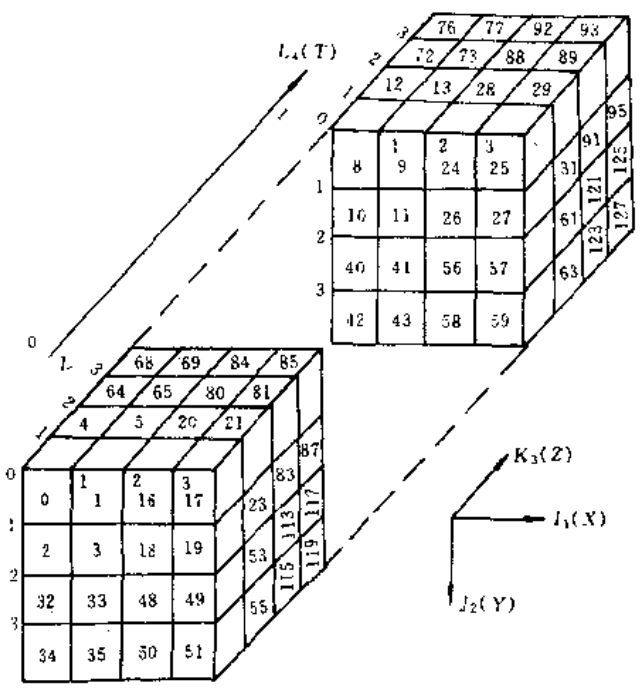


图 2—21　线性十六叉树自然数编码图解（$n=2$）

按上述公式可将四维时空坐标系中每个超体素的时空位置转换为线性 2^4 叉树自然数编码，如位于（1，3，0，1）处的超体素的编码 $N=43$。这样十进制的自然数编码

N 自然地对应一维数组下标，四维图像超体素合并运算时，依次将数组中 2^4 个相邻元素进行合并操作。四维图像恢复时，利用上述逆变换为 4 个坐标值。

线性十六叉树构成的数学模型可扩展如下：

(1) 任何十六叉树的坐标可用下式表示

$$\left.\begin{aligned} X&=c_{n-1}2^{p-1}+c_{n-2}2^{p-2}+\cdots\cdots+c_0\\ Y&=d_{n-1}2^{p-1}+d_{n-2}2^{p-2}+\cdots\cdots+d_0\\ Z&=e_{n-1}2^{p-1}+e_{n-2}2^{p-2}+\cdots\cdots+e_0\\ T&=f_{n-1}2^{p-1}+f_{n-2}2^{p-2}+\cdots\cdots+f_0 \end{aligned}\right\} \qquad (2\text{—}17)$$

(2) 定位码的十六位数字 q_i 为

$$q_i=f_i2^2+e_i2^2+d_i2^1+c_i$$

$(i=n-1,\ n-1,\ \cdots,\ 0)$

举例来说，空间和时间坐标为 $(X, Y, Z, T) = (3, 1, 2, 1)$，则利用上面公式得到地址码为 $5B$，即如图 2 -21 中后面那个立方体中灰色立方单元。而前面那个灰色单元的坐标为 (2, 1, 0, 0)，它的地址码为 12。

第 2—5 节　数据结构的转换

一、矢量与栅格结构的比较

矢量数据结构表示的数据量小而精度高，易于建立和分析图形的拓扑关系和网络关系，且目前最为流行的跟踪数字化输入和高质量的线划地图的输出都需要矢量数据结构。但是它在空间分析运算上比较复杂，特别是缺乏与遥感数据、数字高程数据直接结合的能力。

栅格数据在空间运算方面要简单得多，还能够与遥感数据和数字高程数据直接结合。但它的数据量相对较大，精度相对较低，且难于建立空间实体间的拓扑关系，不利于目标的检索等。

这两种结构各有优缺点，也有各自的特点。因此在当前的地理信息系统中，呈现出两种数据结构并存的局面，并可通过计算机软件实现这二种结构的高效转换。

二、由矢量向栅格的转换

矢量数据的坐标是平面直角坐标 (X, Y)，其坐标起始点一般取图的左下方；栅格数据的基本坐标是行和列 (I, J)，其坐标起始点是图的左上方。两种数据变换时，令直角坐标 X，Y 分别与行和列平行。由于矢量数据的基本要素是点、线、面，因而只要实现点、线、面的转换，就能实现整个线划图的转换。

1. 确定栅格单元的大小

栅格单元的大小就是它的分辨率，应根据原图的精度，变换后的用途及存贮空间等因素予以决定。如果变换后要和一幅卫星图像匹配，最好采用与卫星图像相同的分辨率。如果作为地形分析用，地形起伏变化小时分辨率可以低些，栅格单元就可大些；而地形变化大时，分辨率就应当高些，栅格单元就要小些。

栅格单元的边长在 X, Y 坐标系中的大小用 ΔX 和 ΔY 表示。设 X_{max}、X_{min}和 Y_{max}、Y_{min} 分别表示全图 X 坐标和 Y 坐标的最大值与最小值，I、J 表示全图格网的行数和列数，它们之间的关系为（参阅图 2—22）。

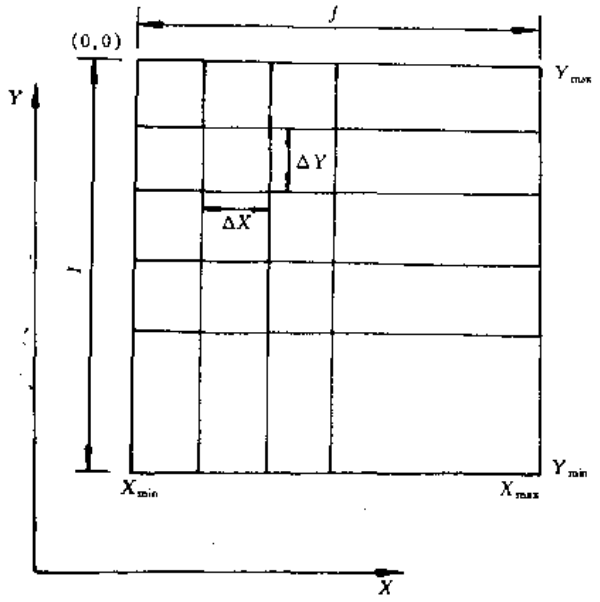


图 2—22 两种坐标关系

$$\left.\begin{aligned}\Delta X&=\frac{X_{max}-X_{min}}{I}\\ \Delta Y&=\frac{Y_{max}-Y_{min}}{J}\end{aligned}\right\}\qquad(2-18)$$

行数和列数根据分辨率确定，取整数。

2. 点的变换

点的变换很简单，只要这个点落在某一个栅格中，就属于那个格网单元，其行、列号 I、J 可由下式求出：

$$\left.\begin{aligned}I&=1+\mathrm{INT}\left(\frac{Y_{max}-Y}{\Delta Y}\right)\\ J&=1+\mathrm{INT}\left(\frac{X-X_{min}}{\Delta X}\right)\end{aligned}\right\}\qquad(2-19)$$

式中 INT 表示取整函数。栅格点的值用点的属性表示。

3. 线的变换

对于曲线可以近似地看成多个直线段组成的折线。每一条直线段需要求算它经过哪些格网单元。

设线段两端点的坐标为 (X_1, Y_1)，(X_2, Y_2)。先求出这二个端点栅格单元的行列号，然后求中间经过的栅格。

如图 2—23，设两个端点的行、列号已经求出，其行号分别为 3 和 7，则中间网格的行号必为 4、5、6。其网格中心线的 Y 坐标应为

$$Y_i=Y_{max}-\Delta Y\cdot\left(i-\frac{1}{2}\right)$$

而与直线段交点的 X 坐标为

$$X_i=\frac{X_2-X_1}{Y_2-Y_1}(Y_i-Y_1)+X_1$$

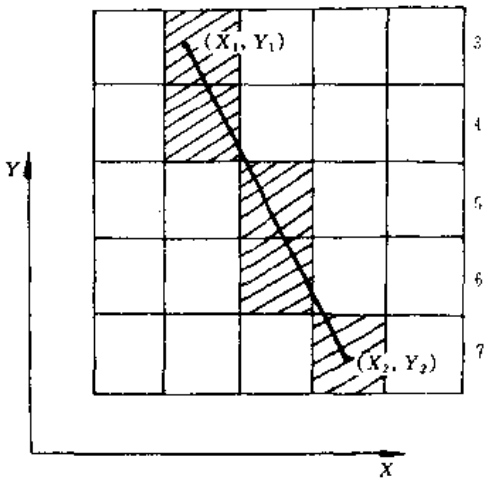


图 2—23 线的变换

由 X 值再根据式（2—19）求出这一点的列号 J。依次求出直线经过的每一个格网单元，并用直线的属性值去充填这些网格，就完成了线段的转换。整个曲线或多边形边界经分段连续运算，即可以完成曲线或多边形边界的转换。与此类似，也可以先算出两端点的列数，知道直线要经过那些列，然后计算各列中心线的 Y 值，再求出相应的行数 I。

4. 面的充填

在矢量结构中，面域用边界的线段表示，面域中间则是空白的。而在栅格结构中，整个面域所在的栅格单元都要用属性值充填，而不能用背景值 0。因此边界线段转换后，多边形面域中，还必须用属性值充填。

充填的方法很多，而关键问题是使计算机能正确判断哪些栅格单元在多边形之内，哪些栅格在多边形之外。为此多边形必须严格封闭，没有缝隙。这在数字化时和线段转换后

都要进行必要的编辑检查。

1）射线法　该法中常用的方法有平行线扫描法和铅垂线跌落法。前一种方法是从待检验的栅格单元作一平行于 X 轴的扫描线，当与多边形边界相交的点数为偶数时，则该栅格在多边形之外，当交点为奇数时，则该栅格在多边形之内。如图 2—24 的 P 点，从它所作的扫描线与多边形相交于 a、b、c、d 四个点，故 P 点在多边形之外。而从 Q 点作扫描线，则交点为 b、c、d 三个点，故 Q 点在多边形之内。有时边界会出现极值点的情况，例如图中的 e 点。此时如在 R 点作扫描线，将只有 e、f 二个交点，就会出现错误的判断。

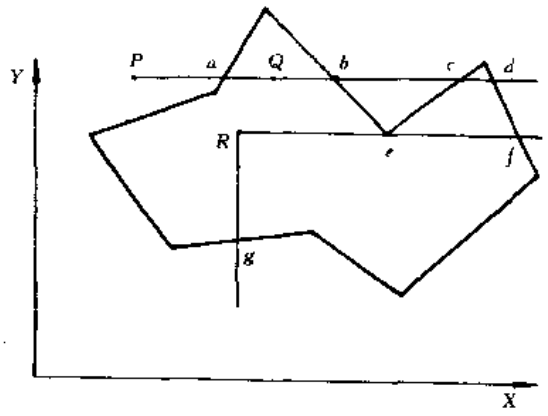


图 2—24　射线法

铅垂线跌落法则是从待检查的栅格作一垂直于 X 轴的直线，检查它与多边形边界交点的个数，奇数在多边形之内，而偶数在多边形之外。例如从 R 点作垂线只交多边形边界于 g 点，故 R 点在多边形之内。

为了避免误判，可以同时采用这两种方法检验，只要一种方法交点为奇数，该点就在多边形之内。遍历所有栅格单元，凡在多边形内的点均充填同一属性值。

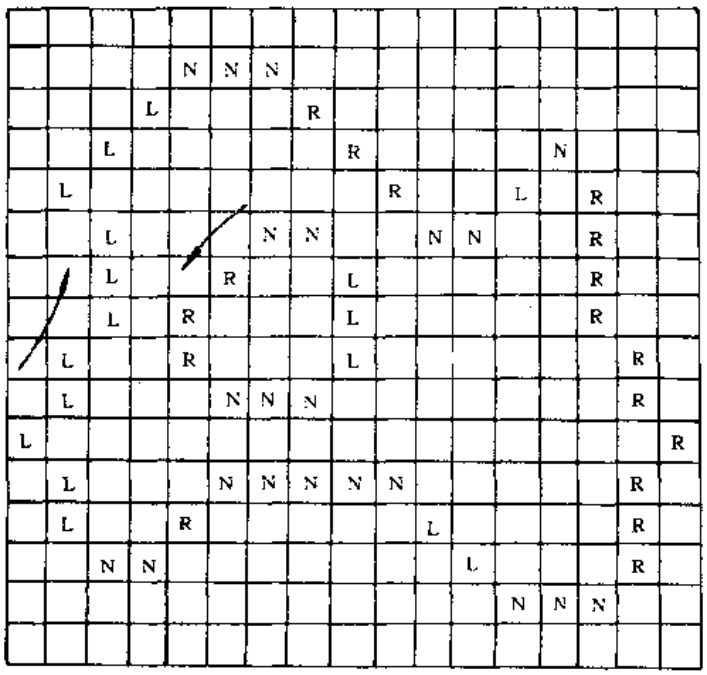


图 2—25　边界点跟踪法

2）边界点跟踪法　多边形边界的栅格单元确定后，从边界上的某栅格单元开始，按顺时针方向跟踪边界栅格，以保证多边形位于前进方向的右方。如图 2—25 所示，将边界经过的每个格网赋于字符 R、L、N 中的一个，直至回到起始点结束。其中 R 代表右，行数一直增加的单元为 R；L 代表左，行数一直减少的单元为 L；N 代表中，与相邻单元行数相同或出现极值现象的单元均赋值 N。最后逐行扫描，以多边形同一属性值充填所有 L 与 R 之间的栅格单元。对于多边形中的岛，则按逆时针方向跟踪。这样，岛区内将不充填。

3）扩散法　在多边形边界栅格确定后，寻找多边形中的一个栅格作为种子点，然后向其相邻的八个方向扩散。被扩散的栅格如果是边界栅格，就不再作为种子点向外扩散，否则就继续作为种子点向外扩散。重复上述过程直到所有种子点填满该多边形为止。

三、由栅格向矢量的转换

栅格向矢量的转换过程比较复杂，它有两种情况：一种是本身为遥感影像或已栅格化

的分类图，在矢量化之前首先要作边界提取，然后将它处理成近似线划图的二值图像（简称二值化），最后才能将它转换成矢量数据（简称矢量化）；另一种情况通常是从原来的线划图扫描得到的栅格图，二值化后的线划宽度往往占据多个栅格，这时需要进行细化处理后才能矢量化。其处理流程见图 2—26。



图 2—26 栅格-矢量转换流程图

1. 边界提取

边界提取是遥感图像处理中的一个专门问题，这里只简要介绍一种方法。这种方法是用一个 2×2 栅格的窗口，按顺序沿行列方向对栅格图像进行扫描。如果窗口内的四个格网点值相同，它们就属于一个等值区，而无边界通过，否则就存在多边形的边界或边界的结点。如果窗口内有两种栅格值，这四个栅格则均标识为边界点，同时保留原栅格的值。如果窗口内有三个以上不同的值，则标识为结点。对于对角线上两两相同的情况，由于造成多边形边界的不连通，也作为结点处理。图 2—27 为边界点的六种可能结构，图 2—28 为结点的八种可能结构。如果栅格的值中包含多边形的编号信息，则这种方法提取的边界线段也就带有左右多边形的编号信息。

边界搜索按线段逐个进行。从搜索到的某一边界窗口开始，下一点组的搜索方向由进入当前点组的搜索方向和将要搜索的后续点的可能走向决定。例如图 2—27（3）所示边界点只可能有两个走向，即下方和右方。如果该边界点组由其下方点组被搜索到，其后续点一定在其右方，而边界左右多边形的值分别为 a 和 b；反之，如果该点在其右方的点组之后被搜索到，其后续点一定在下方。其它情况依此类推。

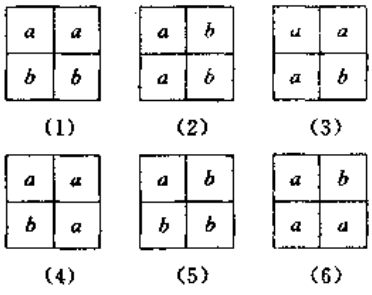


图 2—27 边界点的六种结构

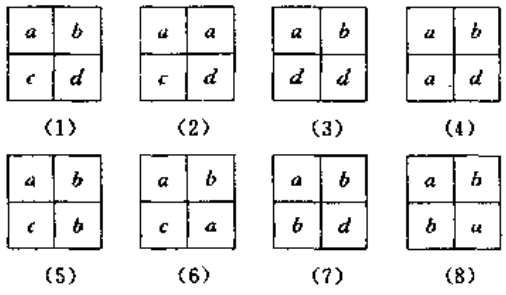


图 2—28 结点的八种结构

2. 二值化

所谓二值化就是将图像中的灰度取一个阈值，凡高于阈值的灰度取 1，低于阈值的灰度取 0。设阈值为 τ，则二值化后的像元灰度值为

$$T(i,j)=\begin{cases}1 & \text{当 } f(i,j)\geqslant\tau \text{ 时}\\ 0 & \text{当 } f(i,j)<\tau \text{ 时}\end{cases}$$

式中 $f(i, j)$ 为原像元灰度。

二值图像中的图形用 1 表示，背景用 0 表示。

3. 细化

细化也称为栅格数据的轴化，就是将占有多个栅格宽的图形要素缩减为只有单栅格宽的图形要素的过程。细化的方法很多，这里介绍两种较常用的细化方法。

1) 剥皮法　剥皮的概念就是每次删掉外层的一些栅格，直到最后留下彼此连通的由单个栅格组成的图形。剥皮的办法也有多种，其中一种的具体做法是，用一个 3×3 的栅格窗口，在栅格图上逐个检查每个栅格单元。被查栅格能否删去，由以该栅格为中心的组合图来决定，其原则是不允许剥去会导致图形不连通的栅格，也不能在图形中形成孔。

例如图 2—29 中，图 (1)，(2)，(3)，(4)，(5) 的中心栅格均可删去，因为它符合上面的原则，删去后不会造成不连通的情况，也没有形成孔。而图 (6)，(7)，(8)，(9)，(10) 的中心栅格则不能删去，因为前三个图删去后将造成不连通，后二个图删去后将生成孔。

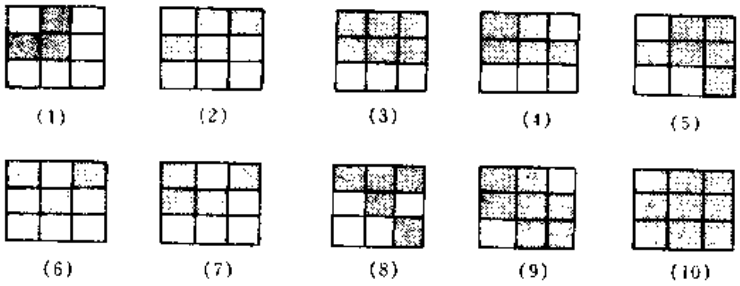


图 2—29　3×3 栅格窗口

这种组合图形共有 2^8 种排列格式。若将差异为转置 90°、180°、270°或互为镜像反射的图形合并，则总共有 51 种排列格式，其中还有 19 种可以删去。

2) 骨架法　这种方法就是确定图形的骨架，而将非骨架上的多余栅格删除。具体做法是扫描全图，凡是像元值为 1 的栅格都用 V 值取代。V 值是该栅格与北、东和北东三个相邻栅格像元值之和，即

$$V=f(i, j)+f(i-1, j)+f(i, j+1)+f(i-1, j+1)$$

如图 2—30 中的图 a 为二值化后的图像，而图 b 为其 V 值图。在 V 值图上保留最大 V 值的栅格，删去其他栅格，但必须保证连通。因为最大 V 值的栅格只能分布在图形的中心线上 (骨架上)，因此选取最大值栅格的过程就是细化的过程。

4. 矢量化

栅格数据矢量化的过程如下：第一步类似于栅格采用链码的栅格跟踪过程，找出线段经过的栅格；第二步将栅格 (i, j) 坐标变成直角坐标 (X, Y)。它是公式 (2—6) 的逆运算，每个网格中心点的坐标为

$$\left.\begin{aligned} X&=X_{\min}+\left(\Delta X\cdot j-\frac{\Delta X}{2}\right)\\ Y&=Y_{\max}-\left(\Delta Y\cdot i-\frac{\Delta Y}{2}\right)\end{aligned}\right\} \qquad (2—20)$$

矢量结构的数据点不需要像栅格那样充满路径，因此对于多余的中间点可以删除。可以用每三个点是否在一条直线上作检查。如在一直线上，则中间点可删除。对于曲线弧段，必要时还可用其他方法删除过多的中间点。

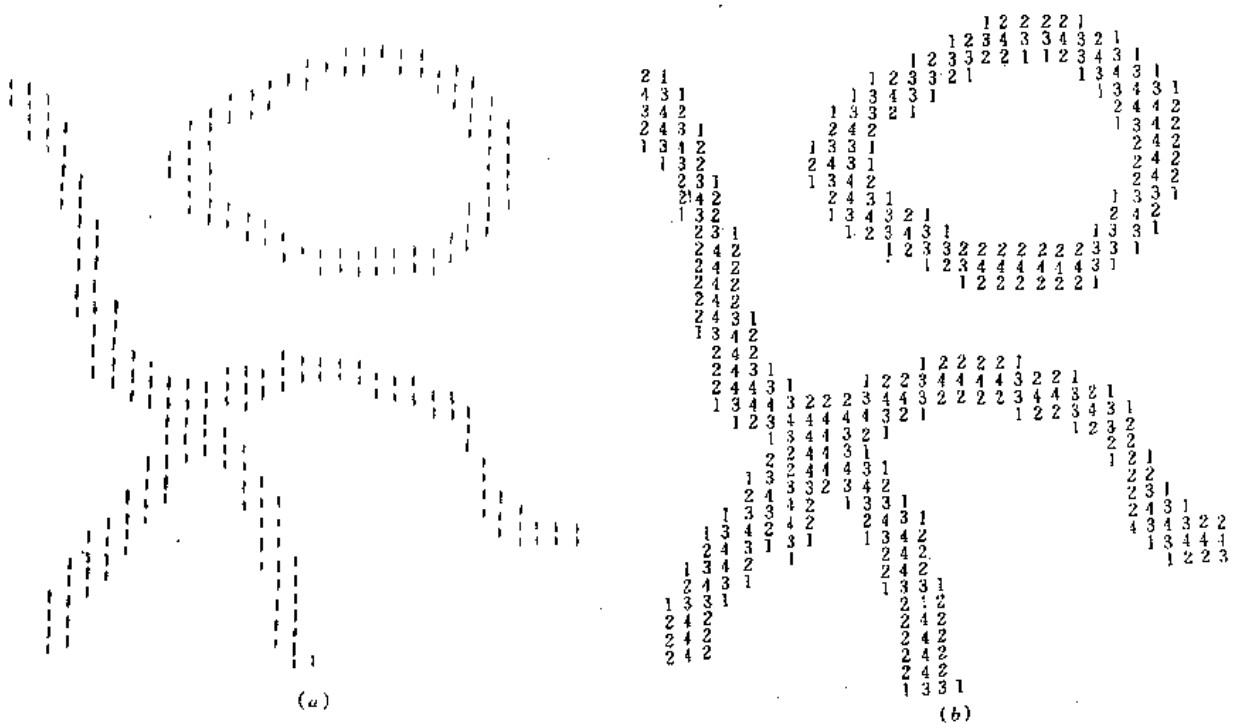


图 2--30 骨架法细化

a—原始栅格数据；*b*—骨架法栅格数据

第 2—6 节 栅格-矢量一体化数据结构

一、栅矢一体化结构的基本概念

目前，有的 GIS 软件同时具有矢量和栅格两种数据结构，并能实现两种数据结构之间的转换。但这需要增加很多的存贮空间和运算处理时间，因而并非是理想的方案。为使系统能用于多种目的，需要研究一种同时具有矢量和栅格两种特性的一体化数据结构。Molenaar 和龚健雅等对该问题作过较深入的研究，读者可参阅他们的论著。

对于面状空间目标，在基于矢量结构的 GIS 中，主要使用边界表达的方法；而在基于栅格结构的 GIS 中，一般用元子空间充填表达的方法。后者更符合于计算机数据表达的特点。对线状目标，人们习惯使用矢量方法表示。如果将矢量方法表示的线目标也用元子空间充填表达的话，就能够将矢量和栅格的概念统一起来，进而发展矢量栅格一体化的数据结构。假设在一个线性目标数字化采样时，恰好在所经过的栅格内都获得了取样点，这样的数据就具有矢量和栅格双重性质。一方面它保留了矢量数据的全部特性，目标具有明显的位置信息，并能建立拓扑关系；另一方面又建立了栅格与地物的关系，即路径上的任一点都与目标直接建立了联系。

因此，可采用填满线性目标路径和充满面状目标空间的表达方法作为一体化数据结构的基础。每个线性目标除记录原始取样点外，还记录所通过的栅格；每个面状地物除记录

它的多边形周边外，还包括中间的面域栅格。因而它既保持了矢量的特性，又具有栅格的性质。这是栅矢一体化数据结构的基本概念。从本质上说，这是一种以栅格为基础的数据结构。

为了设计点、线、面地物的具体数据结构和编码方法、需先作如下的约定：

（1）地面上（包括矿井中）的点状地物是地球表面上的点，它仅有空间位置，没有形状和面积，在计算机内部仅有一个位置数据。

（2）地面上的线状地物是地球表面的空间曲线，它有形状但没有面积。它在平面上的投影是一条连续不间断的直线或曲线，在计算机内部需要用一组元子填满整个路径。

（3）地面上的面状地物是地球表面的空间曲面，并且有形状和面积，它在平面上的投影是由边界包围的紧致空间和一组填满路径的元子表达的边界组成。

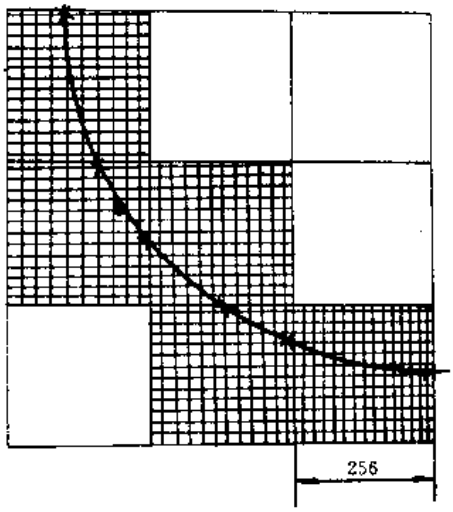


图 2—31　细化格网

由于栅格数据结构的精度较低，可利用细分格网的方法，来提高点、线（包括面状目标边界线）数据表达的精度。如图 2—31 中，在有点、线目标通过的基本格网内，再细分成 256×256 个细格网（精度要求较低时，可分成 16×16 个细格网）。为使数据格式一致。基本格网和细分格网都采用线性四叉树的编码方法，将采样点和线性目标与基本格网的交点用两个 Morton 码表示（均用十进制 Morton 码，简称 M 码）。前一 M_1 表示该点（取样点或附加的交叉点）所在的基本格网的地址码，后一个 M_2 表示该点对应的细分格网的 Morton 码，亦即将一对 X，Y 坐标转换成两个 Morton 码。例如 $X=210.00$、$Y=172.32$，可以转换成为 $M_1=275$，$M_2=2690$。

这种方法可以将栅格数据的精度提高 256 倍，而存贮量仅在有点、线通过的格网上增加两个字节。当细分格网为 16×16 时，则精度提高 16 倍，存贮量仅增加一个字节。

二、栅矢一体化数据结构设计

1. 点状地物和结点的数据结构

根据基本约定，点仅有位置、没有形状和面积，不必将点状地物作为一个覆盖层分解成四叉树，只要将点的坐标转化为 Morton 地址码 M_1 和 M_2，而不管整个构形是否为四叉树。这种结构简单灵活，不仅便于点的插入和删除操作，而且能处理一个栅格内包含多个点状目标的情况。所有的点状地物以及弧段之间的结点可以用一个文件表示，其结构如表 2—14 和图 2—32 所示。这种结构几乎与矢量数据结构完全一致。

2. 线状地物的数据结构

根据对线状地物的约定，线状地物有形状而没有面积，并且表达形状应包含整个路径。没有面积意味着线状地物和点状地物一样不必用一个完全的覆盖层分解四叉树，而只要用一串数字表达每个线状地物的路径即可。也即把该线状地物所经过的栅格地址全部记录下来。一个线状地物可能由几条弧段组成，所以应先建立弧段的数据文件，如表 2—15 和图 2—33 所示。

表 2—14　点状地物和结点的数据结构

点标识号	M_1	M_2	高程 Z
……	……	……	……
10025	43	4082	432
10026	105	7725	463
……	……	……	……

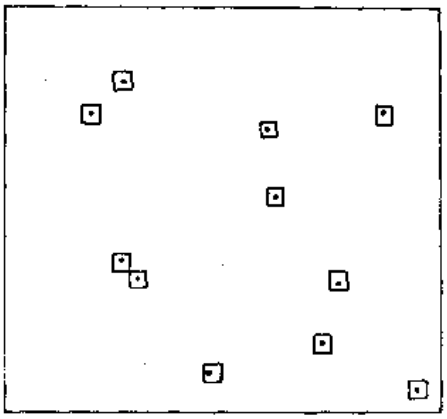


图 2- 32　点状地物数据结构示意图

表 2—15　弧段的数据结构

弧标识号	起结点号	终结点号	中间点串（M_1、M_2、M_2）
…	…	…	……
20078	10025	10026	58，7749，435，92，4377，439…
20079	10026	10032	90，432，502，112，4412，496…
…	…	…	……

表 2—15 中的起结点和终结点是该弧段的两个端点，它们与表 2—14 联接可以建立起弧段与结点之间的拓扑关系。表中的中间点串不仅包含了原始取样点（已转换成用 M_1 和 M_2 表示），而且包含了该弧段路径通过的所有格网边的交点，它所包含的 M_1 码填满了整个路径。这种结构也顾及了线性地物在地表的空间特征。

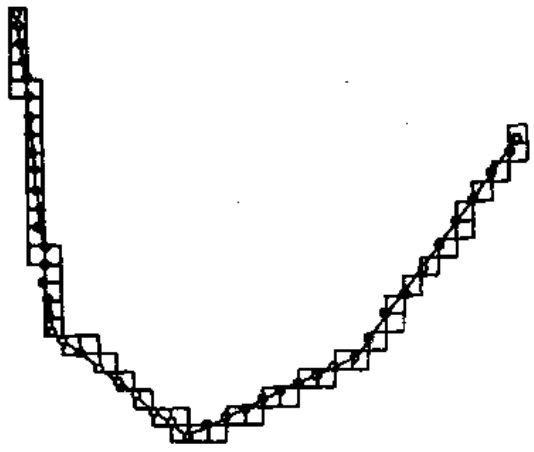


图 2—33　弧段的数据结构

虽然这种数据结构比单纯的矢量结构增加了一定的存贮量，但它解决了线性地物的四叉树表达问题，使它能与点状和面状地物一起建立统一的基于线性四叉树编码的数据结构体系。从而使点状地物与线状地物的相交，线状地物相互之间的相交，以及线状地物与面状地物相交的查询问题变得相当简便和快速。

有了弧段的数据文件，线状地物的数据结构只是它的集合表示，如表 2—16 所示。

3. 面状地物的数据结构

按照基本约定，一个面状地物应包含边界和边界所包围的整个面域。面状地物的边界

表 2—16 线状地物的数据结构

线标识号	弧段标识号
…	……
30031	20078，20079
30032	20092，20098，20099
…	……

由弧段组成，它同样可用表 2—15 那样的数据文件表示。此外，它还应包含面域的信息，而这种信息则由线性四叉树或二维行程编码表示。

各类不同的地物可以形成多个覆盖层，例如建筑物、广场等可为一个覆盖层，土地类型和煤层分布又可形成另外两个覆盖层。这里规定每个覆盖层都是单值的，即每个栅格内仅有一个面状地物的属性值。每个层可用一棵四叉树或一个二维行程编码表来表示。叶结点的值可以是属性值，也可以是目标的标识号，并且可以用循环指针将同属于一个目标的叶结点链接起来，形成面向地物的结构。图 2—34 是链接的情况，表 2—17 是对应的二维线性表。表中的循环指针指向该地物的下一个子块的记录（或地址码），并在最后指向该地物本身。如表中的 0、2、4、8 为地物的属性值，其余数字均为循环指针。只要进入第一块就可以顺着指针直接提取该地物的所有子块，避免像栅格矩阵那样为了查询某一个目标而历遍整个矩阵，从而大大加快了查询速度。

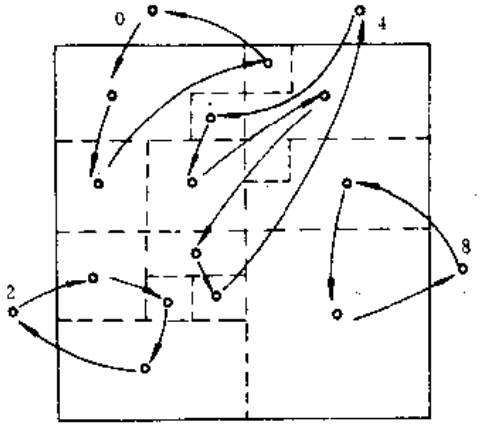


图 2—34 链接面块的指针

表 2—17 带指针的二维行程表

二维行程 M 码	循环指针属性值	二维行程 M 码	循环指针属性值
0	8	32	38
7	12	36	39
8	16	38	40
12	17	39	4
16	0	42	2
17	36	48	8
25	48		

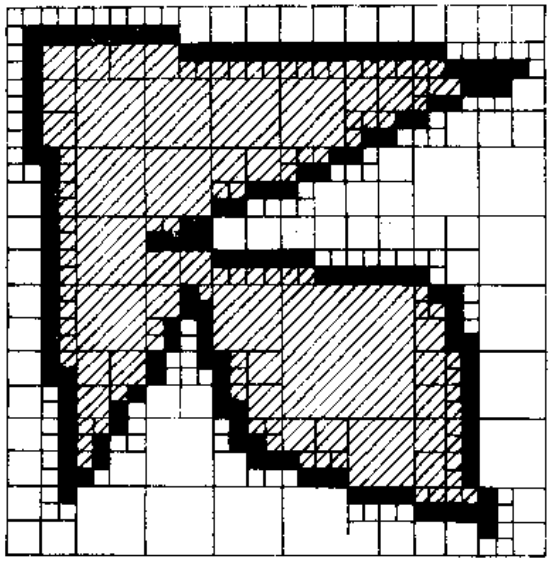


图 2—35 面状地物的数据结构示意图

对于面状地物中的边界格网，采用以面积为指标的四舍五入的方法确定其格网值，即两地物的公共格网值取决于地物面积比重大的格网。如果要求更精确地进行面积计算或叠置运算，则可进一步引用弧段的边界信息。

表 2—15 的弧段文件和表 2—17 的二维行程编码文件是面状地物数据结构的基础。图 2—35 为其示意图，文件结构如表 2—18 所示。

可见，这种数据结构是面向目标的，并具有矢量的特点，此外，通过面状地物的标识

表 2—18　面状地物的数据结构

面标识号	弧标识号串	面块头指针
40001 (A)	20001, 20002, 20003	0
40002 (B)	20002, 20004	16
40003	20004, ……	64
……	……	…

号可以找到它的边界弧，并顺着指针可提取出所有中间面块。同时这种结构又具有栅格的全部特征。表 2—17 中的 M 码表达了位置的相互关系，前后两个 M 码之间隐含了该子块的大小。一个覆盖层形成一个二维行程表，全部记录表示的面块覆盖了研究区域的整个平面。给出任意一点的位置，都可以从表中顺着指针找到面状地物的标识号，并确定是哪一类地物。

三、Voronoi 动态空间数据结构

动态空间地理数据的存储、处理和分析是 GIS 的基本要求。Voronoi 的数据结构方式是一种较好的、值得进一步研究的动态空间数据结构，它具有一些有用的特性，有利于发展综合的 IAS/GIS（图像分析系统/地理信息系统）。Voronoi 数据结构中邻域间的拓扑关系很明显，其 Voronoi 区带（zones）具有公共的边界。

这种数据结构的图形要素是 Voronoi 图形和 Delaunay 三角网。设在平面上有一组点集 S，由点 P_1、P_2、P_3……P_n 组成，其中的任一点 P_i 是最靠近它的周围各个点的重心。这样，平面上的几个点就可构成多个多边形区域，称之为 Voronoi 多边形和 Delaunay 三角网，如图 2—36 所示，其中图 a 表示 Voronoi 图形和 Delaunay 三角网的构成方法，图 b 表示由 16 个数据点集构成的 Delaunay 三角网和 Voronoi 多边形图形。

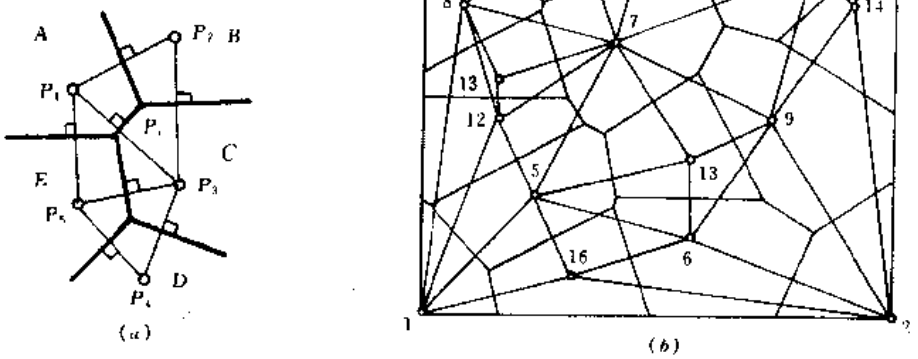


图 2—36　Voronoi 数据结构图形

a—图形的构成方法；b—由 16 个点集构成的 Delaunay 三角网和 Voronoi 图形

由图可见，Voronoi 多边形的边数多少与其周围数据点的个数有关。这种多边形具有下列特性：

（1）多边形的边界线为两邻近数据点连线的垂直二等分线。

（2）每个多边形包含一个数据点。

（3）多边形内的任何点比多边形外的任何其它点更靠近于该多边形内的数据点（只有

一个)。

由数据点构成的 Delaunay 三角网有如下特性:

(1) 三角网的网形是唯一的,即点的次序不变。

(2) 每一个三角形的内角为可能的最大角度,符合最优化的条件。

基于上述特性,若将 Voronoi 数据结构应用于综合 IAS/GIS 时,则将使之具有如下优越性:

(1) 数据结构简单,同时适用于栅格数据方式和矢量数据方式,简化了图像分类处理中辅助数据(点、线或多边形)的输入和操作。

(2) 这种数据结构和内插技术可使图像数据保持非规则的结构形式,便于将几何操作与辐射校正分别进行。此外,在管理空间数据的不确定性方面也很有潜力,可发挥其多分辨率描述的性能。

(3) 拓扑关系明显,利于作图像分割及滤波处理。

(4) 重采样算法便于按标准技术运行。

在第七章中介绍数字地面模型(DTM)和数字高程模型(DEM)时将要指出,由于不规则三角网(TIN)在表示复杂表面方面具有许多优越性,因而被广泛应用于数字制图、地形表面的模型化及分析,以及 GIS 中。如上所述,由于 Delaunay 三角网所具有的特性,它在构筑 TIN 数字地面模型中也得到广泛应用。

第三章　GIS 的数据模型

第 3—1 节　空间目标数据模型化

GIS 是人类认识世界、改造世界的强有力工具。GIS 的开发研究和应用需要经历由现实世界到观念世界的转化，而观念世界的建立是人类通过对错综复杂、色彩缤纷现实世界的认识与抽象，即各种不同专业、用户的研究和系统分析，形成 GIS 软件工具和应 用系统所需的概念模型。逻辑模型设计的任务就是把概念模型结构转换为与选用的计算机数据库系统所支持的数据模型相符合的过程。逻辑模型设计时最好应选择对某个概念模型结构最好的数据模型，然后再选定能支持这种数据模型，且最合适的数据库管理系统。而内部模型是指计算机硬件存储介质中的数据组织。这一过程如图 3—1 所示。

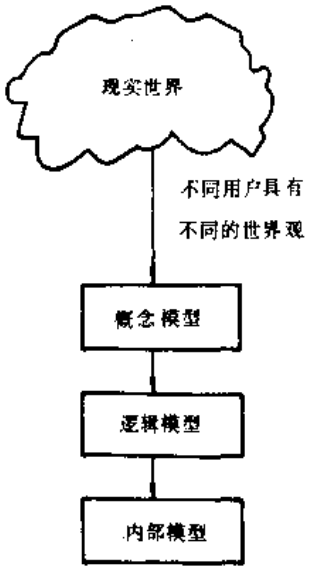


图 3—1　信息系统设计过程

所谓 GIS 的概念模型，是人们在计算机的环境下从不同角度出发和考虑，对现实世界中各种事物、彼此的关系及其发展过程认识、研究和抽象的产物。具体地说，主要包括实体对象的特征描述、关系分析和过程模拟等内容。这些内容在 GIS 的软件工具和应用系统研究中往往被抽象、概括为数据结构的定义、数据模型的建立及专业应用模型的抽象等重大问题。它们构成 GIS 基础研究的主要内容。尽管国内外的许多学者对此已作了大量的研究，但由于研究对象的复杂性和不确定性，以及应用领域的多样性，问题还远未妥善解决。

第二章介绍的数据结构是对事物实体的一些最基本的描述形式。现实世界是一个三维空间的世界。空间特性的最基本实体是点、线、面和体四大类。这些实体类型又可进一步分解为特征点、特征线、特征面和结点、界线、界面两大亚类。此外，这些实体类型的关系又是相对的：一方面，线是由点组成的，面是由线圈定的，体又是由面所包围的。它们有内在的联系，只是在构成上属于不同的层次；另一方面，随着观察这些事物的坐标系统的维数、取向及比例尺的不同或变化，它们之间的关系和内容可能按照一定的规律相互转化。例如，由三维坐标系统变为二维坐标系统后，体则变成面，面可以部分地变成线，线可以部分地变成点。取向（视角）变化之后，将使某些实体发生变化。坐标系统的比例尺缩小时，部分的体、面、线实体均可能变为点实体。图 3—2 概要地描述了三维观念世界中上述实体的分类、构成及其变化状况。点虚线框出了在二维观念世界中的情况。它们是 GIS 数据结构的基础。

然而，客观世界中的各种空间实体并不是孤立存在的，而是具有某种关系。这些关系大体上可以从实体的属性、空间及时间三个方面考察。

所谓属性关系即为对实体对象的描述，主要体现为属性多级分类体系中的从属关系、聚

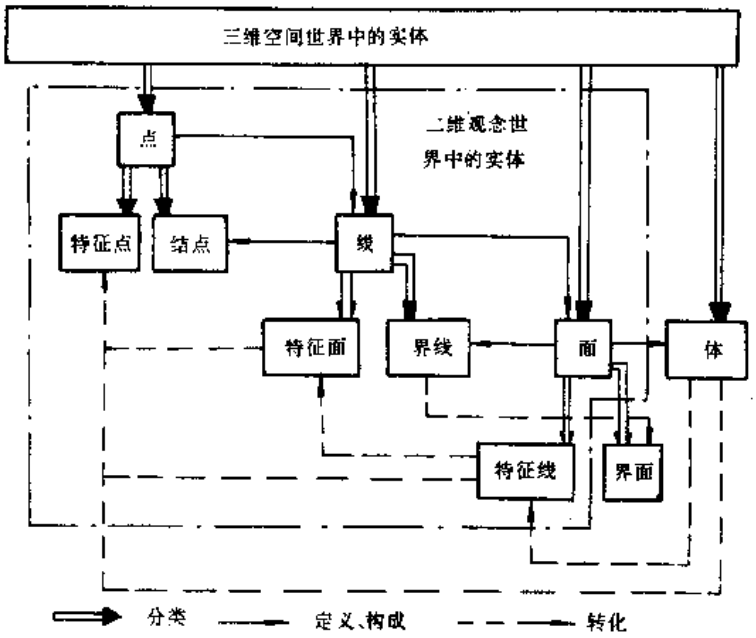


图 3—2　三维空间实体的分类与关系

类关系和相关关系。从属关系主要反映各实体之间的上下级或包含关系；聚类关系是反映各实体之间的相似程度及并行关系；相关关系则反映不同类实体之间的某种直接或间接的并发或共生关系，这种关系可以是比较确定的因果关系，也可以是比较模糊的统计关系等。

实体之间的时间关系（生命周期）一般可以通过实体变化过程来反映。有些实体数据的变化周期很长、很慢，如地球的演变，岩石、土壤、一些地形数据。而有些数据则变化很快，需要及时更新，如土地利用、城市建设中的土地覆盖、生产矿山的采场和巷道情况等。

实体的空间关系一般比较复杂。各种实体两两之间的关系可抽象、简化成如表 3—1 所示的三维或二维观念世界中的空间关系。有的学者认为，空间关系信息可以进一步分解为空间位置、空间分布、空间形态、空间关系、空间相关、空间统计、空间趋势、空间对比和空间运动等关系。空间位置描述的是空间物体的个体位置信息；空间分布是描述空间对象的群体定位信息，且通常能够从空间概率、空间结构、空间聚类、离散度和空间延展等方面予以描述；空间形态反映空间实体的形状和结构；空间关系是基于位置和形态的实体关系；空间相关是空间实体基于属性数据上的关系。例如，火成岩侵入体与围岩之间的关系就是一种复杂的空间关系和空间相关，它是地质作用的结果；空间统计是描述空间物体的数量、质量信息，又称为空间计量；空间趋势反映客观实体空间分布的总体变化规律；空间对比可以体现在数量、质量和形态三个方面；空间运动则反映空间物体随时间迁移或变化。以上种种空间信息基本反映了空间分析所能揭示的信息内涵，彼此互有区别又有联系。

对于上述实体及其关系的数学描述，可以用数据模型这个概念予以概括。建立系统数据模型的目的是，揭示客观实体的本质特性，并对它进行抽象化，使之转化为计算机能够接收、处理的数据。在 GIS 研究中，所谓空间数据模型就是 GIS 的设计者对空间世界实体进行 描述、表示的观点和数学手段，使之能反映实体系统的某些结构特性和行为功能。数据模型是衡量 GIS 功能强弱、优劣的主要因素之一。数据组织的优劣直接影响到数据库中

表 3—1 实体的空间关系

关系 / 实体	点	线	面	体
点	重叠、距离、方向	在线上、距离、方向	在面里、在界线上、距离、方向	在体内、在界面上、距离、方向
线		重叠、平行、相交、距离、方向	在面里、穿过、与一界线相交、与界线重叠、距离、方向	在体内、穿过、与一个界面相交、在界面上、距离、方向
面			重叠、平行、相交、相邻、包含、距离、方向	相截、部分相截、在体内、在界面上、距离、方向
体				重叠、包含、相交、相邻、距离、方向

数据查询、检索的方式、速度和效率。

目前，常用的数据模型有层次模型、网络模型和关系模型，以及近年来兴起的面向对象（object-oriented）模型。M. F. Goodchild 将设计空间数据模型的技术概括为两大类：分层法（layered）和对象法（object）。所谓分层法是将每一层看作一个变量，而地球表层上的每个点则为观测值。这样，地球表层可被分成许多大小不同的区域。土地利用和土壤类型是这种数据 模型技术应用之一例。对象法（或称目标法）是把世界实体看作由对象构成的地理空间。DLG-E 软件属于这种数据模型。还有的学者将 GIS 数据模型概括为面向特征和面向对象两大类。如 ARC/INFD 软件属于前者，而 system 9、MGE 软件属于后者。

实体空间的数据模型与数据结构之间的关系，并非是一一对应的关系，而往往表现为混合的交叉关系。因为客观世界是复杂的、多种多样的。在设计数据模型时应遵循的基本原则，是应使空间数据便于模型化、存储、检索和分析，即确保系统的实用性，而不是基于空间数据结构。

尽管客观实体千差万别，需要解决的问题和具体使用的专业模型有很大差别，但仍可用一个概念化的统一表达式予以表述：

$$O_{X,Y,Z,t}=f\ (E_{1X,Y,Z,t},\ E_{2X,Y,Z,t},\ \cdots\cdots E_{4X,Y,Z,t})$$

式中，$O_{X,Y,Z,t}$为专业模型中的对象实体，$E_{1X,Y,Z,t}$，$E_{2X,Y,Z,t}$，$E_{4X,Y,Z,t}$分别为模型中与 $O_{X,Y,Z,t}$ 有关的诸实体。这些实体和对象实体都与其空间坐标 X、Y、Z 以及时间坐标 t 有关，且由某种规律 f 彼此联系在一起。因此，在建立模型时需要 根据任务或问题的性质、要求，首先确定模型中的对象实体 $O_{X,Y,Z,t}$，然后再找出与它有关的诸实体 $E_{1X,Y,Z,t}$，$E_{2X,Y,Z,t}$，…… $E_{4X,Y,Z,t}$的模型边界，进而深入分析、揭示 O 与 E_i 之间在时空上的内在联系和规律，最后根据这些规律给出描述模型的数学或逻辑表达式。这就是一般建立专业模型的方法。这种模型可以是一组复杂的模型或一个包含众多相关模型的模型系统。

应该指出，目前大多数 GIS 系统均以数据库为其核心，而模型在系统中仅处于从属的地位，这是需要着力改变的状况。因此，应该加强以模型或模型库为中心的 GIS 研究。

第 3—2 节 关系模型、层次模型与网络模型

GIS 一般采用数据库管理系统对数据进行统一的管理，帮助用户查询、检索、增删和修改数据，保障数据的独立性、完整性和安全性，以利于改善对数据资源的使用和管理。而数据库由许多数据文件和文件中的大量数据组成。为了便于从一个或多个文件中存取数据，

必须采用某种方式来构造或组织数据库。那些不同的构造方（模）式就称为数据模型。目前，有三种常用的数据模型，即层次模型、网络模型和关系模型。下面以一幅由两个多边形Ⅰ、Ⅱ构成的地图 M 为例（图 3--3），说明这三种数据模型中的数据组织方式、特点及其优缺点。

一、层次数据模型

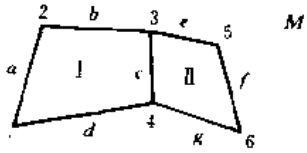


图 3—3 地图 M 及其实体Ⅰ、Ⅱ

层次数据模型是将数据组织成一对多的关系，如同一个有向树结构，即一个父记录对应于多个子记录，而一个子记录只对应于一个父记录。在土地利用的类型划分，动物、植物、岩性等的分类，环境科学管理中常用这种数据组织模型。对于如图 3--3 所示的多边形地图，可以构造出图 3—4 所示的层次模型。可见，在层次模型中各种数据之间构成一种依次的、从上至下的层次关系，层次清楚，数据存取、查询方便。

层次模型的不足之处是，不能表示多对多的联系，结构缺乏灵活性。在GIS 中，它难以顾及公共点、线数据共享和实体元素间的拓扑关系，并且不得不保留大量的索引文件，实体元素往往重复存储，导致大量的冗余数据。

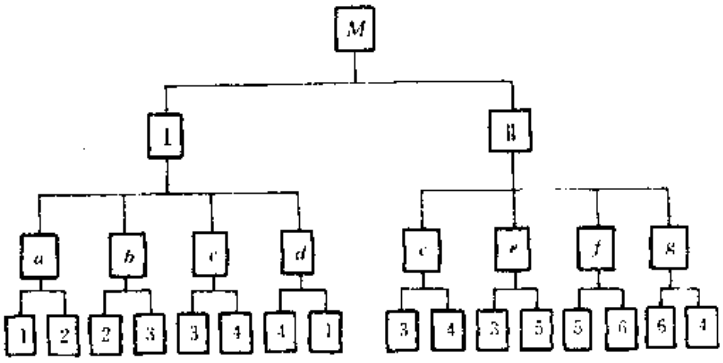


图 3--4 层次数据模型

二、网络数据模型

在层次数据模型中实体元素（数据）间的连接只限于上下通路。但在许多情况下要求多通路连接，即具有多对多的数据组织方式。例如，一个复杂的图形中，有许多图形元素需要存储在数据库的不同部分，或存储在不同的数据文件中，要把这些元素组织在一起，则需进行多通路连接，才能迅速形成图形。网络数据模型能够满足这种要求。

图 3—5 是图 3—3 的网络模型，图中的每个方格称为一个结点，代表一个实体元素，不同元素之间的联系用线联接，即一个子结点可能有多个双亲。在这种模型中使用连接指令或指针来确定数据（元素）之间的显式连接关系。

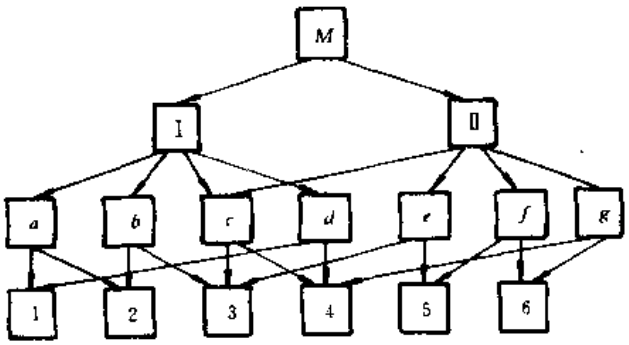


图 3 5 网络数据模型

在基于矢量的 GIS 中，图形数据通常采用拓扑数据模型，这种模型很类似于网络模型，但拓扑模型一般采用目标标识来代替网络联接的指针。例如，图 3—3 中的结点 3、4 及线 c，在层次模型中要重复存储，线 c 成为多余线。但在网络模型中就可以避免这种情况。图 3—5 中每一条线及每一对坐标都只出现一次，要去掉 c 线也是很简单的事情。网络模型的这一特点为制图综合提供了很大的方便。

网络模型的主要缺点是：指针数据项使数据增大，当数据复杂时指针部分将占去大量的存储空间；数据增删时指针也必须随着变化，因此指针的建立和维护较麻烦。

三、关系数据模型

关系数据模型是一种数学化的模型，它把数据的逻辑结构归结为满足一定条件的二维表中的元素，这种表就称为关系。一个实体由若干个关系组成，而关系表的集合就构成为关系模型。

图 3—3 所示的多边形地图，可用表 3—2 所示的关系来表示多边形与边界及结点之间的关系。其中多边形序号Ⅰ、Ⅱ称为特征值或关键字。此外，属性数据（与多边形、线或点有关的属性数据）也可用属性数据表表示之。属性数据可能有多种多样，例如，线段 a、b、c……g 可以表示道路、河流或巷道等。所有这些关系表通过某种共同特征连系起来。

表 3—2 关 系 表

多边形与边长关系

多边形号	边号	边长
Ⅰ	a	16.0
	b	15.9
	c	10.9
	d	21.2
Ⅱ	c	10.9
	e	10.3
	f	13.9
	g	13.8

边结点关系

边号	起结点号	终结点号
a	1	2
b	2	3
c	3	4
d	4	1
e	3	5
f	5	6
g	6	4

结点坐标关系

结点号	X 坐标	Y 坐标
1	10.0	8.1
2	13.6	23.5
3	29.9	23.4
4	30.5	13.0
5	39.9	22.4
6	44.2	9.9

关系模型具有数据结构简单，能够直接处理多对多的关系，数据修改和更新方便，容易维护和理解等优点。对数据（元素）之间的联系不是用指针表示，而是由数据本身通过公共值隐含地予以表示，还可以用布尔逻辑运算和数学运算规则来操作数据。因此，关系模型是当前传统数据库中最常用的数据模型。并且目前大多数 GIS 软件中的属性数据，甚至几何图形数据也采用关系数据模型，如 system 9，Geovision 等使用一种优化的面向目标的全关系型数据库，把实体的几何关系特征和属性存储在同一个网络分布式关系数据库中，去掉了空间数据库与属性数据库之间的繁琐联接关系，提高了系统的效率。图 3—6 比较了传统的专用数据库及其联接关系和全关系型 GIS 软件的数据库。

关系模型数据库的主要缺点，是许多操作都要求在文件中顺序查找满足特定关系的数据，如果数据库涉及的目标很多、很复杂，将需要花费很多时间，影响效率；在程序交互、目标标识和模型扩充方面也存在一些不足。

上述三种数据模型均可用来描述包括图形数据和属性数据的空间数据，但各有优缺点、如表 3—3 所示，在选用时应根据应用要求择优而用，或取长补短。

由于上述三种传统的数据模型都是以记录为中心的，难以表达和处理复杂的空间地理实体，难以实现快速查询和复杂的空间分析，难以实现真三维空间模型化和四维时空的模型化，也不便于系统扩充，没有提供存贮信息的解释及与用户友好的接口。因此，探索研究更接近于人类通常的思维方式，便于操作管理空间实体数据的模型，已成为 GIS 研究的重要课题。

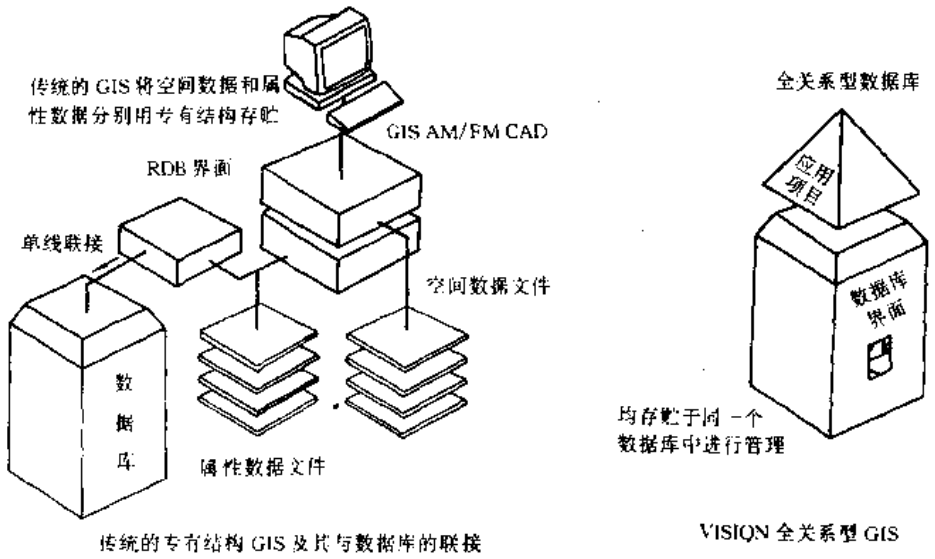


图 3 -6 全关系型数据库与传统数据库对比

表 3—3 三种常用的传统数据模型比较

数据模型	优 点	缺 点
层次模型	1. 容易理解，更新与扩充 2. 通过关键字数据访问易于实现 3. 事先知道全部可能的查询结构，数据检取方便	1. 访问限于自上而下的路径，不够灵活 2. 大量索引文件需要维护 3. 一些属性值重复多次，导致数据冗余，存贮和访问的开销增加
网络模型	1. 地图中相邻的图形特征，乃至其坐标数据易于连接 2. 在很复杂的拓扑结构中搜索，有环路指针是一种很有效的方法 3. 避免数据冗余、已有数据可充分使用	1. 间接的指针使数据库扩大，在复杂的系统中可能占据数据库的很大部分 2. 每次数据库变动，这些指针必须更新维护，其工作量相当大
关系模型	1. 结构灵活 2. 可以满足布尔逻辑和数学运算表达的各种查询需要 3. 允许对各种数据类型的搜索、组合和比较	1. 为找到满足指定关系要求的数据，许多操作涉及到对文件的顺序搜索，对大型系统而言，很费时间 2. 为保证以适宜速度进行搜索的能力，商用系统一般需经过十分精心的设计，因而价格很昂贵

第 3--3 节 综 合 数 据 模 型

如前所述，目前多数 GIS 软件是建立在关系数据库管理系统(RDBMS)基础上的，然而 RDB 的显著缺点是不能存储和操作地理空间数据，不便进行空间查询。为了克服这些缺点，当前商品化或研制的 GIS 软件大多采用综合数据模型。归纳起来有如下三种构造方式：

一、双元模型

它的基本思想是用两个子系统分别存储和检索空间数据与属性数据，其中属性数据存储在 RDBMS 中，两个子系统之间使用一种标识符联系起来。图 3—7 为其原理框图。在检索目标时必须同时询问两个子系统，然后将它们的回答结合起来。

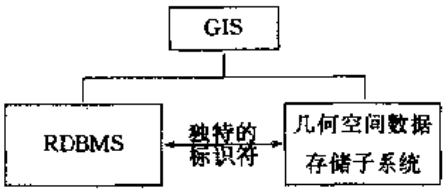


图 3—7　双元模型

由于这种双元模型的一部分是建立在标准 RDBMS 之上，故存储和检索空间数据比较有效、可靠。但因为使用两个存储子系统，它们有各自的锁定规则，查询操作难以优化，存储在 RDBMS 外面的数据有时会丢失数据项的语义；此外，数据完整性的约束条件有可能遭破坏，例如在几何空间数据存储子系统中目标实体仍然存在，但在 RDBMS 中却已被删除。

属这种模型的 GIS 软件有 ARC/INFO (ESRI)、MGE (Modular GIS Environment，Intergraph)、SICAD、ARGIS 等。

二、分层模型

双元模型的缺陷大都源于两个存储子系统具有各自的职责。那么能否在 RDBMS 中存储空间数据，而又克服其不足呢？现在人们利用标准关系数据库顶层的某些“智能”处理技术，即利用该层将地理结构查询语言(Geo SQL)转化成标准的 SQL 查询，借助索引数据的辅助关系实施空间索引操作。这种模型的空间数据存取速度较快，但由于是间接存取，在效率上总是低于 DBMS 中所用的直接操作过程，且查询过程复杂。图 3—8 为其原理框图。

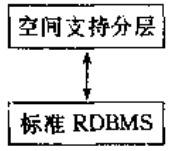


图 3—8　分层模型

这种模型的代表性 GIS 软件有 system 9，GEOVIEW（爱登堡大学）、SIRO-DBMS（澳大利亚）等。

三、综合模型

这种综合数据模型又称全 DBMS 模型。它不是基于标准的 RDBMS，而是基于可扩充的 DBMS 开发的。如图 3—9 所示，空间扩展完全包含在 DBMS 中，用户可以使用自己的基本抽象数据类型（ADT）来扩充 DBMS。在核心 DBMS 中进行数据类型的直接操作很方便、有效，并且用户还可以开发自己的空间存取算法。该模型的缺点是，用户必须在 DBMS

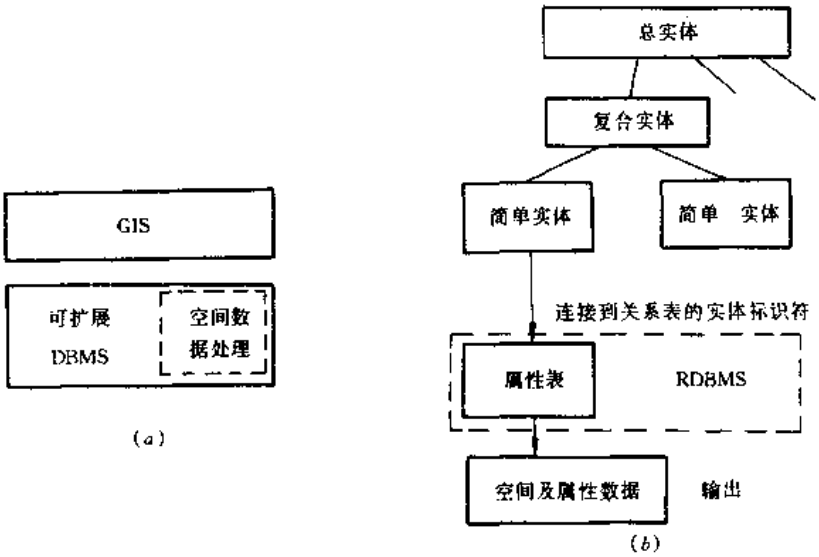


图 3　9　综合数据模型

环境中实施自己的数据类型，对有些应用将相当复杂。

属于此类综合模型的软件如 TIGRIS (intergraph)、GEO^{++}（荷兰）、GéoTropics（巴黎大学）等。而 Smallworld GIS（英国剑桥）则兼有分层和综合模型的特点。此外，经扩充后的 GEO^{++}GIS 软件充分利用 R 树存取方法的优点，具有存储、显示栅格数据的能力，即将栅格数据存储在分辨率可变的“Clusters”组件（如 32×32 或 64×64）中。

第 3—4 节　超图的概念与多媒体空间信息系统

进入 90 年代以来，计算机技术在容量及运算速度方面都发生了质的飞跃，为数据库提供了表现更丰富信息的环境。人们习惯于根据联想的思维方式采取行动，这就要求数据也具有这种基本功能，于是超文本（Hypertext）和超媒体（Hypermedia）的结构方式被提出。属性数据部分按相互关系组织成正文网，可以从任意一点以联想方式浏览查阅，其中包括图像、图形、声音、动画和连续播放等多媒体（Multimedia）方式。其中，最基本的组成部分是节点（node）和链（link）。节点用来存储对象的信息，链用来表示信息之间的关系。节点又分为正文节点、图像节点、声音节点、图形节点；链可包括导航链、名称链、聚焦链、全景链、组织推理链、索引链、属性链、组成链、隐含链等，从而构成一个更灵活的查询环境和智 能化的系统。

空间信息的综合性、动态性、空间性和快速反应特点，要求采用遥感、地理信息系统和全球定位系统（通称“3S”技术）作为动态空间信息采集和处理的技术支撑。然而，若不能解决三者之间的逻辑联接，实现它们之间的“一体化”，就难以发挥各自的优势，建立真正有效的地理空间信息系统。这里的关键问题之一是建立一体化的数据结构，以一体化数据结构为核心来组织整个系统。超图数据结构（Hypergraph Based Data Structure，HBDS）是一种很有发展前途的空间数据组织方式。“在不久的将来超图将广泛应用于空间信息系统中，或者更确切地说，未来的空间信息系统将建立在超图概念的基础上。”（Laurini，Robert and Derek Thompson，1992，Fundamentals of Spatial Information Systems，Academic Press，P. 618）因为要较好地表示包括空间与非空间属性联系的复合制图对象，就要求有性能完备的空间数据结构和模型，而超图数据结构是建立在超图和集合论的基础上，以“现象结构”理论取代传统的关系结构（Relational Structure）、层次结构（Hierarchical Structure）和网络结构（Network Structure）理论，以类别、目标、属性和联接这四种抽象数据类型作为最基本的结构框架，来描述、表示事物实体，为实现和建立一体化数据结构提供了重要的理论基础和有效的可操作的手段。

一、超图数据结构的理论框架

HBDS 最早是由法国巴黎大学的 F. Bouille 教授于 1979 年提出来的。后来经过中国学者崔伟宏研究员和美国的 Rugg 教授等的研究逐步完善起来。它的理论框架和数据组织简介如下。

根据图论有　　$G=(V.R)$

其中，V 称为顶点或端点子集，R 称为边缘或弧的子集，并且

$$V=(V1,\ V2,\ V3,\ V4,\ \cdots\cdots Vn)$$

$$R=(R1,\ R2,\ R3,\ R4,\ \cdots\cdots Rn)$$

这就是说，一个 G 的子集是在 G 内具有所有端点的弧的图。它的基本要素是，端点

(Vertex) V 是有穷的集合，VR 是两个端点间关系的集合。

假若〈$\dot{X}Y$〉$\in VR$ 表示从 X 到 Y 的一条弧线，其中，X 称为初始点或者弧头，Y 称为终点或者弧尾。此时的图称为有向图。因此，如果〈XY〉$\in VR$，就一定有〈YX〉$\in VR$，这个弧又叫边缘 (edge)。

超图是在图的基础上发展起来的。1970 年著名数学家 Berge 给超图下的定义是：一个超图是在图中两个或两个以上的端点被同一个边缘所联接。在图 3—10 中，如果

$R=(Ei/i\in I)$ 是 V 的子集

其中，Ei 为边缘，$i\in I$，$Ei\neq 0$ (i 是指标号)

$UEi=V$

则 $H=(V.R)$ 为超图表达式

$|V|=n$，称为超图的次序

$V1$，$V2$，$V3$，$V4$……Vn 为端点

$R1$，$R2$，$R3$，$R4$……Rn 为边缘

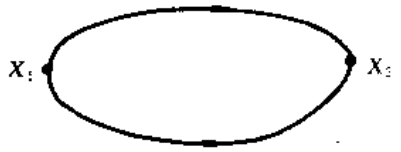


图 3—10 超图概念之一

超图也可以用图形（如图 3—11）来表示

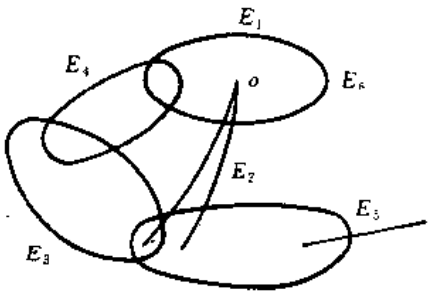


图 3—11 超图概念之二

如果 $|Ei|>2$，则 Ei 为一个圆

$|Ei|=2$，则 Ei 为一个边缘

$|Ei|=1$，则 Ei 为一个图的循环

根据上述规则，对于图 3—11 而言

$|E5|=2$，$|E2|=2$，$|E6|=1$，$|E4|>2$

超图 $H=(V.R)$ 的入射矩阵为 $((a_i^j))$，它代表 H 边缘的 m 行和端点的 n 列。

如果

$X_j\in Ei$ 则 $a_i^j=1$

$X_j\notin Ei$， $a_i^j=0$

F. Bouille 在超图和集合论的基础上，引进了抽象数据类型 (Abstract Data Type, ADT)，共有四种抽象数据类型：

〈数据〉::=〈集〉〈要素〉〈属性〉〈相关〉

在此基础上又将上述关系发展成为 HBDS 理论集的四个基础成分：

类别 Class

目标 Object

属性 Attribute

联接 Link

以及 〈数据〉::=〈类别〉〈目标〉〈属性〉〈联接〉

〈属性〉::=〈类别属性〉〈目标属性〉

〈联接〉::=〈类别间的联接〉〈目标间的联接〉

将 HBDS 的超图理论与集理论和图论作一比较，列于下表 3—4。

但是，F. Bouille 在 HBDS 模型中没能区分存在于现实世界的“实体”与空间数据中表达现实世界的“实体”。有关学者在中国和美国进行信息系统的研究中，采用更加概括的概念来覆盖客观现象，即用“entity”来描述数据库中的实体和客观现象，用“属性”和“属

性值”分别等价于 Bouille HBDS 中的“类别属性”和“实体属性”。于是在新的 HBDS 数据结构中共有六种抽象数据类型：

表 3—4 比　　较　　表

集　理　论	HBDS 超图理论	图　　论
集 Set	类别 Class	超图的边缘
要素 Element	目标 Object	超图的端点
特征 Property	属性 Attribute	端点赋值或边缘赋值
相关关系 Relation ship	相关 Relation	联接

实体（Object）

类别（Class）

属性（Attribute）

属性值（Attribute Value）

实体间的联接（Link between Objects）

类别间的联接（Link between classes）

以上六种 ADT 之间的关系可用图 3—12 予以表述。

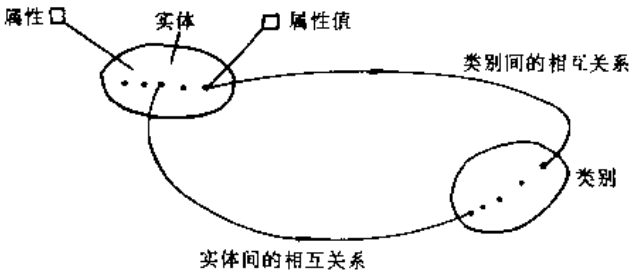


图 3—12　HBDS 中 ADT 的相关关系

二、超图数据库设计基本原理

超图数据库的设计是以超图数据结构（HBDS）为基础，以 HBDS 的数据链系统（DSS—Data String System）作为数据库中数据的基本单元，并且在 Corlet“单元结构”理论的支持下建立起来的。它将超图数据链的头、体分开，分别建立 0 单元库、1 单元库和 2 单元库，在指针快速搜索系统的支持下，进行数据库的自动维护与更新，以及 0 单元、1 单元和 2 单元的快速搜索，以实现图形及属性数据的直接联接。美国的 TIGER 数据库以及中国几个地区的农业可持续发展决策支持系统的数据库可称为超图数据库。

1. 单元结构理论

J. Corbett 的单元结构理论是在二维拓扑理论的基础上，并进而提出：二维空间平面是一个连续的空间；空间目标是多样性的，这种多样性是与拓扑等价的；任何空间目标将定位于该二维空间上，且空间目标由单元结构组成。即，

0 单元（0 Cell）是指已知的端点或结点；

1 单元（1 Cell）指连续的弧段或边缘（edge）；

2 单元（2 Cell）指区域（areas，regions）或多边形（Polygons）。

根据“单元结构”理论形成四个基本定义：

0 单元指 0 维目标，代表点或端点；

0 单元构成一单元的边界，即 1 单元为 1 维目标，其起点和终点在 0 单元上；

1 单元联接 0 单元，包括 0 单元的所有点，但不包括 0 单元自己，1 单元不能自己定义

或者与其它 1 单元相交；

2 单元是一个二维目标，以一个闭合的 1 单元作为边界。

根据上述定义，1985 年 Corbett 定义了两种相关的系集，即边界关系和结边界关系，并规定：

单元边界是小于已给单元维数 1 维的偶然点的集合；

单元的结边界是大于已给单元维数 1 维的偶然点的集合。

2. DSS 数据链系统

数据链是超图数据结构中的数据记录单元。数据链由三部分组成，即超图数据头、超图数据体和超图数据尾。DSS 数据链系统是一种相当灵活的系统，其主要优越性在于：

(1) 超图数据头与超图数据体分开。数据头将包括 ADT 抽象数据类型和存贮结构中的地址，即：C_i（类别），O_i（目标），A_i（类别属性）；aip（目标属性），$\underset{C_iC_j}{L}$（类别的联接）和$\overset{O_i^m}{\underset{O_j^m}{L}}$（目标间的联接）。

(2) 应用指针穿插技术对分别存贮的数据头和数据体进行维护，效率较高。

3. 超图数据结构数据库的构建

根据单元结构理论和数据链系统的特点，构建超图数据结构数据库的要点包括：0 单元库，1 单元数据头数据库，1 单元数据体数据库，2 单元数据头数据库和 2 单元数据体数据库。这就是说，这种数据库系统采用 HBDS 的 DSS 作为数据库组织、存贮的基本单元。一个 HBDS 的 DSS 可包括以下要素：图幅控制点，弧段，图斑，多边形标识。以上各要素分别存贮在不同的文件中。具体说明如下：

图幅控制点，是存贮地图的四个角的控制点坐标；

一单元，表示线状要素的空间位置；

二单元，记录组成最小多边形的弧段号及其相邻弧段的拓扑关系；

二单元标识，存贮多边形的构成（含岛屿的情况），记录多边形的记录号、面积和属性。

上述文件之间是相互联系的，其连接操作是通过数据项的关键码和属性进行的。这三个数据文件分别是：一单元文件 COORI；二单元文件 CHAIN 和二单元标识文件 POLY。

文件格式

* 一单元文件（COORI）

  属性 S

  $KEY1$，点坐标 $X_1Y_1 \leftrightarrow X_NY_N$（END 结束标识）

* 二单元文件（CHAIN）

  属性 S

  $CKEY$，弧段号 $KEY_1 \leftrightarrow KEY_N$（END 结束标识）

* 二单元标识文件（POLY）

  属性 S

  $PKEY$，$AREA$，$CKEY$ $CKEY'$ END

  （关键字）（面积）（岛屿）（结束标识）

二单元拓扑关系建立时，搜集外边界弧是最根本的功能，而二单元的拓扑信息由外边界弧、岛组弧及单弧组成。

搜索程序实质上是一个迭代过程。将所有同一属性弧段的起、终点坐标和拓扑信息排

列，并读入结构 reswlt 中，依次序选取 reswlt 中的弧段作为当前弧段 cid，记录该弧段的终点信息，依次序搜索 reswlt 结构中剩余的未处理过的弧段 qid，当弧段 cid 与弧段 qid 满足连接关系时，将弧段 qid 作为当前新的弧 cid 进行下一轮的搜索。每一步的迭代搜索满足如图 3—13 所示的条件时返回，构成多边形封闭或单弧段封闭。

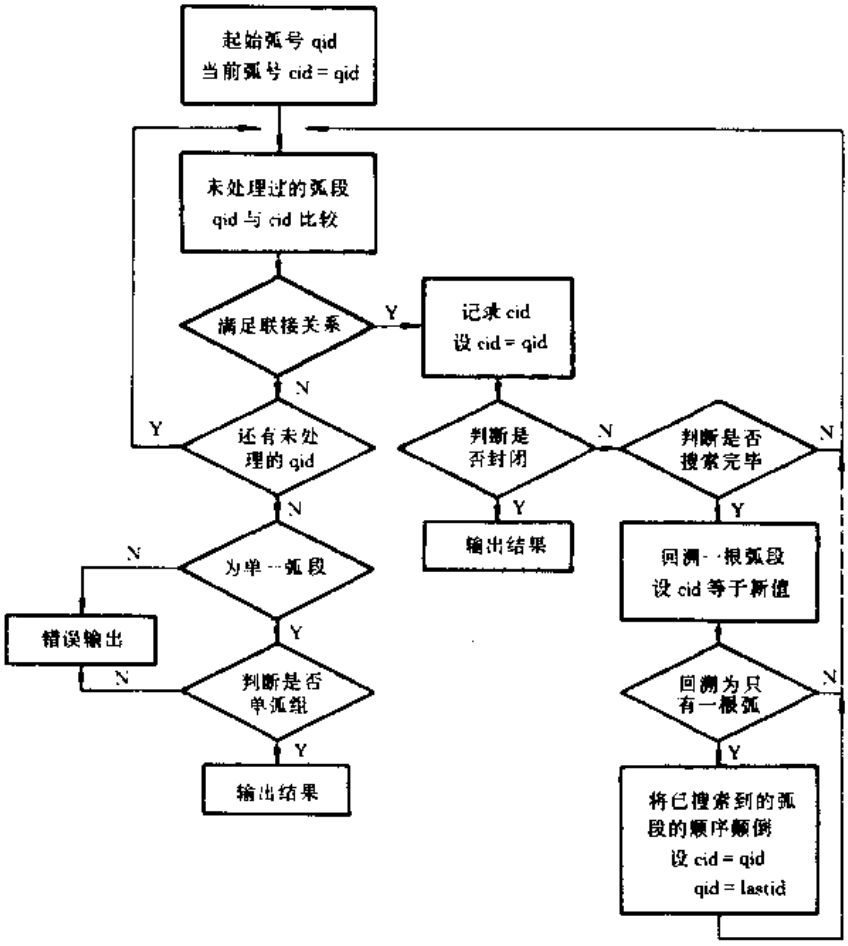


图 3—13　HBDS 数据库搜索流程图

第 3—5 节　时空 GIS 数据库系统的概念

时间是现实世界中各种事件的一种普遍属性。任何人类活动和自然现象的变化都需要经历一定的时间。虽然某些现象或事件并不一定必要记录和存贮时间，但如果 GIS 具有支持时间的功能，能将时间变量加入到 GIS 空间分析过程中去，则对广大用户的吸引力和作用就更大。在采矿工业中，特别是地下开采过程中，矿山的空间状态和环境条件是随着时间和生产的发展而变化的。今日的采场即为明日的采空区。这种采场的定位数据及其成为采空区的时间，在相当长的时期内都必须予以存贮。此外，勘测人员还要对一些未开采区进行详细探测，并将有关信息资料存贮在信息系统中。对环境生态系统发展变化过程的分析研究，也需考虑时间和空间两方面的状况和变化。在研究全球变化问题时，科技工作者往往需要以不同的时间尺度来研究自然现象所传递的信息等等。因此，实现时空复合操作，将空间分析的问题进一步拓展为时空分析的范畴，已成为新一代 GIS 的重要研究课题，也

是新一代GIS的重要标志。时空GIS是一种四维（X、Y、Z、T）或（S、T）信息系统，其中（X、Y、Z）或（S）表示空间系统，（T）表示时间，这是一种具有时空复合分析功能和多维信息视觉化的系统，甚至还具有人工智能的功能。

根据目前的计算机技术发展水平来看，四维（4-D）时空数据库系统或数据处理技术研究，应解决的关键问题或难点是数据模型和数据结构，多维信息的复合、分析，查询和可视化等。

一、数据模型

过去十多年中有不少论著讨论时空一体化数据的数据结构和数据模型。这些模型在表达数据的时空一体化特征方面各有优劣，但都为研制较完善的时空一体化数据模型和数据库做出了有益的贡献。这些模型可大体归纳为三种：

1. 连续快照（time-slice snapshots）模型

常规的时空数据模型（如遥感图像数据的采集方式）即属连续快照模型。它是用一系列状态对应的地图来反映地理现象的时空演化过程。这种快照图像仅代表地理现象的状态，而缺乏对现象所包含的对象变化的明确表现，因此它不能确定地理现象所包含的对象之间在时间上的拓扑联系。在捕捉能力方面，由于没有时间维拓扑联系的存在，这一模型将无法实施按时间联系而确立的捕捉规则。此外，由于连续快照是对状态数据的完整存储，故数据冗余度极大。

2. 底图叠加（base map with overlay）模型

这一模型由于类似于地理底图的修订方式而得名。它的基本思路是，首先确定数据的初始状态，然后按照适宜的时间间隔记录发生变化的区域。通过叠加操作利用事件数据（变化区域）来建造状态数据，每一次叠加则表示状态的一次变化。

3. 时空合成（Space-time composite）模型

该模型是在底图叠加模型的基础上由Chrisman于1983年提出的，其设计思想是将每一次独立的叠加操作转换为一次性的合成叠加。这样，变化的累积即形成最小变化单元，由这些变化单元构成的图形文件和记录变化历史的属性文件联系在一起，则可较完整地表达数据的时空特征。这一数据模型一方面保持了底图叠加模型的优点，另一方面与当今GIS软件表达空间、属性数据的手段相似，因而便于以矢量为基础的GIS软件实现。另外，其时间变化特征存储在关系型数据库中，操作简便、高效。

在GIS中由于空间和非空间信息并存，使空间目标及其拓扑关系比较复杂，因此空间数据的体系结构极其重要。如何把空间和非空间数据溶合于一个系统中？比较广泛采用的一种方法是使用二个独立的数据库来分别处理空间目标及其属性数据。该法的主要优点是可以利用现有的DBMS。另一种较直观的解决办法是把时间看作附加的属性数据。传统的数据库可用来记录时间信息，即将目标的时间属性和其它特定的指标结合起来，以检索感兴趣的目标。一些土地信息系统常采用这种办法。图3—14和图3—15所示为二维和三维目标随着时间的变化，描述不同时期的地块状态以历史记录的方式存贮在系统中。

由图可见，在两个数据库中对目标进行综合时，难免会出现重复存储。当数据更新时，需要对目标重新描述，并且为了保持目标的历史链，旧目标也不能删除，这将使存储空间迅速增长。此外，由于目标的时间与空间关系彼此分离，两者间无内在联系，致使要求知道空间、时间及其它属性的应用时，不得不建立更多的附加关系。这又将使主存容量的需

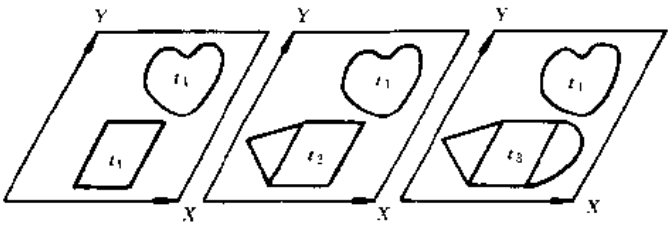


图 3—14　三个不同时间的二维目标表示

求大大增加，并且在人机对话环境中操作麻烦。

通过上面的论述可见，造成实施一个空间和时间信息系统困难的主要原因是对存储量需求的增加和操作的复杂性。无疑，存储时间信息会增加一定的数据存储量，这几乎是不可避免的。随着计算机容量的扩大这一问题已日渐不突出了。然而，研制一种良好的数据组织对改善存贮和操作性能将会有重大作用。

任何目标都可能包括空间和时间数据，在设计四维时空 GIS 系统时，中心问题是要把时间纳入空间系统中综合考虑。为此，首先应该熟悉空间与时间的一些基本属性，即其通用性、不可分割性、连续性、一致性、相对性和可量测性。在时空 GIS 中的目标具有不可分割的空间和时间数据，对一个目标来说时间与空间数据同等重要。

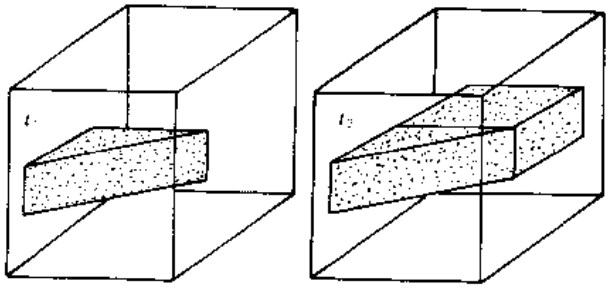


图 3—15　不同时间的三维目标表示

对于一个三维空间目标，在同一坐标系统下的四维数据模型可由一个四维坐标系统（X、Y、Z、T）唯一确定，即每一个三维坐标数组（X、Y、Z）必定有，而且仅有一个时间 T 值与之相对应。反之则不然。在时空数据模型中表示时间的方法，基本都采用如同空间数据一样来处理四维数据问题，并假定三维空间数据包含在具有连续的第四维时间的四维直角坐标系统中，且时间坐标垂直于三个空间维。时间可以前后变化，但不能分割。尽管在数据模型中各维数据以相同的方法处理，但这并不意味着在内插时间数据时采用如同内插空间数据相同的算法，也就是说，可以用不同的方法来分别处理空间和时间数据。4-D 数据集可以表示成（X、Y）和时间轴 T 的系统（图 3—16），也可以表示成如 3-D（X、Y、Z）和时间轴的系统（图 3—17）。在某些情况下，对一些用户来说可能前一种表示方式（图 3—16）更容易提取信息，并且数据模型比较灵活。其它的好处还有，已有的重要处理技术，如布尔数据集操作，都能够方便地扩展，象数据询问等算法可被简化。在四维系统中不同维之间的比例关系，进行四维距离的量测等主要取决于算法技术，而不是数据模型。

由于层状数据模型具有规则的不断分割的能力，以及如下的优点（原因）常被选用为时空四维 GIS 中存储空间数据的模型：

（1）层状模型便于压缩栅格数据。在四维模型的数据输出时经常需进行数据压缩，即作一定程度的数据概括。这种数据压缩的副产品是减少了许多操作的执行时间。然而，应该指出，对于复杂的充满空间的图像信息，在层状模型中要如实地存储数据集，实际上将增加存储容量。

（2）层状模型的空间检索能力有利于实施智能数据存取，因而速度较快。

（3）层状模型有利于产生比例尺可变的空间数据处理技术。由于四维数据集往往非常

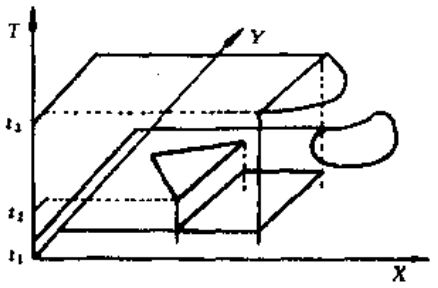


图 3—16　时空模型中三维目标表示

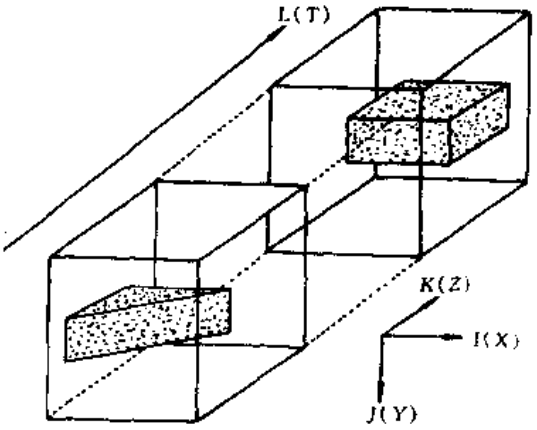


图 3—17　时空模型中四维目标表示

巨大，而客观现象的整体概念（如通过卫星遥感图像来分析地面的一些现象）往往是在较低分辨率的宏观条件获取的，但是通常又要求在高分辨率的条件下考察数据集的某一部分（如详细判读局部现象或结构特征）。层状模型结构能够在较低节点的周围存储较高的节点，所以低分辨信息也能够通过考察高等级节点来获取。

（4）层状模型结构便于对所有维的数据作相同的处理。层状模型的缺点是，为了最大程度地压缩数据集，要求所有维具有与二维平面坐标同样的压缩能力。一般二维平面有同样的坐标长度或度量，但时间和垂直坐标轴则往往与之不同，或者彼此也不同。在时间维中有时采用不规则取样（不等时间间隔取值），此时同规则取样相比，可能导致数据压缩丢失。

二、目标数据处理和数据结构

在资源环境等信息系统中，目标数据很多，也是用户很感兴趣的内容。这些目标通常是具有多种形式的空间和时间数据的目标。例如，矿体的空间位置、属性及其变化过程，环境科学中污染源的种类、程度及其动态变化等。在四维数据模型中计算机对数据目标的表示和处理方式大体上包括两方面的内容：首先：需要设计一种方案，使在某些特定应用环境下能够表示四维时空目标，此时，应处理好对目标描述的完整性、精确性和实施的综合性、有效性之间的关系；其次，是设计一种数据结构能方便、快速地存贮、查询和显示这种目标。

在表示具有不规则表面的现实目标时有多种不同的数据结构。目标的维数实际上可从0维变化到4维。对地矿应用来说，八叉树（Octree）表示方式是比较理想的一种。八叉树结构是从四叉树（Quadtree）结构直接发展来的。四叉树结构原本用来描述层状数据类型，其主要优点是可以对空间目标进行不断分解。这里应考虑的基本问题是，所表示数据目标的类型，分解过程应遵循什么原则和分辨率高低及是否需变化。八叉树结构的基本思想与四叉树相同，区别仅在于四叉树用来表示二维目标，而八叉树用来表示三维目标。还有一种扩展的八叉树称为十六叉树（Hexadecimaltree），则用来表示四维目标。

在区分数据结构时主要根据它们是确定目标的边界，还是组织目标的内涵。扩展八叉树是利用纯八叉树方式来表示一个实体的内部状况和原始边界的例子，这种结构同时顾及了存贮效率和表示目标的近似程度，是一种较好的描述和组织现实目标的方法。

四叉树有两种：指标四叉树和线性四叉树。前者是在子结点与父结点之间设立指标，因

此需要额外的存贮容量来储存非叶结点以及各结点的指标，难以实现数据压缩的目的。线性四叉树不需要记录中间结点，也不必使用指针，只需记录叶结点，并用地址码表示叶结点的位置。因此线性四叉树，以及类似的线性八叉树结构广泛应用于数据压缩和 GIS 中的数据结构，它们在节省存储量，提高存储和处理效率上具有明显的优越性。

1. 一种存储和表示四维目标的方案

目前国际上一些学者提出了一种存储和表示四维目标的方案（Mason D. C 等，1994）。对四维目标边界以一系列 2-D（X、Y）截面的方式进行存储。就某一时刻 t 来说，对每个 3-D（X、Y、Z）截面内的每一个 2-D（X、Y）截面，是存储表示每个 2-D 边界的一系列规则多边形的地址。这些规则多边形边界在直角坐标系统中的地址只在 X 或 Y 方向上有不同，而不能在两个方向上都不同。同理，相邻 2-D（X、Y）截面的地址只在 Z 方向上不同，而相邻 3-D（X、Y、Z）截面则只能在 t 方向上不同。如图 3—18 所示，这种表示方式实际上是一种四维线性二叉树结构。这种表示目标边界的方法无需内插两取样点间边界的细节，在数据结构上简化了节点与目标相交点的标识。尽管八叉树结构在表示目标边界方面有不少独到之处，但这种结构方案较容易操作，还允许将截面简化成一个点，从而便于采集和存储低于四维的目标。

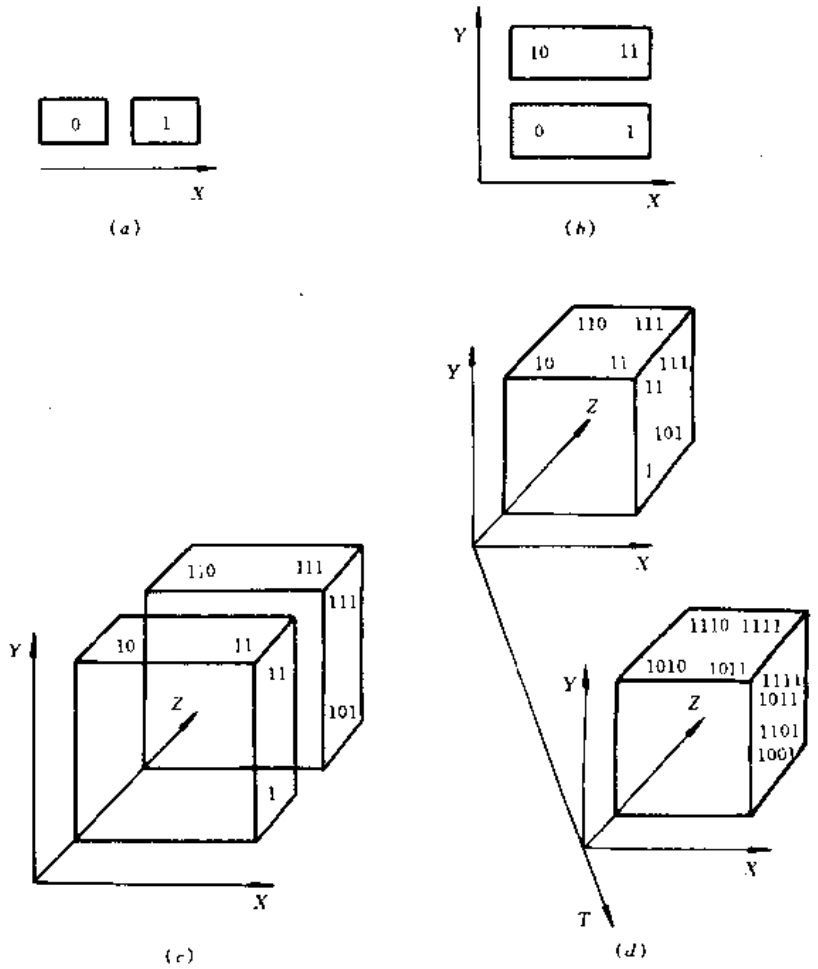


图 3—18　四维空间目标的线性二叉树表示

由 Marconi 研究中心研制的空间数据库管理系统，即 PUD（Picture Understanding Database）系统是建立在网状模型基础上的，可存储矢量数据。每种目标有统一的标识符，存储目标的属性包括目标类型，不同 t 值时 3-D 截面的数量，每一 t 值 2-D 截面的数量，目标的大小以及边界标识符。在构建原理上 PUD 系统能够存储由目标集组成的超目标，因此需要的话，复杂的四维目标还可以分割为两个以上的子目标进行处理。

2. 八叉树和十六叉树数据结构的表示方式（见第 2—4 节）

三、四维时空数据的组织和时空复合分析

1. 时空数据库的组织

在许多实际应用中四维时空数据库的数据量往往非常巨大，并且数据目标的空间和时间关系也相当复杂，同时这种数据库系统又必须具有反映和利用这些关系的能力。为了解决时空数据库容量与数据快速查询、利用之间的矛盾，在时空数据库的组织方式上提出了两种方案：

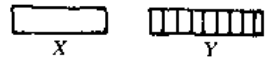
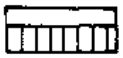
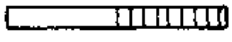
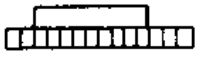
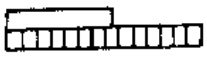
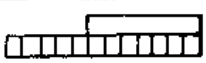
（1）双重库方式。即将时空数据库分割成历史库和当前库两部分。在当前库中存储地理实体最新的数据，而在历史库中保存地理实体的历史版本序列。

（2）空间数据库（Volume database）与目标数据库（Object database）独立方式。由于各种目标数据并非处处、时时都存在，因此也可将目标数据库与空间数据库彼此独立构建。

2. 空间与时间拓扑

客观对象在空间和时间上存在着内在的逻辑关系，即拓扑关系。拓扑关系是较比几何关系在更高层次上对地理实体空间关系特征的描述。它能够从下列三个方面提高空间信息的处理能力：减少空间搜索，简化算法；捕获数据错误，检验数据的空间结构是否遵循模型的拓扑一致性；具有最小的数据冗余，节约存储空间。时空一体化的数据结构必须具有拓扑特性，才能提高数据的质量控制与分析能力。一般来说，空间对象与时间的拓扑联系大体上如表 3—5 所示。

表 3—5　空间对象的拓扑时间联系

关　系	运算符	逆运算符	图　示
X 在 Y 之前	$<$	$>$	
X 等于 Y	$=$		
X 与 Y 相接	m	mi	
X 与 Y 重叠	O	Oi	
X 居于 Y 期间	d	di	
X 在 Y 的前段	S	Si	
X 在 Y 的末段	f	fi	

考虑到连续时态邻域在拓扑结构上的联结性，时态数据拓扑结构宜将时态按时序先后连接起来，以避免在时间和空间层次上的大量重复搜索。如同空间拓扑数据结构提供了在空间中实体间的导航方法一样，时态拓扑数据结构也能够提供在时间上从一个状态 或版本到其邻接状态或版本的导航方法。由于时态拓扑在逻辑上存在明显的两方向性，所以时态搜索可从正向（前向）和反向（后向）两个方向搜索，实现以下的查询：某一状态或版本在数据库中的前或后序状态或版本是什么？在某一段时间内或在某区域里什么内容发生了变化？如此等等。

在一些系统中代表四维子目标的结点是作为定位码来操作的，则很容易用特征码和特征 id 将它们联系起来，然后用同一特征码和 id 将所有邻接节点压缩成为一整体目标。Delaunay 三角形（双重 Voronoi 图示图形）可用来表示一组数点之间近似关系的数据结构。一种以四叉树为基础的 PMR 结构可用来表示线性特征。此外，分类生成和聚集的数据抽取技术，以及具有二进制关系的“部分指令集法”（poset）都获得使用。

3. 可视化（Visualization）与内插（interpolation）技术

在时间与空间数据库系统的数据分析、处理中，如同第一章所述的一般 GIS 系统那样，涉及到许多技术，而可视化与内插技术在资源环境信息的分析评价中往往特别有用。

可视化的基本含义是把一些数据转换成图形的形式，并以视觉的方式去研究这些图形及其变化，进而获取新的知识的一种科技手段。运用可视化技术可以弥补静态的，以文本为媒介的传输信息方式的不足，可以对数据进行实时的跟踪、引导和后处理，从而使观察者能更容易、更直观、更自然地解释、理解，并推理由各种方式获取的大量具有时态变化的地理对象数据。由于时空数据的多维性、复杂性，可视化操作的实现在很大程度上依赖于计算机科技（包括硬件、软件和算法）的发展。在现有条件下可视化操作主要通过以下四种途径：多窗口同步显示技术、幻灯片仿真技术、剖面显示技术、真三维造型技术，来实现诸如显示、缩放、移动、旋转、变换视角、剪切等功能。

内插操作需要解决的问题很多，例如获取同一地理对象在不同时间的空间位置；在某一时间点上，一组地理对象的空间位置；某一地理现象在某一空间位置的发生时间；在连续分布的地理对象中，利用某一范围中的属性时空信息推导某一特定位置处的属性变化时空规律。在时空信息的预报方面，可由历史信息预计地理对象的变化趋势。此外，时空数据的统计，剪裁和变换等也是时空信息的复合与分析模型中需要解决的问题。内插技术便于从一组离散的数据点生成概略景观，并可改善数据的可视化。当高维空间数据库属于稀疏数据集时，内插操作特别有用。它既能够加密、修匀稀疏数据，改善某些不准确性，又能够得到每一内插值的误差大小。

空间内插方法很多，但其共同的本质问题是根据已知的数据点情况选择恰当的内插模型，即首先使模型与已知数据点相对应，然后根据内插模型计算未知点的值。主要的内插方法有：梯森（thiessen）多边形法，趋势面分析法，样条函数法，移动平均法，克里金(kriging) 法等。在选择内插方法时，必须仔细考虑已知数据与内插法的固有假设条件是否一致，若一致才能获得最佳的内插结果。表 3—6 对一些内插法进行了总结、比较，以便于用户选用。

表3-6 内插方法比较表

方法	限定的随机的	局部整体	跃变渐变	精确内插否	局限性	最适合的场合	输出的数据结构	计算量	假设条件
EYE BALL	主观的限定的	整体	跃变	否	不可重复生产	外业数据航片判读	多边形	无	空间过程的直观理解边界内同性
边沿检测算子	限定的	整体	跃变	否	需确定和存储形状，更适合于人造地物	遥感获取的栅格图象	栅格	中等	边界内同性
邻近法	限定的	局部	跃变	是	每象元一数据点，不能估计误差，块模式取决于数据点的分布	由点模式产生的名义数据	多边形	少/中等	"最邻近"求得最佳信息
趋势面	随机的	整体	渐变	否	边沿影响，复杂多项式的意义不大，误差很少是空间非相关的	展现的特征很多，将这些特征转送到其他内插方法	栅格上的点	少/中等	趋势面的多重回归表象解释，独立高斯误差
傅立叶级数	随机的	整体	渐变	否	不适用于数据缺乏周期性	沙丘、波纹标记，或人造地物等周期特征	栅格中的点	中等	感兴趣现象的严格周期性
B一样条	限定的	局部	渐变	是	无误差估计，曲面不确定性	非常平滑的曲面	栅格中的点	高/中等	变化绝对平缓
移动平均法	限定的	局部	渐变	否/除非强制	结果取决于数据点的分布及窗口的大小；简单模型假设同一性；无误差估计	中等平滑数据的快速等值线绘制	栅格中的点	中等	连续可微分的曲面
最佳内插法(含kriging法)	随机的	局部	渐变	是	数据中非稳定性的实践与理论问题；制图的计算费用高	详细细节估计且需估计误差的情况	栅格中的点	很高	固有假设(齐性一阶差分)；局部平均值可用连续曲面表示

注：此表引自兰运超等编译的《地理信息系统原理》。

第四章　整体空间数据库

第 4—1 节　从传统数据模型到面向对象模型

数据库把相关的数据集合以综合的方法进行组织，使用户能有效地处理数据。数据的结构和表示以及其性质和特征相当复杂，需要有形式化的方法来描述数据的逻辑结构和各种操作，于是产生了数据模型的概念。数据模型可看作表示现实世界的规格说明。具体地说，数据库的数据结构、操作集合和完整性规则集合组成数据库的数据模型。

数据模型可分为三类：一是面向记录的传统数据模型；二是注重描述数据及其之间语义的语义数据模型；三是面向对象模型，它是在前两类数据模型基础上发展起来的面向对象的数据模型技术。

一、传统数据模型的弱点

传统数据模型主要指层次、网状和关系三种模型，它们是文件系统中数据模型的继承和发展。从私有到共享，从孤立的文件到彼此互相联系的数据模式，从简单的文件操作发展为较复杂的数据库操作和相应的数据语言，这无疑是一个质的飞跃。传统数据模型不但在数据库发展的历史过程中起了重要作用，而且现在仍然是商用数据库的主要数据模型。但由于历史和技术条件的限制，传统数据模型带有很深的面向计算机或面向实现的烙印。在非事务处理的应用中，例如 CAD、CAM、办公自动化、基于知识的系统等，人们越来越感到它们的局限性和不足。其主要弱点可概括为：

1. 以记录为基础，不能很好地面向用户和应用

传统数据模型基本结构是记录。而人类对现实世界的认识往往以某个事、物或概念为单位。这些可以互相区别的事、物或概念统称为实体。实体与记录不一定相对应。记录的组成往往从实现的角度考虑，而不一定反映人们对现实世界的认识。

记录的属性以及每个属性的域都是在记录的类型中定义的，也即对于给定类型的记录，要求其水平方向(有哪些属性)和垂直方向(每个属性的域)都是均一的。这在有些情况下基本上是可行的，例如学生的花名册。但是在有些情况下是不适用的，例如服装、家具的记录就很难满足水平方向均一的要求。因为，服装、家具的属性不一定都一样。在有些情况下，垂直方向均一性的要求也难以满足，例如对于“车主”这样的属性，属性值可能是部门，也可能是个人。总之，记录未能以较灵活的方式来描述纷繁的现实世界，而是从实现的方便出发做了简单化的规定。因而有时会使用户有削足适履之感，甚至削足也不能适履。

2. 不能以自然的方式表示实体之间的联系

除了实体的描述之外，对实体间联系的描述也是数据模型的一个重要方面。层次和网状数据模型虽然提供了显式的联系描述手段，但这些联系形式不是实体间联系的自然表示，而是联系的物理实现的描述。把本来应该对用户隐藏的物理实现的细节当作数据模型的组成部分呈现在用户面前，既不自然，又不便于用户理解和使用，而且也有损于数据的物理独立性。关系数据模型避免了这个缺点，实体间的联系是用关系来表示的。但关系数据模

型很难从数据模型看出实体间的联系。一个实体与哪些实体有联系，必须检查每个关系才能知道。现实世界中的实体联系被掩盖在记录、属性之中。总之，三种传统数据模型都难以自然地表示实体间的联系，是它们的一个主要弱点。

3. 语义贫乏

三种传统数据模型都或多或少含有一些语义，例如关系数据模型的主键，层次数据模型中的父子关系，网状数据模型中的系都提供了一定的语义信息。但是就总体而言，它们的语义是贫乏的。记录各个属性间的关系及其作用不够明确。这就要求用户通过文字来描述这些记录的语义，但这些文字描述又不是数据库的一部分。

此外，它们对实体间联系的语义描述，也缺少必要的手段。“交互关系”、“从属关系”、“性质”、“组成”等都是代表语义不同的联系，却用同样的记录结构表示。并且缺少描述这些联系的相应结构，在语义上无法区分这些联系。

4. 数据类型太少，难以满足应用需要

传统数据模型原来都是面向事物处理的。它们一般只提供最常用的一些简单数据类型，例如整数、实数、字符串等，而且不允许用户自己定义新的数据类型。在有些应用中，希望属性值不直接给出，而是由公式、规则或过程推导而得。有时不但要知道数据的当前值，还要知道不同时间的数值。这些功能在传统数据模型中都不支持。

由于传统数据模型的缺陷，从70年代后期开始，陆续出现了各种其它的数据模型，并且统称为后关系数据模型。它们主要是语义数据模型和面向对象数据模型。

二、语义数据模型

语义数据模型发展的基本目的是克服传统数据模型的缺陷，提供不受具体的实现结构的限制，更多地面向用户的模型。它能提供一种“自然”的机制来说明数据库的设计，还能更准确地表示数据及数据间的关系。

语义数据模型提供了强有力的抽象机制，如概括和聚合。概括允许设计者将相似对象形成组，集中到一个更普及的组对象上，如从旅社、宾馆和招待所中可得到组对象饭店。聚合允许设计者从对象的性质或属性中模型化出抽象实体，抽象实体中可包括异质部件，如一个聚合对象地址由街道、城市、州和邮编组成。除概括和聚合外，还有联合、分类等。

除抽象构造外，语义数据模型的另一重要特点是支持“派生数据”的概念。派生数据是指并不实际存于库中，而是当需要时可由库中数据及其之间的关系派生出来的数据。

语义数据模型主要包括E-R模型、RM/T模型、TAXIS模型、SDM模型、函数模型、SAM＊模型、事件模型以及SHM＋模型等。下面只介绍RM/T模型和函数模型。

RM/T是对关系模型进行扩充用以描述应用中的语义信息的语义模型。RM/T认为数据库由实体集组成，实体集用来描述现实世界的对象和概念及其相互关系，每个实体拥有一个关联特性集，可通过操作集来管理实体。数据库中每个实体定义成至少具有一个实体类型的一个实例，给出类型的所有实体共享该类型的所有特性，RM/T支持了类型和超类型两种关系。在RM/T中各实体都有系统定义的唯一标识号。

函数模型中将实体集和属性之间的关系抽象为一个实体到属性域的映射函数，实体与实体之间的关系也用函数表示。实体间及实体与属性间的关系，既可以是单值函数也可以是多值函数。函数模型具有易于理解的可视性表示和先进的内部语义完整性。

三、面向对象的数据模型

面向对象（object-oriented，oo）的概念起源于程序设计语言。对象是客观世界实体的抽象描述，由信息（数据）和对数据的操作组合而成。类是对多个相似对象共同特性的描述。消息是对象之间通信的手段，用来指示对象的操作。方法是对象接收到消息后应采取的动作序列的描述。实例是由一特定类描述的具体对象。对象具有封装性，对外部只提供一个抽象接口而隐藏具体实现细节。类具有继承的特性。研究人员在面向对象程序设计中试图加进数据持久性的概念，以支持数据库。另外，语义数据模型中概括、聚合和派生的概念被面向对象模型吸收。当然，传统的数据库管理中持久性、二级存储管理、并发、数据恢复和查询语言的概念，也必须包括在面向对象数据模型中，这样，就形成新的面向对象的数据模型，如图 4—1 所示：

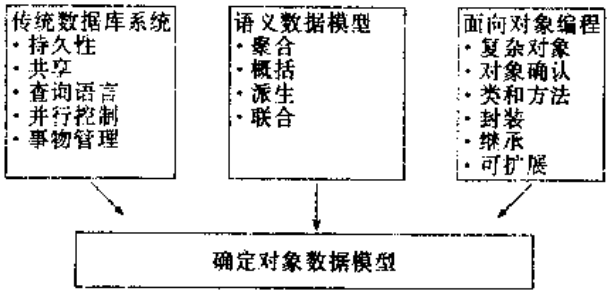


图 4—1　面向对象数据模型

面向对象数据模型提供了表示复杂对象的能力。可在任意层次上嵌套各种类型构造符（数组、表、元组等），同时可表示数据之间各种特殊关系。数据之间特殊关系的引入主要是通过元类的概念得到的。元类是相似的类的共同属性的抽象，元类的实例是类，类的实例是对象。面向对象的数据模型除数据与操作的封装外，还涉及到分类（classification）、概括（generalization）、聚集（aggregation）、联合（association）、继承（inheritance）和传播（propagation）等概念。

由于面向对象的数据模型是将现实世界的实体都抽象成对象，即将模拟实体状态的数据和操作封装成相对独立的模块，然后进行分类和概括，以人们认识问题的自然方式将所有的对象构建成一个分层结构，来描述问题领域中各实体之间的相互关系和相互作用，从而建立起一个比较完整的结构模型，使得对现实世界的构成与人们认识问题的方式直接对应。因此，面向对象的系统分析与设计方法和面向对象的数据模型，为目前 GIS 所面临的问题提供了较好的解决途径，成为 GIS 尽快进入决策应用阶段的关键技术之一，并受到计算机技术界的重视。

第 4—2 节　面向对象技术的原理与方法

一、对象、消息和封装

（一）对象和消息

1. 对象

世界上的事物无论如何千变万化，都离不开一个词，那就是“对象”。可以说，我们身边的一切事物都是对象，小至一粒米、一根针、一个结点、一条弧段，大到一个人、一块

面状地物、一所学校，都是对象。从哲学概念上讲，它们都是客观对象。我们大脑中的概念和认识也是一些对象，是主观对象。

如此看来，一个对象既可以极其简单，又可以相对复杂。但它们都具有相似性。例如一粒米，它首先是一粒米；其次它具有颜色；再次它可以食用，它可以用作工业原料等等。又如一个人，首先他有名字，如张小东；有一些身体素质情况，如性别、身高、体重等；还有一些有益于社会的技能，如精通无线电技术等。再如一所学校，它有名字，如“中国矿业大学”，还有师资、设备、在校学生数等情况；还有教学楼、图书馆、居民区、运动场、道路网、通讯网等。

从上面的分析看来，一个对象无非就是这样一个实体，它具有一个名字标识，并有自身的状态和自身的功能。世界上所有事物就是如此简单，这恰恰是面向对象技术所追求的目标——将世界上的问题求解尽可能地简化。

现在，对象已经形成，但还不能生存。就象人的生存离不开社会一样，对象的生存也离不开与其它对象的交互。对象又是如何交互的呢？

2. 消息

仍以人这一对象来进行分析。张小东是一个无线电工程师，他能为其他人修理无线电器。但他要吃饭、穿衣、娱乐等等，这就要请求他人帮助解决这些问题。这里的“请求”便是一个人与其他人进行交互的手段，并且是面向对象技术中的专业术语——消息。此外，日常生活中不仅有请求，而且还有命令。比如，上级对下级的命令，军官对士兵的命令等等，这些“命令”也是一种消息。因此，消息就是对象之间相互请求或相互协作的途径。

3. 对象的划分与确定

面向对象程序设计的一个重要方面是确定对象以及对象之间相互传递的消息。程序员任何时候说明的消息都可发送给对象，对象的选择依赖于设置对象的目的和所需进行的操作。例如，在某个 GIS 中，有一条高速公路，路上的加油站、立交桥、交通灯等作为点状对象，它的中心线作为线状对象，它的路面作为面状对象。

在面向对象程序的设计中，对象的选择是最关键的一步。对象选择得好，便于程序扩充，还可以为其它应用提供基础。如何选择对象，没有固定的方法。但有一个基本原则，即寻求共性——所有具有共性的系统成份便可以确定为一种对象。

4. 对象的状态

在面向对象系统中，一个对象的状态是通过域来描述的，也称为“私有存储单元”。对象的一个重要特征表现在，它的私有存储单元只能由它自己的操作进行处理，消息是外界能够引用对象操作及获取对象状态的唯一方式。这个特征保证了对象的实现只能依赖于它本身的状态和所能接受的消息，而不依赖其它对象的内部细节（包括状态）。

5. 公有消息与私有消息

公有和私有消息的含义是：一批消息都属于同一个对象，但有些是可由其它对象向它发送的，叫做公有消息；另外一些则是由它自己向本身发送的，叫做私有消息。

如果一种复杂对象是由一批简单对象组成，这种复杂对象中实现了许多消息，一批是直接针对这一复杂对象的，可由外界对象发送，这些就是公有消息；还有一批消息是针对该对象所包含的更简单对象的，这些消息是不对外开放的，外界也不必了解它，这些消息则称为私有消息。

（二）协议与封装

对象和消息对自然界的实际事物进行了良好的模拟，也简化了人们对世界事物的理解。但是，现实事物仅靠这两个概念还不能充分表达。例如，一个人有各种各样的能力。有一些能力，他乐意向外人宣告；有些能力只向一部分人通告；还可能有些能力，不想让任何人知道。另外，即使外人知道了他的某些能力，但他也不向外界提供这些服务。这是经常出现的情况，上面的私有消息便是这样的例子。这里就有一个协议的问题。

协议是一个对象对外服务的说明，它告知一个对象可以为外界做什么。外界对象能够并且只能向该对象发送协议中所提供的消息，请求该对象服务。因此，它是由一个对象能够接受并且愿意接受的所有消息构成的对外接口。也就是说，请求对象进行操作的唯一途径就是通过协议中提供的消息来进行。即使一个对象可以完成某一功能，但它没有将该功能放入协议中去，外界对象依然不能请求它完成这一功能。从私有消息和公有消息上看，协议是一个对象所能接受的所有公有消息的集合。

对象、消息和协议是面向对象设计中的支柱性概念，这些概念一起又引入了一个新的概念——封装。封装就是将某件事物包围起来，使外界不必知道其实际内容。封装的意义如图 4—2 所示。

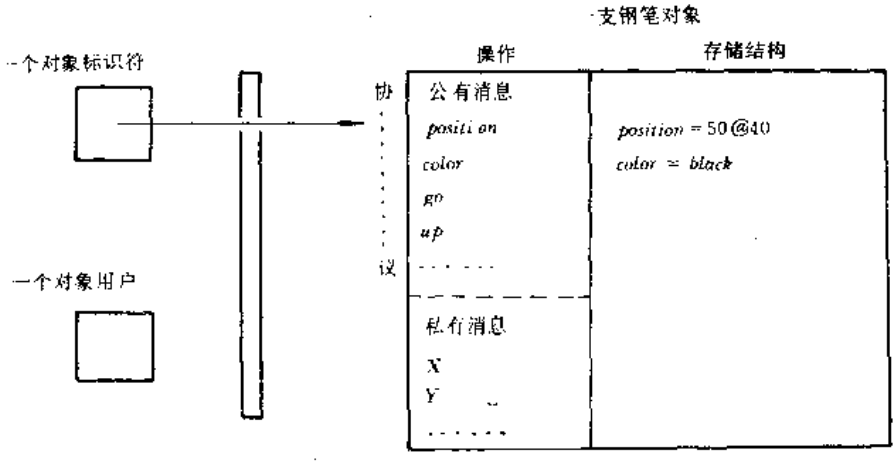


图 4—2　封装的意义

从图中可以清楚地看到，钢笔对象的用户只能通过对象标识符访问到这个对象提供给外界对象的协议，并且只能通过协议上提供的操作，向该对象发送消息请求它工作。

二、类和实例

（一）基本概念

类是对一组客观对象的抽象，它将该组对象所具有的共同特征集中起来，以说明该组对象的能力和性质。在系统构成上，类则形成了一个具有特定功能的模块和一种代码共享的手段。

类和实例这两个概念，在日常生活中随时可能用到。比如，你拿到一个苹果，说："这是一个苹果。"翻译成专业术语，就是："这是一个苹果类的实例。"而你见到一个饭店，说："这是一个饭店。"其实就是说："这是建筑物类的实例，该建筑物作饭店用"。再进一步抽象，可看到"矿大内招"是饭店的一个实例，"矿大外招"是饭店的一个实例。从这个角度

来说，饭店成了类，而“矿大内招”和“矿大外招”则为实例。由此可见，类和实例之间的关系是抽象和具体的关系。实例是类的具体事物，类是多个实例的综合抽象。

类的确定方法是归纳。这就要求在观察问题时，要注意提炼所遇到对象的共同特征，以确定其类。例如，居民住宅、饭店、商店、工厂可分别是一个类，它们也可以抽象为建筑物类。又如，建筑物是一个类，湖是一个类，公园是一个类，它们也可以抽象为面状地物类。

（二）类的描述

在实际工作中，单纯描述一个独立的对象如何工作是没有什么意义的，而描述一个类的工作则是有效的。因为类是一种对于一类事物的概括工具，是一种提高工作效率的工具。

在图形处理过程中，一般需要定义一个钢笔类。要描述一个钢笔类需要提供以下信息：

1）数据说明　所有的钢笔都具有存放墨水的胶囊，都有笔头，都有起笔位置等，也就是结构相同。不同的是，各钢笔的胶囊中所盛的墨水颜色不同，笔头的粗细不一，起始位置各异，这些就是每个钢笔对象的私有存储内容。因此，要为钢笔说明一个存储区，用于存放每支钢笔的不同特性，即在类中要定义保存颜色的域、表示粗细程度和保存起始位置的域。

2）操作定义　所有的钢笔都用同样的方法响应消息。比如，可以响应向不同方向移动的消息，可以响应画面状地物的消息等，这些是共享信息。

（三）对象创建——实例化

一个类一般来讲不做实际工作，它只不过是构造实例的模板，真正工作的是对象，也就是类的实例。在设计过程的初期，首先要产生一个类的描述，再使用这个描述来构造它的实例。如在 Smalltalk 和 c^{++} 这样的面向对象程序设计语言中，对象都是由定义好的类通过实例化操作而生成的。例如

aPen←Pen new.

就是 Smalltalk 语言中创建一个对象的格式。

这个语句的意义是，创建一个钢笔的实例，用 aPen 作为它的标识。也就是说，aPen 指向一个存储区，其中可以存放这支钢笔的私有特性，如颜色、笔头粗细和初始位置等。

（四）实例的初始化及其删除

对象的初始化与实例化是两种不同的概念。实例化一般只给出一个对象的存储结构，实例化过程可以包含初始化过程。而初始化才具体给对象的存储结构填充初始内容。在 c^{++} 中，初始化是定义构造函数。构造函数是实际上初始化一个对象新实例的消息，不论这些实例是根据源代码中定义新实例的语句创建的，还是通过使用 New 关键字动态创建的。这一过程叫做对象的实例化。根据一个类创建一个新对象，就实例化了这个类的一个对象。构造函数的名字与类本身的名字是相同的。c^{++} 还要处理对象删除（即去构）的问题。去构函数只做名字所指出的工作，即清除一个类的特定实例。

三、继承及类之间的层次关系

（一）继承

1. 继承的引入

继承是一种现实世界中对象之间独特的关系，它使得某类对象可以继承另外一类对象的特征和能力。从一般概念上看，继承有“接过来”、“照搬使用”的意思。例如饭店是建

筑物类的子类，它的一些操作，如显示和删除等，以及一些属性如房主、地址、建筑日期是所有建筑物公有的，仅在建筑物类中定义它们，然后遗传给饭店等子类。当人们知道招待所是饭店类的子类的时候，通过继承概念可知道招待所也具有建筑物类的操作和属性。继承性能使许多事情简单化，因为继承性使许多概念相互依赖，记住了一个主要概念，其它相关概念的主要特征也就知道了。继承性具有双重作用，即一是减少代码冗余，二是通过协调性来减少相互之间的接口和界面。

2. 继承的分类

继承概念具有多样性。从狭义上说，比如一条可航行的河流，既可以继承自然交通线的特征，又可以继承河流的特征，这是多重继承的概念。如是一条不可航行的河流，只能继承河流的特征，这就是单重继承。一个人可以全部继承父辈的财产，也可以继承其部分财产。儿子只能继承父亲的部分性格特征，而不是全部，这就是全部继承与部分继承的概念。

从广义的继承概念来讲，如一徒弟从其师傅那里学到师傅所有的技术，则在任何需要师傅的地方都可以由该徒弟来取代，这就是取代继承。

“水系”是一类对象，“湖”是一种特殊的水系。“湖”继承了“水系”的所有特征，并且任何一个湖都是一个水系，这便是包含继承，即湖包含了水系具有的所有特征。

“鸵鸟”是一类特殊的鸟，它们不能继承鸟会飞的特征，这就是受限继承。

“工程师”是一类特殊的“人”，他们比一般人有更多的特有信息，这是特化继承。所有这些继承具有两类不同的性质。多重继承和单重继承是一类，是用于描述继承源的；而另一类，如取代继承、包含继承等则是描述继承内容的。

从系统论的观点看，多重继承和单重继承用于控制系统的组织策略；而另一类继承则是用于控制系统组成单元的内部要素。

3. 继承与封装

有人认为继承和封装是两个矛盾的机制。因继承机制的引入削弱了对象概念的封装性。其实，这两个概念没有实质性的冲突。在面向对象系统中，对象依然是封装得很好的实体，其它对象与之通讯的途径只有一条，即发送消息。类机制是一种静态机制，它丝毫不影响对象的封装性。从另一个角度来看，它们还有一些相似性，即它们都是一种共享代码的手段。比如，继承是一种静态共享代码的手段，通过子类创建的实例可以接受某一消息启动由该类的超类所定义的代码段，从而子类与超类共享了这一代码段。而封装基础上的消息，则是提供了一种动态的代码手段。通过封装，将一段代码定义在一个类中，由另一个类中定义的方法可以通过创建该类的实例，并向它发送消息而启动这一段代码，同样达到共享的目的。

4. 继承与委托

面向对象系统中每个元素本身都是一个对象。为了减少每一个操作的实现代码，要尽量使用向其它对象发送消息的办法。这就是委托的意义。在日常生活中很常见，经理要求班长完成一系列任务，班长又分别让班里的工人处理各种事务，他本身或许并不做实际工作，只做监督和验收工作，而他的工作主要是通过“委托”来完成。

对象之间的关系，一方面有继承——静态定义的类可以继承其它类的能力；另一方面还有委托——一个对象可以依赖其它对象为它完成某些操作。

（二）类的层次

引入类的继承，就出现了类的层次结构，也就有了超类（基类）、子类（派生类）的概念。以一个城市 GIS 为例，类的层次如图 4—3 所示。

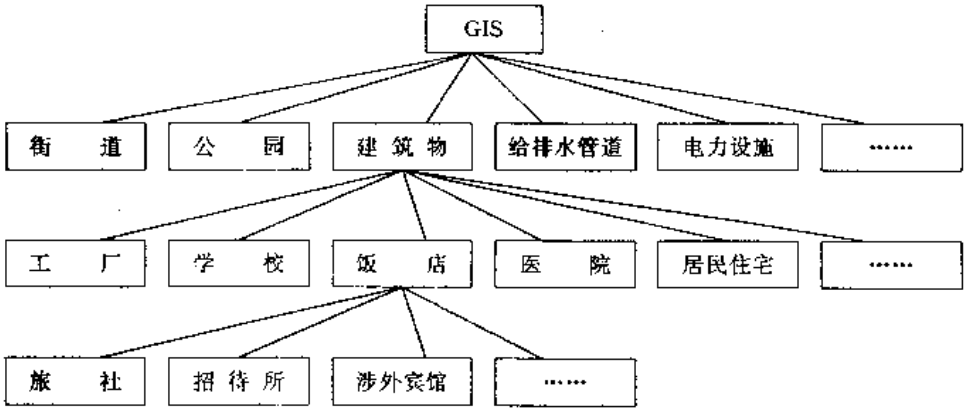


图 4—3　类的层次关系

在这个层次结构中，GIS 系统包括街道、公园、建筑物等多个超类，而建筑物是工厂、学校、饭店等类的超类。超类和子类的概念是相对的。类的层次关系是用于描述和组织世界事物的一种简单、有效的工具。

（三）继承机制在面向对象系统中的应用

如前所述，类的一大作用是可重用性，继承机制进一步增强了类可重用的灵活性，它不但使一个类可以通过委托来实现共享，而且可以使用继承来共享已有的事物。继承的意义不仅在于代码的共享和重用，还在于它带来了结构的共享和重用。

1. 域（或称为结构）的继承

一个对象的域由类中的结构说明来定义。有了继承机制，要定义一个类的子类，就自然定义了超类中定义的数据结构，也就是说，子类继承了超类的域。另外，子类除了拥有其超类所具有的数据结构之外，还可以定义自己新的数据结构。

如前所述，继承与封装没有冲突，那是从功能角度上看的。现在再从结构的角度上看封装与继承。在 c^{++} 面向对象系统中，类的使用方法有二种：一种是实例化使用，即建立一个类的实例，并通过向这个实例发送消息来完成一定的任务；另一种是继承使用，即建立一个类的子类，形成一种更具体的一类对象的模板。

在这种使用观点下，对象的域对于该对象的使用者来说是不可见的（封装），而我们又允许子类不受限制地使用超类中定义的对象域，这就会破坏该类的封装。但是，如果完全不让子类使用超类中的域，只能通过超类提供的方法来访问超类中的域，又可能失去有效性。为了保持一致性和有效性，在 c^{++} 中提供了一种折中的解决办法，即在类的定义中提供了公有（public）、私有（private）和保护（protected）三种定义域（数据结构）和方法，同时保证了方法使用上的一致性和有效性。其中，公有域和方法是指这个类的任何用户（包括实例用户和继承用户）都可以直接使用；私有域和方法是指只能由类本身自己使用，用户不能访问；保护域和方法是指只能继承使用，不能实例使用。

2. 方法的继承及抽象类

继承使得定义新的子类自然具有其超类的一切功能。对于子类的功能定义，仅仅继承超类中的方法是远远不够的。比如，同样一个方法可以很好地处理超类的数据结构，但由于子类中增加了新的数据结构，它的处理功能就不够了。这就需要子类重新定义这个方法，并且要保证方法名相同，否则超类的方法仍为子类所用，也就是说，子类定义的方法要覆盖超类中的相应方法。

有了继承概念，才使得类的定义更加灵活和实用，抽象类的定义只有有了继承机制才能发挥作用。顾名思义，抽象类不是一种可生成具体对象的类，也就是不能产生实例的类。抽象类的作用在于将许多有关的类组织在一起，提供一个公有的根基，让其它一些具体的类从这个根基上产生，从而可以进行正常的工作。抽象类刻划了公有行为的特征，并通过继承机制传送给它的子类。这不但能够简化子类的定义，还能够利用方法的继承和覆盖来处理重载及多态问题。

（四）功能重载和多态性

谈到类和继承，还得涉及另一个相关的机制，即功能重载和多态性。子类对超类方法的覆盖功能就是功能重载和多态性的体现。面向对象系统中消息传递的实现是以重载和多态为支撑概念的，也可以说，重载和多态概念是消息传递模式的自然体现。功能重载实际上意味着特定的函数不仅以名字来区分，而且用它所带的参数来区别。c++用了以下规则来解决功能重载问题：唯一确定函数的是它的名字及其所接受的参数，并要注意它的类型和顺序。假设有一个函数，需从一个数据库中取一个记录，叫做 fetchFromData—Base（ ），并且要求这个函数既能够返回数据到一个形参结构中，又能返回一个及时创建的有新标记的结构中，这就需要定义下面两个重载函数：

- fetchFromDatabase（Key）
- fetchFromDatabase（Key，Receiving Structure Address）

编译器根据传给这个函数的参数指出要使用哪个函数。这就意味着不必仅为要接受不同的参数，而给这个函数起个不同的名字。

编译器的第二个保证是：消息名与它们所属的对象是紧密结合在一起的。例如在打一份建筑物明细表时，按传统的方式，一个函数名对应一种类型的待打印对象，就要考虑“是 printhouse（ ），还是 house Print（ ）”，而在面向对象程序设计中，只要简单地实现消息 print（ ），并为它定义每个要打印的对象即可。

多态的意思是同一个消息可以根据发送消息对象的不同采用多种不同的行为方式。例如，在一个图形系统中，有建筑物、公路、控制点等几类对象，在 object 类中定义一个 draw（ ）方法，只定义其协议，而在建筑物、公路、控制点类中分别再次定义 draw（ ）方法，这样，在 c++中就可能做这个聚合工作。

```
for (i=0, i<intmax, i++)
    s [i]. draw;
```

在 *Smalltalk* 中为

```
aBuildingdraw.
aHighwaydraw.
aControllpointdraw.
```

四、概括、聚集、联合与传播

1. 概括

概括是把一组具有相同特征和操作的对象归纳在一个更一般的超类中。例如，对象类“住宅”是一种“建筑物”，“住宅”是“建筑物”的子类，而“建筑物”是“住宅”的超类。

由于概括的层次性，所以某一层次的子类对于更特殊的子类来说，它又是超类。在“建筑物”/“住宅”中概括能延伸到“乡村住宅”，这样，“住宅”是“建筑物”的子类，同时“住宅”又是“乡村住宅”的超类。

注意，这里的超类和子类是对一个对象的抽象化，而不是指两个不同的对象。淮海路30号楼是“住宅”类的一个实例，同时也是它的超类“建筑物”的一个实例。

2. 联合

联合是相似对象的抽象组合，可看作是更高层次的集合对象。术语“集合”是用来描述联合的，有联合关系的对象叫成员。因此联合指的就是“成员”关系。例如北京市是东城区、西城区、海淀区等区的联合。

在联合中，一个成员对象的具体细节被忽略了。强调的是整个集体对象的特征。一个集合对象的实例可以分解成一系列其成员对象的实例。联合通过其对象（成员）产生集合数据结构，一个集合的操作是由该集合每个成员的操作组成的。

3. 聚集

聚集是类似于联合的抽象化概念，它模拟了所谓的合成对象，合成对象是由一些其它对象组成的。术语“合成对象”是用来描述更高层次对象的，而“部分”是指合成对象的组成部分，部分和合成对象的关系是“部分”关系；反之，合成对象和部分的关系是“组成”关系。例如“建筑物”是“墙”、“窗”、“门”、“房顶”和“楼板”几部分的聚集体。

考虑聚集时，不强调整个合成对象的具体细节。每一个聚集对象的实例都可以分解成其成员对象的实例，每部分都保持其自有功能。聚集的操作与其它部分的操作是不兼容的。聚集的每一个操作是由其每个部分产生的不同操作所组成，其数目是固定的。

4. 传播

传播作为联合和聚集的一种工具，通过一种强制性的手段将子对象的属性信息传播给复杂对象。就是说，复杂对象的某些属性值不单独存于数据库中，而是从它的子对象中提取或派生。例如一个县的人口可由关联的乡镇人口求和派生得到。一家旅社的客房总数和职工总数可以从客房数据库和职员数据库中派生得到。这一概念保证了数据库的一致性。

为对上述诸多名词、概念有一个较系统的理解，现结合图4—4的例子作一个小结。分类是将类似的对象归于一类，如玩具部、鞋部等可归于购物中心。概括是进一步将种类分成子类或超类。购物中心对于各个售货部来说是超类，而对于建筑物来说又是子类。因此分类是相对的。在概括这一体系中，子类中的对象继承了上一层次的属性，并且借助一套操作程序进行管理。联合是对体系中类似对象之间相互关系的描述，并可定义一套数据结构。聚集有些类似于联合，但聚集是由下往上地将几个不同特征的对象集合成一个更高水平的合成对象。例如许多售货部通过聚集形成购物中心。合成对象里的操作程序不同于各个子对象，但合成对象可以从它们那里派生得到一些信息。子对象的属性也可以传播给合成对象。

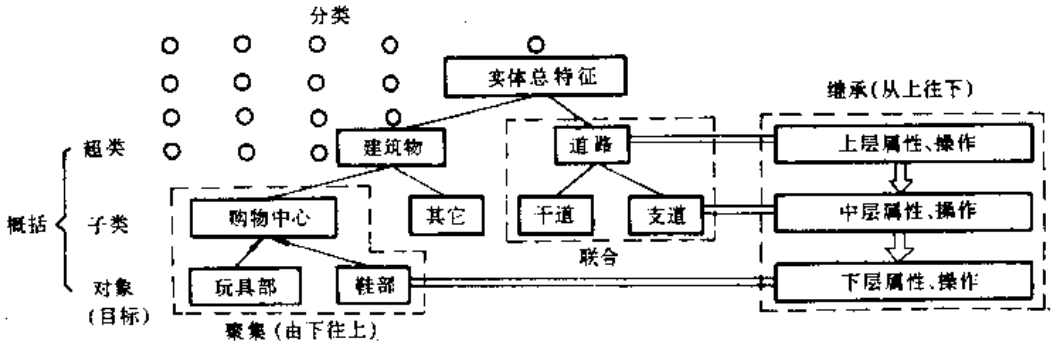


图 4—4　面向对象技术方法构造原理举例

第 4—3 节　面向对象数据库的分析与设计

数据库系统开发是一项软件工程，按照 1993 年 6 月公开发行的《国家经济信息系统设计与应用标准化规范》之五——《数据库开发工程实施细则》，数据库系统开发过程划分为规划论证、需求分析、数据库设计、数据库实施和数据库运行与维护五个阶段。

下面概略地介绍各个阶段的主要任务。

1）规划论证阶段　本阶段是属于开发数据库系统的立项阶段。其主要任务是对特定的信息系统做概括性的调查了解，将调查所收集到的情况，进行分析归纳、合并，对其中与数据及信息管理有关的，再进行深入分析，并根据归纳分析的结果确定该特定信息系统数据库的总目标，形成总体的结构，然后进行可行性分析论证。

2）需求分析阶段　详细调查用户需求；分析数据的现状和使用；登记数据资源词典；确定数据库的应用目标、结构及用户界面；写出需求分析和数据库设计说明书。

3）数据库设计阶段　进行概念模式设计、逻辑模式设计和物理数据库的设计；写出数据库结构设计报告。

4）数据库系统实施阶段　主要任务是数据库创建；数据库模拟；数据库加载；数据库部分服务及应用程序设计、调试。

5）数据库系统运行与维护阶段　建立数据库管理员的责任制度；培训系统用户；支持系统正常运行；系统功能维护和新项目开发。

由于数据库设计是系统开发的中心环节，本节主要介绍数据库设计。

一、概念模式设计

概念设计是指采用面向对象数据模型表示系统需求分析阶段所得到的信息及信息之间的联系。概念设计一般分为三步：一是设计局部概念模式；二是将局部概念模式综合成全局模式；三是组织评审。

（一）设计局部视图

局部视图设计过程如图 4—5 所示。

1．确定局部结构范围

范围划分的方式有两种。一种是依据系统的当前用户进行自然划分，若各用户对信息内容和处理的要求明显不同，则应对它们分别设计局部视图。第二种是各用户对信息内容及处理的要求没有明显区别，此时可把数据库提供的服务归成几类，然后为每类应用设计

一个局部视图。地理信息数据库一般采用第二种方式，可以从几何图形数据、影像数据、地形数据和属性数据等几类设计局部视图。

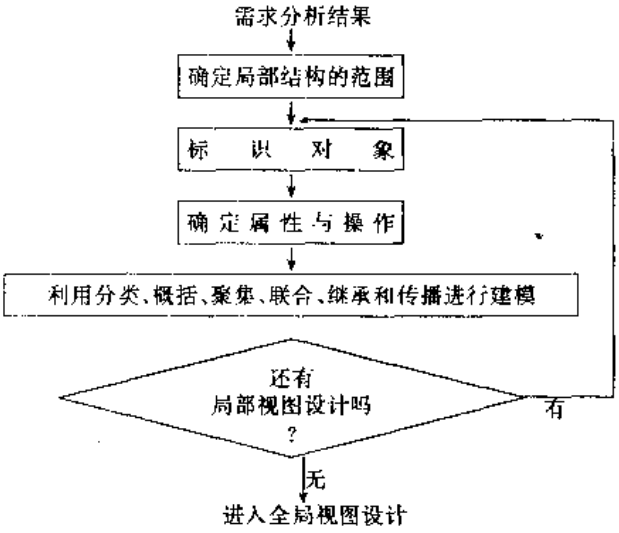


图 4-5 局部视图设计过程

2. SIS 中面向对象的几何数据模型

以 GIS 为代表的空间信息系统（Spatial Information System，SIS）作为计算机、摄影测量与遥感、地理学、自动制图等多种科技发展的综合产物，已得到迅速发展。

考察 GIS 中的各种地物，GIS 分为复杂地物和简单地物。就其几何性质而言不外乎四种类型的地物，即点状地物、线状地物、面状地物以及由它们混合组成的复杂地物。因而这四种类型可以作为 GIS 中各种地物类型的超类。从几何位置方面抽象：点状地物为点（M_1，M_2，Z）；线状地物由弧段组成，弧段由点组成；面状地物由弧段和面域的二维行程编码（2DRE）表组成。因此，弧段聚集成线性地物或面状地物，简单地物组合成复杂地物，结点的坐标通过标识号传播给线状地物和面状地物，进而还可以传播给复杂地物。2DRE 表中的 Morton 码通过指针传播给面状地物。

综上所述，几何数据的面向对象模型如图 4—6 所示。

3. SIS 中面向对象的属性数据模型

GIS 中的地物可根据国家分类标准或实际情况划分类型，如一个大学的 GIS 对象可分为建筑物、道路、水系、植被、电力线等几大类。地物类型的每一大类又可以进一步分类，如建筑物可再分成住宅、工厂、医院、子弟学校、教学楼、图书馆等子类。住宅类又可分为职工住宅和学生住宅；职工住宅进而分为双职工住宅和单职工住宅。教学楼可分为教学主楼、教 1 楼、教 2 楼、…等等。

将几种类型中某些具有公共特征的属性和操作抽象出来，可形成一种更一般的超类。例如工厂、商店、招待所属于产业，它们有收入和税收等属性，可把它们组成一个更高水平的类型——产业类。又如教授、讲师、助教属于教师类；教师、职员、工人属于教职工类；研究生、大学生、专科生属于学生类；教职工、学生属于人员类。

在地理属性数据建模中，经常要用到聚集的概念。例如学生住宅可能具有房间总数、床

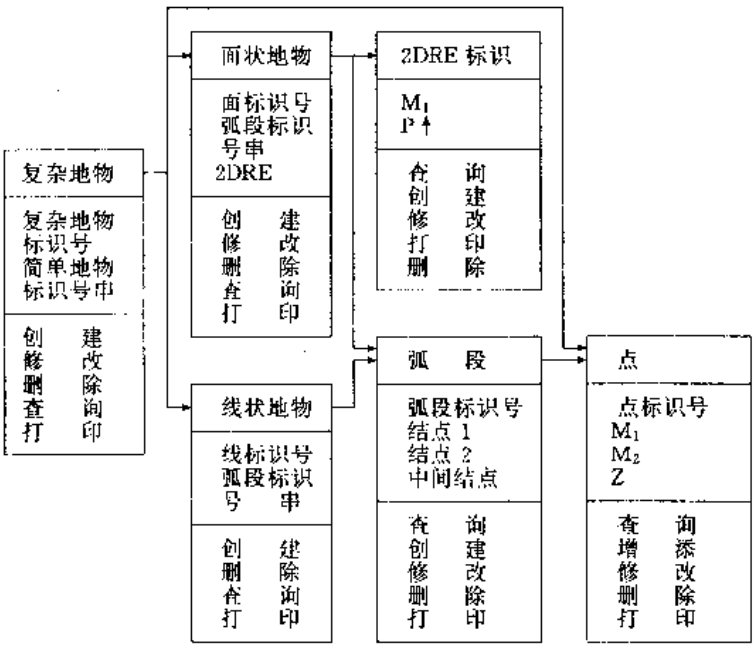


图 4--6　面向对象的几何数据模型

位总数和住入人数等属性，这些属性必须从学生库和学生住宅设施库中获得。又如教学大楼由房顶、楼板、墙、门、窗等组成。图书馆由情报室、流通室、期刊室、采编室组成。

联合也可用来建模，如学校的“建筑物”和门牌为“西 27 栋”的“建筑物”是通过“内部”关系联合的。又如“某宅某套房的房主”和“门牌号”二个属性可组成一个新的对象。

图 4—7 表示了 GIS 地物的分类及其关系。

4. SIS 中面向对象的 DEM 模型和 RS 模型

GIS、DEM 和 RS 的结合已成为这三个领域发展的必然趋势，而建立能兼容图形、DEM 和 RS 数据的整体空间数据也是 SIS 或 GIS 的核心问题。

数字高程模型（DEM）中的各种地形特征可以看成是简单地物，因而可用前面所述的点状地物、线状地物和面状地物描述。DEM 的等高线可用弧段来表示。格网 DEM 用双向行程编码表示。

RS 数据可用 2DRE 的紧凑格式表示。

（二）设计全局视图

设计全局视图的过程如图 4—8 所示。

1）确定公共对象类　公共对象类的确定并非一目了然。特别是当系统较大时，有很多局部视图，而且是由不同的设计人员确定的。因而对同一现实世界的对象可能给予不同的描述，有的作为对象类，有的作为属性。即便都表示对象类，其名称与标识号也不可能都相同。在这一步仅根据对象类名和标识号来认定公共对象类。

2）局部视图合并　合并的原则可以是：首先进行两两合并；先合并那些现实世界中有联系的局部结构；合并从公共对象开始，最后再把独立的局部结构加进去。

3）消除冲突　消除冲突是本设计中最重要、最复杂的部分。“同名异义”与“异名同义”是冲突的两个类型。

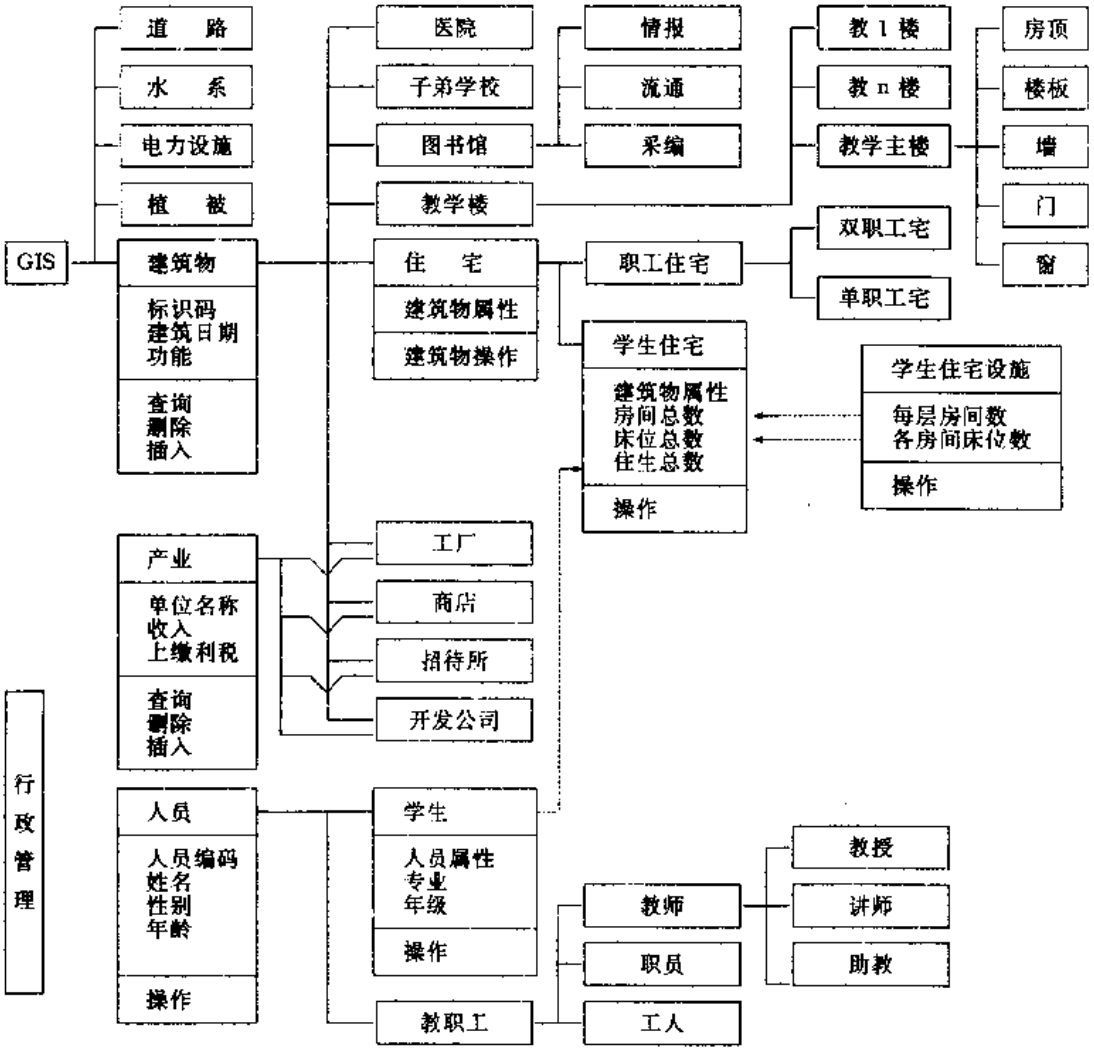


图 4—7 各种对象的分类及其关系

按照上述过程，面向对象的几何数据模型与面向对象的属性数据模型相联系，这种联系可以通过对象的标识号来实现，标识号的第一位数字表示几何类型，例如“1”表示点状地物，“3”表示线状地物，“4”表示面状地物，“5”表示复杂地物。其余位数表示属性标识号，如某公路的标识号为 30027。

DEM 模型的各种地形特征和等高线数据可以和 GIS 中点状地物、线状地物、面状地物合并。GIS 中的 2DRE 的链指针与 DEM 中的双向行程编码和 RS 中的 2DRE 的紧凑格式可以聚集成一个粗网格索引对象。因此，面向对象的整体空间数据模型如图 4—9 所示。（为了清晰起见，图中对象框内略去属性和操作）

二、逻辑模式设计

逻辑模式设计的目的是把概念设计的结果转换成由特定 DBMS 支持的数据库模式。所以，逻辑设计是直接依赖于数据模型和 DBMS 来选择的。目前，由于面向对象数据库系统不太完善，建立数据库系统有下列三种方法：

第一，在现有的面向对象系统或语言上加上数据库管理功能以丰富系统的库管理能力，从而建立一个面向对象的数据库；反之，也可以在现有的数据库管理系统上引入面向对象

的数据模型，使其支持面向对象的设计方法，而形成一个面向对象的数据库系统。

第二，在现有的面向对象系统和数据库系统之间建立交互联系，使两者之间互相补充，互相支持。

第三，另起炉灶。根据面向对象概念和数据库原理的要求，重新设计并实现一个面向对象的数据库系统。

现用第一种方法。几何数据用面向对象语言加数据管理功能方法，属性数据用 DBMS 加面向对象数据模型方法。

1. 几何数据的类描述

由于GIS功能和应用范围还处于发展之中，目前还无法对一些几何对象类型定义一个完整的操作集。例如，对面状目标的几何类型，除需要的最基本的操作，如创建、删除、修改、查询、显示目标外，还要涉及包含、相邻、相

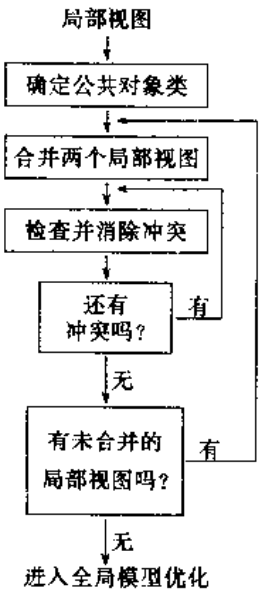


图 4—8 全局视图设计过程

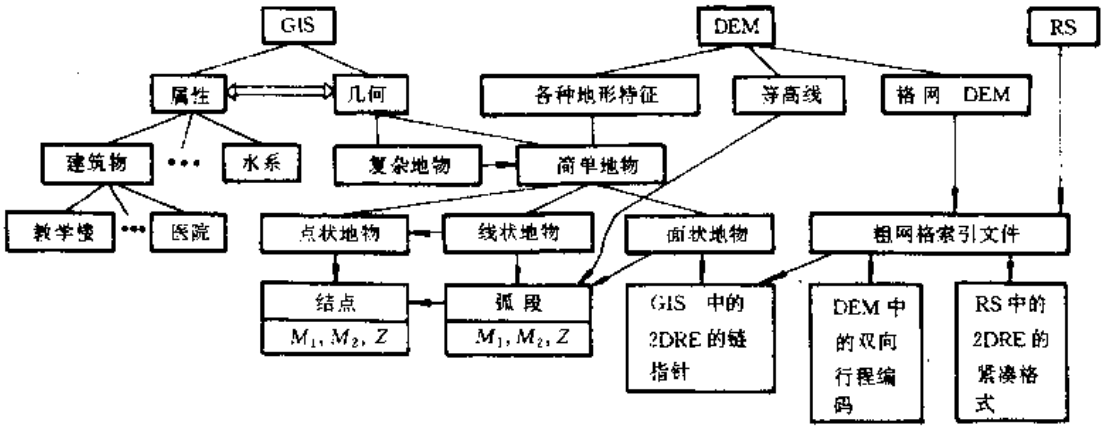


图 4—9 面向对象的整体空间数据模型

交等拓扑分析操作以及各种空间集合逻辑运算等几十种操作，而且随着GIS应用范围的扩大，GIS的功能操作还将不断增加。因此，可将几何类型最基本的操作（称基本操作）封装在类型中，而将大量的分析操作（称为附加操作）放在类型之外，并通过数据接口，将私有数据传递出来。如点基本类的数据类型及基本操作定义如下：

```
class Npoint
    { int cpointid;
      int M1, M2;
      float Z;
    public:
      select ( );
      modify ( );
      delete ( );
      print ( ); }
```

有关类型的数据变量、基本操作集、附加操作集的关系如图 4—10 所示。

```
class 类名                          附加操作集定义：
{ 数据类型说明；                    main ( )  /* 主函数 */
  public：                            { 公共数据变量说明；
  基本操作集定义：}                     调用基本操作；
                                        调用附加操作；}
```

图 4—10 几何对象类的描述

2. 属性数据模型向关系模式转换

由于是在现有关系数据库系统 oracle 上引入 c^{++}，实现属性数据库，故属性数据模型中各类的属性域必须转换成关系模式，各类操作用 c^{++} 实现。其转换规则一般为对每个类转换成一个关系模式，类的属性即为关系模式的属性，对象的标识号是关系模式的键。例如：

建筑物（建筑物标识码，建筑日期，功能）。

学生住宅（建筑物标识码，房间总数，床位总数，住生总数）。

学生（人员编码，专业，年级）。

人员（人员编码，姓名，性别，年龄）。

为相应类转换后的关系模式。

三、物理数据库设计

一般来说，在关系数据库的物理设计阶段，设计人员需要考虑以下内容：

1. 选择数据库主文件结构

所谓选择主文件的结构，是指选择合适的主存取路径，包括文件内部记录的物理顺序安排和聚类索引的选择等。oracle 系统允许在不同的关系模式之间建立聚类索引，它可按照相同数据项值把关系记录存放在相连的物理存贮区中，按此索引访问数据可达到较高的效率。

2. 选择索引

选择索引的目的在于实施关键字的约束和提高访问效率。通常关键字约束是通过指定在关键字属性或属性组合上建立唯一索引来实现，在其它属性或属性组合上建立的索引称为辅索引，用以提高访问效率。因此应根据需求分析，对用户在经常性应用中要求访问的数据项建立索引或组合索引。

3. 计算空间占有量

可按 RDBMS 所提供的计算公式计算空间占有量。计算时要考虑当前数据量、增长率和索引选择情况，并为未来的扩展留有充分余地。

4. 分配存贮设备

设计人员必须根据对应用的分析，合理地分配存贮设备，以避免对某一设备访问过于频繁的不均衡现象。

四、面向对象的属性数据库实现方法

把关系与 oo 技术统一起来是今后 RDBS 技术最重要的基础。为了实现 oracle 与 c^{++} 相结合的 OODBS 的功能，应在以下几个方面把它们统一起来。

1. 统一结构

c++要与oracle数据管理层接口，亦即通过SQL语言与RDB连接，或者与数据存储管理层接口，即通过层的过程调用与RDB连接，把对象映射到关系的元组上，并生成对象标识符(OID)，把它们作为元组的一个属性传给RDB，并把在一个对象中找到的OID映射到存储在RDB的相应对象上，从而实现对对象的操作映射为对元组的操作。

2. 扩充RDB的数据模型

第一个扩充允许关系的一个列的值是用户定义的关系，而不仅仅限于系统定义的数据类型。这就意味着用户可以把关系的一个列的域定义为用户定义的关系。

第二个扩充封装，即把数据与对数据进行操作的程序统一起来。允许用户对一个关系加上过程进行操作，或对元组的列值进行操作的过程。这样就封装了元组的状态与行为，状态是列值的集合，行为是列值进行操作的过程集合。用户可以写任意过程，并将它加到关系上对任意一个或多个元组进行操作。

第三个扩充继承，允许用户把DB中所有关系组成继承层次。这样，对关系A与B来说，A可以作为B的父关系，B可以继承A中定义的所有列与过程。允许一个关系有多个父关系，可以继承它们的列与过程。

3. 扩充查询和数据操纵

扩充RDB的数据模型是不够的，还必须扩充查询和数据操纵，才能在扩充的数据模型下增加查询和数据操纵类型。

4. 提供对象导航

c++与oracle结合的OODBS，应提供自动管理内存中大量对象的工作空间管理机制。工作空间管理机制使DB应用程序能通过内存指针追踪，在内存中对象间进行导航，把各个对象的变化情况以集合为单位传播到DB中。

第4—4节　人工智能与专家系统

人工智能简称为AI。近20多年来AI获得了迅速的发展，应用领域广泛。什么是人工智能呢？一般认为用计算机模拟人的智能行为就属于人工智能的范畴。实际上，人工智能方法是指人工智能研究的一些核心内容，包括搜索技术、推理技术、知识表示、机器学习与人工智能语言等方面。

目前，人工智能研究的主要内容包括：专家系统、知识库系统、决策支持系统、自然语言理解、智能机器人、模式识别、自动程序设计。专家系统简称ES，是目前人工智能应用中最成熟的一个领域。

一、知识表示

(一) 知识表示的基本概念

知识表示是研究用什么形式将有关问题的知识存入计算机，以便进行处理。计算机所处理的知识，按其作用可大致分为三类：

1) 描述性知识　表示对象及概念的特征及其相互关系的知识，以及问题求解状况的知识，也称为事实性知识。

2) 判断性知识　表示与领域有关的问题求解知识如推理规则等，也称为启发性知识。

3) 过程性知识　表示问题求解的控制策略，即如何应用判断性知识进行推理的知识。

例如，从矢量数据到栅格数据的转换，是采用平行线扫描法，还是采用线性四叉树法的问题。有关的知识如下：

描述性知识：矢量数据、栅格数据、平行线扫描、线性四叉树、时间、存贮量。

过程性知识：采用平行线扫描法、采用线性四叉树法。

判断性知识：采用线性四叉树法较快、存贮量大，采用平行线扫描法较慢、存贮量较小。

按照作用的层次，知识还可以分成以下两类：

1）对象级知识　直接描述有关领域对象的知识，或称为领域相关的知识。

2）元级知识　描述对象级知识的知识，如关于领域知识的内容、特征、应用范围、可信程度的知识以及如何运用这些知识的知识。

目前，常用的知识表示方法有逻辑、框架、语义网络、产生式系统等，下面分别予以介绍。

（二）逻辑表示法

这里所说的逻辑是指一阶谓词逻辑。一阶谓词逻辑可以很直观、自然地表示“事实”，即描述客观事物状态、属性值及事物之间的关系等。例：“北京是中国的首都”，“日本位于中国的东方”，“淮海路与中山路的交叉口”等等事实，可以描述为下列命题：

位于（东方，日本，中国）；

首都（北京，中国）；

交叉口（淮海路，中山路）。

对于一些复杂事实，如“北京的春天不是刮风沙就是不下雨”，“天下雨后会湿地皮”，“不是所有的商品房都是没人居住”等，需要采用连接词“与（∧）”、“或（∨）”、“非（¬）”、“蕴含（→）”和“存在量词∃”、“全称量词∀”来组合，可表示为

地点（北京）∧时间（春天）∧〔刮风沙（天）∨¬下雨（天）〕；

下雨（天）→湿（地皮）；

¬∀ x〔商品房（x）∧¬居住（人，x）〕；

或∃ x〔商品房（x）∧居住（人，x）〕。

（三）语义网络

语义网络是一种采用有向图的知识表达方法。结点表示客体、客体性质、概念、事件和行为等；有向边则表示客体间的联系。图 4—11 用语义网络表示了建筑物的一部分知识。从图中可看出，语义网络类似于面向对象数据模型，因而可以采用分类、概括、聚集和联合四个抽象概念来表示知识。

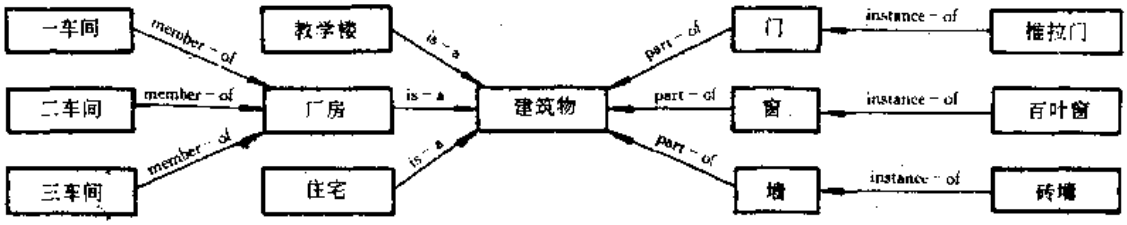


图 4－11　建筑物部分知识的语义网络

（四）产生式系统

产生式系统的基本要素包括：综合数据库、产生式规则和控制系统。综合数据库是产生式系统所使用的主要数据结构，用来描述问题的状态。控制系统根据综合数据库新的状态选择一条规则作用于综合数据库，形成一个“识别—动作”循环，直至综合数据库的状态满足了结束条件或无可用规则为止。

产生式规则的作用是对综合数据库进行操作，使综合数据库发生变化。产生式规则的一般形式是

if〈前提〉then〈动作〉(或〈结论〉)。

例如在城市规划中一些知识为：“沼泽地不许盖房”，“飞机场附近的地不能盖高楼”，“消防站附近不许建医院”等，用产生式规则表示为

IF a zone is a marshland THEN prohibit construction,

IF a parcel is close to an airport THEN limit building height,

IF a parcel is near to a firestation THEN prohibit hospital construction.

（五）框架

一个框架可看作是一个“聚集语义网络结构”，其中的语义网络聚集结点对应框架名，而聚集元素结点为框架的块。例如，图 4—12 所示的语义网络聚集对应图 4—13 所示的框架结构。

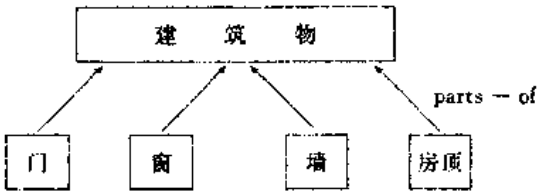


图 4—12 语义网络聚集

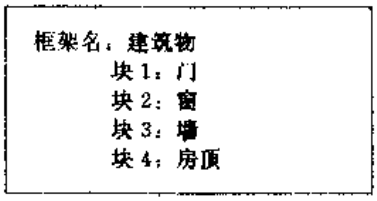


图 4—13 建筑物框架

框架的一般结构如图 4—14 所示。块内的值可以是数字、字符串，也可以是框架名、块名或过程名等。

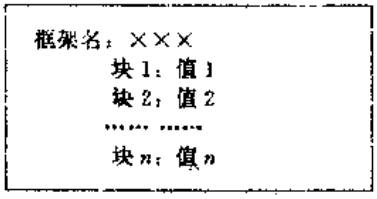


图 4—14 框架的一般结构

二、知识推理

知识推理就是运用知识求解问题。知识推理与知识表示有密切的关系，根据知识表示方式的特点，可将知识推理分为二类：“图搜索”法和“逻辑论证”法。基于图的知识表示，知识推理采用图搜索法。基本的图搜索法有：广度优先搜索法、深度优先搜索法等。当知识表示采用逻辑方法时，知识推理采用逻辑论证法。常用的逻辑推理方式可分为五类：演绎推理、归纳推理、外展推理、非单调推理和不精确推理。下面简要介绍前三种逻辑推理。

1. 演绎推理（Deduction）

演绎推理的基本形式为

$$(P;\ P\rightarrow R)\rightarrow R$$

其意思是如果 P 为真并且规则 $P\rightarrow R$ 也为真，那么通过演绎推理导出 R 为真。演绎推理的

另一种推理形式为

$$(P\rightarrow Q;\ Q\rightarrow R)\Rightarrow(P\rightarrow R)$$

其意思为如果规则 $P\rightarrow Q$ 和 $Q\rightarrow R$ 都为真，那么可导出 $P\rightarrow R$ 为真。这种推理又称为推理传递性。

2. 外展推理（Abduction）

外展推理的基本形式为

$$(R;\ P\rightarrow R)\Rightarrow P$$

其意思为如果 R 为真并且规则 $P\rightarrow R$ 为真，那么推导出 P 为真。

3. 归纳推理（Induction，类比推理）

归纳推理的基本形式为

$$(P;\ R)\Rightarrow(P\rightarrow R)$$

其意思为如果 P 和 R 都为真，那么导出规则 $P\rightarrow R$ 为真。

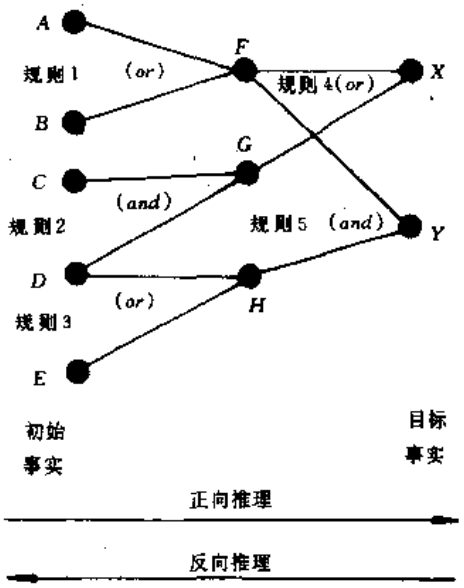


图 4—15 推理示意图

．演绎推理又称为正向推理，外展推理称为反向推理。为了对这两种推理过程有一个初步理解，现结合图 4—15 所示的例子作一个简单的推理。从图中知：有初始事实 A、B、C、D、E 和目标事实 X、Y，五条规则为

规则 1：$A\ or\ B\rightarrow F$

规则 2：$C\ and\ D\rightarrow G$

规则 3：$D\ or\ E\rightarrow H$

规则 4：$F\ or\ G\rightarrow X$

规则 5：$F\ and\ H\rightarrow Y$

在正向推理情况下，假设 B 和 D 为真，A 和 C 为假，E 为未知；首先运用规则 1 导出 F 为真，然后运用规则 4 导出 X 为真；再运用规则 3 导出 H 为真，因为 F 为真，所以运用规则 5 导出 Y 也为真。

在反向推理的情况下，假设 X 为真，Y 为假，可能导出：F 必为真，H 必为假，G 无关紧要，在这种状态下，可能导出：A 为真，D 和 E 为假，B 和 C 无关紧要。

在实际应用中，有时运用双向推理，即混合使用正向推理和反向推理。

三、知识获取

所谓知识获取是将解决问题的专门知识从知识源传递和转换成程序的过程。知识获取的方法和途径可分为以下三类：

1. 人工移植

所谓“人工移植”，是依靠人工智能系统的设计师、知识工程师、程序编制人员、专家或用户，通过系统设计、编制程序及人机交互或辅助工具，将人的知识移植到知识库中去。

2. 机器学习

所谓“机器学习”，是人工智能系统在运行过程中，机器通过学习，获取知识，进行知识积累，对知识库进行增删、修改、扩充与更新。

3. 机器感知

所谓“机器感知”，是通过机器视觉、听觉、触觉等途径，直接感知外部世界，输入自然信息，获取感性和理性知识。

四、专家系统

专家系统是为了解决具有适当规模的专家级问题，采用一定的策略，具备知识库，具有学习机制，可以对知识库进行修改，以增进解题能力的计算机程序系统。它根据用户提供的数据、事实和信息，动用系统存储的专家知识，进行推理判断，最后得出结论，同时给出结论的可信度，以供用户决策。或者简而言之，专家系统是同专家具有相等解决问题能力的智能程序系统。

专家系统的基本结构是围绕知识库和推理机组织的，粗略地说是：知识＋推理＝专家系统。当然，一个实用的专家系统还需要知识获取，解释功能，人机交互界面等功能模块。图 4—16 是专家系统的基本结构示意图。

图中知识库用于存放系统求解所需要的知识(即以事实和规则表述的领域知识及元知识)。推理机负责使用知识库中的知识去解决实际问题。解释系统负责解释系统本身的推理结果，回答用户的问题。知识获取系统负责管理知识库的知识，包括根据需要修改、删除或添加知识及由此引起的一切必要的改动，维持知识库的一致性、完整性等方面。人机接口将专家和用户的输入信息翻译成系统可以接受的内部形式，同时把系统向专家或用户的输出信息转换为人类易于理解的形式。中间数据库是专家系统在执行与推理过程中用以存放中间结果或论据等的工作存贮器。

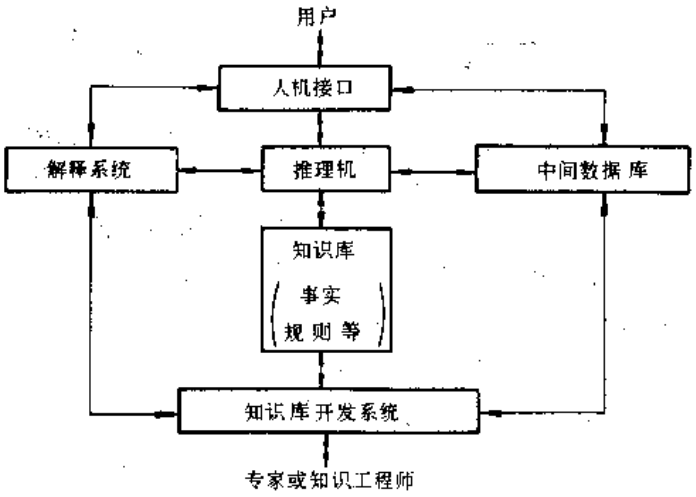


图 4—16　专家系统的基本结构

第 4—5 节　空间知识的表示

空间知识不仅十分丰富，而且相当复杂。按知识的属性，可分为静态的、可靠的、确定性的知识和变化的、模糊的，不确定的知识。按其推理过程中知识所起的作用，又可分为说明性知识、与问题有关的求解知识、描述控制策略的过程性知识和元知识。常用的知识表示方法有五种：谓词逻辑、规则、框架、语义网络和过程。这些方法各有所长，又各有特点，且各有一定的适用范围。但如何充分、有效地表示空间的复杂知识还有待于作进一步的探讨。

一、谓词逻辑用于空间知识表示

(一) *形式化方面*

关系数据库可以认为是采用了一阶谓词的形式，每一个关系模式对应着一个谓词。例如 GIS、DEM 和 RS 三位一体的整体空间数据库可以按表 4—1 所示的关系模式组织。每个

关系可以看成一个谓词，因此表中的信息用一阶谓词逻辑可表示为 SISDB（1，418，260，105），SISDB（2，2406，520，320），等等。

上述这种对应关系，可以在空间信息领域得到直接应用。例如，在可视知识编码（Visual knowledge encoding）中，按下面三步识别地理物体（图 4—17 所示）：

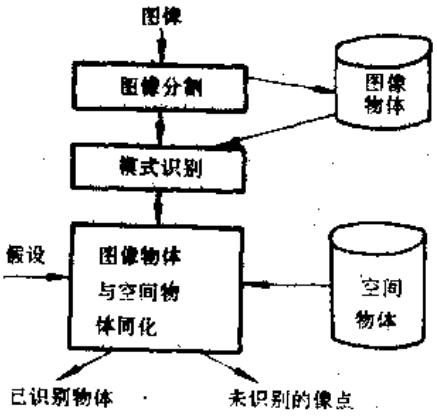


图 4—17　图像识别的主要步骤

表 4—1　SISDB 表

M_0	GIS	DEM	RS
0	0	0	0
1	418	260	105
2	2406	520	320
3	3942	780	482

（1）为了描述图像物体，利用结构和形态分析进行图像分割。

（2）利用模式识别技术对图像物体进行组合，然后把图像物体与库中的理想空间物体进行比较，通过比较使图像物体与空间物体同化。

（3）图像识别的结果是一些已识别的物体和一些未识别的像点。

这个过程可以用下面谓词逻辑来表示。首先从

IMAGE（*Pixel—ID*，*Red—level*，

Green—level，*Blue—level*）

关系开始，可以得到下面关系：

*R*1（*Object—ID*，*Locator*）

*R*2（*Object—ID*，*Shape*）

*R*3（*Object*1*—ID*，*Object*2*—ID*，*Relation—type*）

其中 *Relation—type* 表示物体 1 和物体 2 之间的联系类型。从空间意义讲，可能有如表 4—2 所示的一些联系。其中两个物体可能同类型，也可能不同类型。

表 4—2　物体间的联系类型

*Object*1	*Object*2	*Relation—type*（联系类型）	类型 1/类型 2
345	678	*Vicinity of*（相邻）	区/区
278	804	*Contained within*（包含）	区/区
301	345	*North of*（在北面）	城市/城市
451	901	*Outside*（在外面）	建筑物/道路
341	935	*More than* 1,000*m*（大于……）	海岸线/房屋
136	730	*Near*（邻近）	农场/农场
276	778	*Crossing*（交叉）	公路/公路

（二）非形式化方面

上面所讲的关系数据库就是一个谓词逻辑理论的模型，因此数据库中的事实与现实世

界的映射是直接的。但在下面几种情况下则不是很显然的，须用另一种逻辑模型表示。

1. 空间事实

例如："我住中山路 34 号" 和 "我站在东经 24°23′34″，北纬 56°14′48″，海拔 345m 的这点上"，可以表示为

$SITUATE\ (x, y, z) \wedge Person\ (x) \wedge street\ (y) \wedge Building\ (z)$

$Stand\ (x, y, z, w) \wedge Person\ (x) \wedge N\ (y) \wedge W\ (z) \wedge Elevation\ (w)$

2. 空间关系

如图 4—18 所示，独立空间物体之间有下列三种关系：

(1) 空间关系相对于观测者，例如物体 A 在后面，物体 B 在前面。用逻辑表示为

$Behind\ (A, X) \wedge Front\ (B, X) \wedge Person\ (X) \wedge Object1\ (A) \wedge Object2\ (B)$。

(2) 与坐标轴有关的空间关系，如物体 D 在物体 C 北面。用逻辑表示为

$North\ (D, C) \wedge Object\ (D) \wedge Object\ (C)$

(3) 物体间内部关系，如物体 F 在物体 E 和 G 之间，逻辑表示为

$Intrinsic\ (F, E, G) \wedge Object\ (F) \wedge Object\ (E) \wedge Object\ (G)$

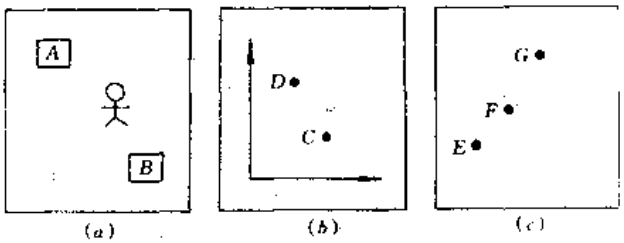


图 4—18　空间关系的类型

a—与观测者相联系；b—与坐标轴相联系；c—内部的联系

二、语义网络的应用

语义网络在专家系统中得到了广泛的应用。地质探矿专家系统 Prospector 的知识表示就是由三级网络构成，它们分别用来描述概念、陈述和推理规则。

最低一级网络是分类学网络，它们由四种弧表示：$N_1 \xleftarrow{s} N_2$ 表示 N_2 是 N_1 的子集；$N_1 \xleftarrow{e} N_2$ 表示 N_2 是 N_1 的元素；$N_1 \xleftarrow{ds} N_2$ 既表示 N_2 是 N_1 的子集，又说明 N_2 与 N_1 的其它用 ds 链接的子集是不相交的概念；$N_1 \xleftarrow{de} N_2$ 表示 N_2 是 N_1 的元素，且 N_2 与 N_1 的其它 de 型元素是不相同的。如图 4—19 所示。

陈述由分块语义网络表示。分块语义网络把整个网络划分成若干个块，每一块表示一个完整的陈述，每一个陈述都可以分解为若干个简单断言。例如，"角闪石部分地转化为黑云母"，可由图 4—20 表示，分解为 7 个简单断言的集合：

$a1$：存在一个实体 E_1

$a2$：E_1 的成分是角闪石

$a3$：存在一个实体 E_2

$a4$：E_2 的成分是黑云母

$a5$：存在一个过程 P

$a6$：在 P 过程中，E_1 转化为 E_2

$a7$：在 P 过程中，转化的程度是部分转化

判断性知识用规则表示，其规则的形式如下：

$$E \rightarrow H\ (LS, LN)$$

其中 LS、LN 为可信度，用来反映证据 E 对假说 H 的影响程度。当 $LS > 1$、$LS < 1$ 和

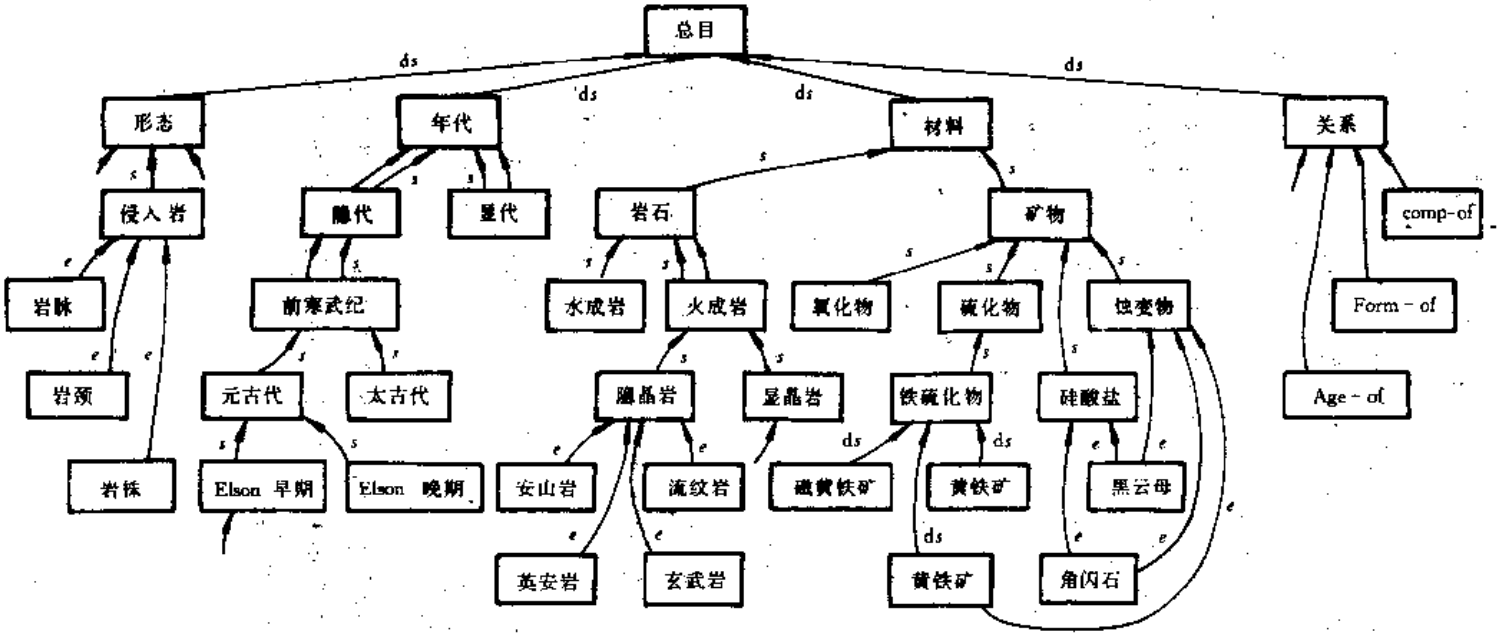


图 4—19 分类学网络（局部）

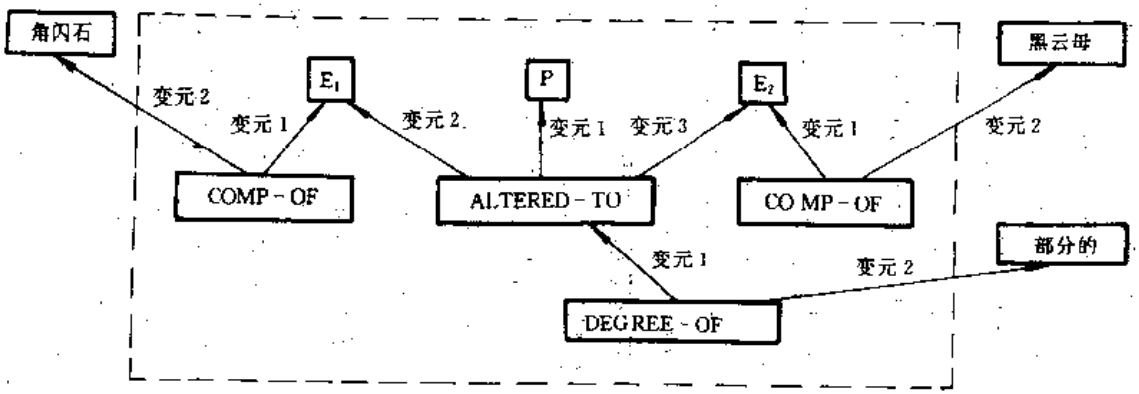


图 4—20 角闪石部分地转化为黑云母

$LS=1$ 分别表示证据 E 出现时，它支持、反对或不影响假说 H 成立。$LN>1$、$LN<1$ 和 $LN=1$ 分别表示 E 不出现时，它支持、反对或不影响 H 成立。（推理网络图略）

三、规则在拓扑关系中的应用

对于图 4—21 所示的空间实体，可用规则表示其拓扑关系。

1. 面块—弧段的拓扑关系

规则 1　如果 S 为面块且由弧段 A_1、A_2、A_3、A_4 组成，则 S 为面块 B_1。

规则 2　如果 S 为面块且由弧段 A_4、A_5、A_6 组成，则 S 为面块 B_2。

2. 弧段—结点的拓扑关系

规则 3　如果 S 是起止结点为 N_1 和 N_2 的弧段，则 S 为弧段 A_1。

规则 4　如果 S 是起止结点为 N_1 和 N_3 的弧段，则 S 为弧段 A_2。

规则 5　如果 S 是起止结点为 N_3 和 N_4 的弧段，则 S 为弧段 A_3。

规则 6　如果 S 是起止结点为 N_2 和 N_4 的弧段，则 S 为弧段 A_4。

3. 结点—弧段的拓扑关系

规则 10　如果 S 为弧段 A_1、A_2 的端点，则 S 为结点 N_1。

规则 11　如果 S 为弧段 A_1、A_4、A_5 的端点，则 S 为结点 N_2。

4. 弧段—面块的拓扑关系

规则 15　如果 S 的左面块为 0，右面块为 B_1 的弧段，则 S 为弧段 A_1。

规则 16　如果 S 的左面块为 B_1，右面块为 0 的弧段，则 S 为弧段 A_2。

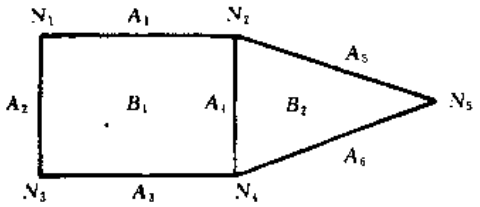


图 4—21　空间实体

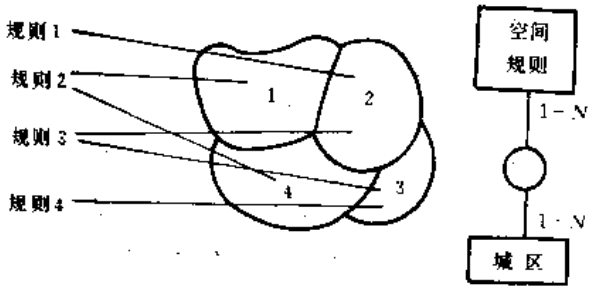


图 4—22　城区与空间规则的关系

规则可用来表示空间元知识。例如，图 4—22 表示了城区与空间规则的关系。对于给定一个点 (X, Y)，要确定使用的空间规则。这个过程首先需要解决查询多边形内的点的问题，然后定义元规则：

IF (X, Y) *belongs to zone*1 *THEN rule*2 *is applicable*

IF (X, Y) *belongs to zone*2 *THEN rule*1 *and* 3 *are applicable*

IF (X, Y) *belongs to zone*3 *THEN rule*3 *and* 4 *are applicable*

IF (X, Y) *belongs to zone*4 *THEN rule*2 *is applicable*。

四、空间知识的框架表示

1. GIS 属性数据描述

框架名：属性数据

super—class：GIS 数据

sub—class：控制点、水系、公园、建筑物、电力设施

框架名：建筑物

super—class：属性数据

sub—class：医院、学校、居民住宅、饭店、工厂、商店

房主：

value—class：字符串

value：未知

街道名：

value—class：字符串

value：中山路

框架名：饭店

super—class：建筑物

sub—class：旅社、涉外宾馆、餐馆、招待所

电话号码：

value—class：字符串

value：0518—5419880

职工人数：

value—class：整数

value：245

2. GIS 几何数据描述

框架名：几何数据

super—class：GIS 数据

sub—class：复杂地物、简单地物

框架名：简单地物

super—class：几何数据、复杂地物

sub—class：点状地物、线状地物、面状地物

框架名：点状地物

super—class：简单地物

控制点：

value：(M_1, M_2, Z)

电视塔：

value：(M_1, M_2, Z)

……

五、空间知识的过程表示

所谓过程表示就是用一段子程序来表示一些知识。在推理过程中经常需要用到经计算后的一些参数值，这些参数值的求解过程就可用过程表示。例如，有一些只由顶点坐标定义的三角形，要查询具有 30°角的所有三角形。假设如图 4—23 所示的三角形一些数据存在数据库中：

$R1\ (Triangle-ID, A, B, C)$

$R2\ (Point-ID, X, Y)$

另一些数据不存在数据库中：

R3 (*Triangle—ID*, *a*, *b*, *c*)

R4 (*Triangle—ID*, *alpha*, *beta*, *gamma*)

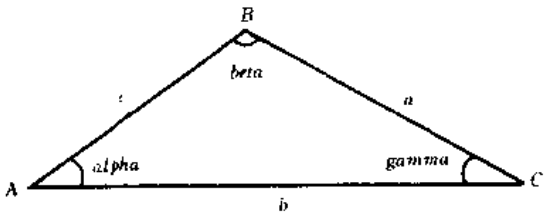


图 4—23 三角形

要回答上述查询，可用过程表示求出线段长度：

$$a^2=(X_B-X_C)^2+(Y_B-Y_C)^2$$

$$b^2=(X_C-X_A)^2+(Y_C-Y_A)^2$$

$$c^2=(X_A-X_B)^2+(Y_A-Y_B)^2$$

然后再用过程表示求出角度：

$$alpha=\arccos\left[(b^2+c^2-a^2)/2bc\right]$$

$$beta=\arccos\left[(c^2+a^2-b^2)/2ca\right]$$

$$gamma=\arccos\left[(a^2+b^2-c^2)/2ab\right]$$

六、不精确的空间知识表示

空间知识常常具有不确定性、模糊性、不完备性。描述这种不确定的知识通常有四种方法：可信度方法，主观贝叶斯（Bayes）方法，证据理论，可能性理论（模糊集合理论）。这些方法可以同前述的知识表示方法结合一起使用。例如，可能性理论与前述知识表示方法结合，形成模糊逻辑、模糊规则、模糊框架、模糊语义网络等方法。模糊知识用隶属度(membership degree) 来描述。例如，假设要查询徐州市东湖区年轻人的数量，已知道某些居委会年轻人占 95%，另一些居委会只占 60%，其它一些居委会只占 30%等等。根据这些隶属度求出东湖区的隶属度，从而求出东湖区年轻人数。

七、空间知识工程示例

这里以一个新医院选址为例，说明空间知识工程的应用。如图 4—24 所示，已知如下知识（用逻辑表示）：

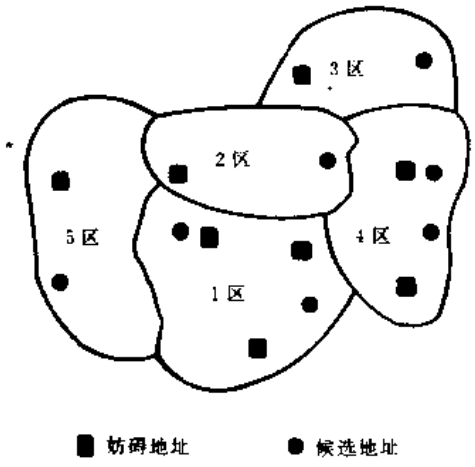


图 4—24 新医院选址的必要信息

(1) 各区对新医院的需求，以各区人数作度量，各区中心坐标：

R1 (*Zone—ID*, *Population*, *X—zone*, *Y—zone*)

(2) 候选地址及各地址的特征，如土地可用性、土地价格、及附近环境等：

R2 [*Site—ID*, *X—site*, *Y—site*, (*Features*)*]

(3) 各区到选址的可达性，以各区到选址的距离表示：

R3 (*Site—ID*, *Zone—ID*, *Km—distance*)

R3 (*Site—ID*; *Zone—ID*, *Time—Distance*)

(4) 某些选址不宜建医院，例如选址离消防站相当近。这些妨碍物以无用度来描述：

R5 (*Hindrance—ID*, *X—hinder*, *Y—hinder*, *Hindrance—type*, *Hindrance—degree*)

从这些已知知识开始，运用类似于运筹学的空间推理，这种推理主要是选取权函数。推理步骤如下：

(1) 根据距离与人口之比求出初始权函数——影响系数。

*A*1 (*Site*－*ID*, *Accessibility*－*potential*－*index*)

(2) 对每一个选址，从妨碍物开始，根据妨碍度和选址到妨碍地址距离的倒数，计算出一个专用统计量——妨碍系数。

*A*2 (*Site*－*ID*, *Hindrance*－*potential*－*index*)

(3) 类似地，根据选址的特征和影响系数及妨碍系数计算出各个选址的实用权系数。

*A*3 (*Site*－*ID*, *Utility*－*weight*)

(4) 根据最大实用权系数，可求出最可行的新医院地址。

使用不同的权系数，例如使平均行程最小，可得到另外的选址方法。

第五章　商品化GIS软件系统介绍与评述

第5—1节　概　　述

80年代以来，随着计算机硬、软件技术的飞速发展，大容量存取设备的使用，为空间数据的录入、存储、检索和输出提供了强有力的手段，用户屏幕和图形、图像卡的发展增强了人机对话和高质量图形显示功能，促使GIS朝着实用化方向迅速发展。据不完全统计，到1993年末为止，商品化的形形色色的GIS软件达一百多种，市场销售额达四亿多美元。有代表性的GIS软件如ARC/INFO，GRASS，MGE，GENAMAP，MICRSTATION，system 9，PHOCUS等。其中ARC/INFO软件首次把关系数据库管理系统与图形信息的存储管理有机地结合起来，并解决了矢量数据处理的关键技术问题，如多边形拓扑关系的自动建立、空间与属性信息的查询检索、数据的叠加技术等。以工作站为平台的ARC/INFO系统版本7.0集矢量数据、栅格数据、行程编码为一体，充分利用面向对象的思想形成了ARCSTONE模块，在软件系统中引入扫描技术形成了ARCSCAN模块，加强了具有图像处理与分析功能的GRID模块，具有分析和计算特点的NETWORK、TIN、COGO等模块，加强了用户开发工具。system 9软件引入了面向对象的思想，把空间信息和属性信息的管理融合在一个关系数据库管理系统中，还成功地开发了独立于system 9软件本身的分析工具箱（ATB）。INTERGRAPH软件MGE则把图像处理与分析、CAD和GIS技术有机地结合在一起。

中国的GIS科技人员一方面适当引进国外典型的有代表性的系统和软件，进行学习和消化；另一方面根据我们的国情，独立研制自己的GIS软件，以满足日益增长的用户需求。中国科学院资源与环境信息系统国家重点实验室联合中国科技大学、中科院遥感应用研究所、北京大学等单位，正在研建有中国特色的面向要素的GIS软件系统。该系统的硬件环境是：SUN670/470服务器，SUN SPARC10/2工作站，PC486以上微机以及以太网分布式网络；软件环境是：SUN OS4.1.3以上操作系统，MS WINDOW3.1以上操作系统，OPEN WINDOW3.0以上窗口管理器，XLIB/XGL图形支撑软件，扩展的关系数据库管理系统软件。这一系统建成后将是一种面向要素的，以物体为中心的行程编码为主的多种数据结构并存，具有面向分析的分析工具包的GIS软件。

由国家测绘局资助，武汉测绘科技大学研制开发的GIS软件——吉奥之星(Geostar)已经推出。该系统采用面向对象的系统分析与程序设计方法，面向对象的数据模型和数据库管理系统，基于Windows等的环境。它具有属性、图形、图像整体结合，多窗口多任务多用户的功能，能够和已有的多种GIS系统（如ARC/INFO，Genamap，system 9等）及Auto CAD进行数据交换、公开数据格式、实现数据共享。

此外，北京大学在原Spaceman GIS软件的基础上，经升级换代已推出城市之星Citystar GIS软件，属高档微机型，适用于城市规划、建设及管理，资源与环境信息采集、处理、分析管理。

然而，可以说至今真正三维的，更不用说四维时空的实用 GIS 软件还没有问世。商品化的各种 GIS 软件都存在着这样那样、程度不同的缺陷或不足，主要表现在以下三个方面：

1）用户接口　由于 GIS 用户的多样性，需要解决的任务、问题的复杂性，以及开发研制 GIS 软件的艰巨性（不象普通二维数据库管理系统那样简单），尚未能设计出一种适合于 GIS 特点的用户接口语言，来解释多种专业用户提出的实际问题，方便缺乏 GIS 专门知识用户的使用。

2）数据结构和数据模型　现有的 GIS 软件系统都是面向图形或记录的，基本上没有摆脱现有关系数据库管理系统的模式，因此难以适应以图形，图像和属性数据整体结合为特征的动态空间数据的管理。

3）三维功能　虽然许多 GIS 软件具有三维分析模块，具有对区域地表景观表现和描述的能力，但也只能表达三维空间的表面，或者说是 2.5 维的表达方式。而真实的三维空间是立体的、多层次的，且具有复杂的内部结构，如矿区的地质资源和复杂的井巷采掘系统，带有悬崖峭壁的地形地貌，现代化城市的基础设施和形形色色的建筑群等。

总之，虽然 GIS 软件系统的研建和实际应用取得了很大成效，但是距离人类社会持续协调发展和各行各业的需求还差得很远，需要全世界多个学科科技人员的共同努力。特别是中国是一个大国，GIS 拥有广阔的市场。我们应该抓住机遇，力争成为 GIS 软件开发和应用的强国。

下面介绍几种有代表性的 GIS 软件。

第 5—2 节　ARC/INFO GIS 软件

ARC/INFO 是美国环境系统研究所（ESRI）于 80 年代推出的通用性 GIS 软件。这一套软件可在多种型号的微型计算机和工作站上运行，支持多种输入输出设备。在 IBM PC/AT286/386/486 及其兼容机上运行的版本称为 PC ARC/INFO。到 1994 年最新的 PC ARC/INFO 版本是 3.4.2，工作站上的最新版本是 ver7.0.2。它们是当前我国使用较广泛的 GIS 软件。

一、PC ARC/INFO 系统的软件结构

在功能上该软件可分为两大部分：ARC，用来管理坐标位置信息，实现用拓扑数据模型定义空间数据；INFO，用来管理属性信息，实现用关系数据模型定义属性数据。系统功能的实现是通过由众多命令组成的“工具箱”（toolsbox）向用户提供不同的命令。在系统内部是一系列 SML（Simple Macro Language）［在 Host-ARC/INFO 系统中是用 AML（ARC Macro Language）］程序、命令语言程序（BAT 文件），以及由 FORTRAN 或 C 语言编写的主程序和子程序的集合。所有的命令都体现为一个与其相应的 SML 程序或命令语言程序。其中包括：

1）SML 程序或命令语言程序　其功能是输入、检查和传递命令参数，调用其下一级的 FORTRAN 程序，以实现不同的功能。

2）FORTRAN 主程序　其功能是实现参数的输入，完成调用子程序的功能。

3）FORTRAN 子程序模块　ARC/INFO 系统中包含一系列子程序模块，每一模块完成不同的功能。按其功能不同，可把它们分成如下几种类型：

（1）接口模块：提供一系列接口，使不同程序调用可以享受系统资源。在不同的操作

系统中，该模块不同。

(2) 数据支持模块：其主要功能是定义不同数据的数据结构及管理这些数据，它不随操作系统的变化而变化。

(3) 功能模块：它主要是根据不同数据处理的要求，用相应的算法对数据进行处理及分析，它也不随操作系统的变化而变化。

ARC/INFO 的主要子系统有

PC ARC/INFO STARTER KIT：它提供了 PC ARC/INFO 的基本功能，如弧段数字化系统（ARC DIGITIZING SYSTEM）用于图形数字化；CLEAN 和 BUILD 用于建立线或面的拓扑关系；TABLES 用于属性数据的管理，以及绘图、打印和拷贝、改名等。

PC OVERLAY：用于多边形叠加、线和点的叠加，以及缓冲带的产生。

PC GRID CONVERSION：用于矢量图形和网格图形的转换。

PC ARCPLOT：用于人机对话方式产生、显示图形和生成绘图文件。

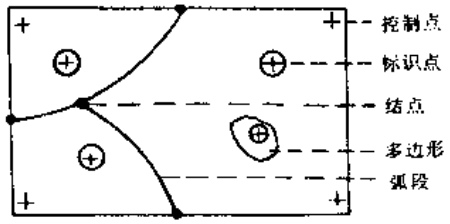
PC ARCEDIT：交互编辑模块。它将数字化和编辑功能结合在一起，可进行图形数字化，生成图形文件和对坐标数据进行编辑、检查、修改。

PC NETWORK：执行多种网络分析功能，包括资源分配（ALLOCATION）、最佳通路分析（ROUTING）和形成新的多边形集合（DISTRICTING）等。

二、PC ARC/INFO 系统的文件结构

从数据模型上看，ARC/INFO 采用的是拓扑数据模型与关系数据模型相结合的混合式数据模型，空间数据与属性数据通过以内部代码（INTERNAL NUMBER）和用户标识码（USER-ID）作为公共数据项，以实现两者的联接。从文件结构上看，ARC/INFO 系统采用图幅工作单元（COVERAGE）作为数据库数据组织、存储和管理的基本单元。

图幅工作单元是对传统地图的一种模拟，它相当于一张地图或传统制图的单要素版，用以表示某地给定区域内地理要素（如土壤、土地利用或森林分布等）的空间位置和专题属性。一个图幅工作单元中的要素可以分五种不同的类型：图幅控制点（TICS）；弧段（ARCS）；结点（NODES）；多边形（POLYGONS）；标识点（LABELPOINTS）（图 5—1）。每一个类型的要素均由不同的文件定义。因此，一个 COVERAGE 作为一目录，它有一组文件存储在计算机中，其文件的组成如图 5—2 所示。



(a)

IDTIC	YTIC	XTIC
1	5.07	6.25
2	5.07	31.25
3	35.07	31.25
4	35.07	6.25

(b)

图 5—1　TIC 文件的构成

(1) TIC（控制点坐标及标识码）每一图层都至少需要四个控制点，用来与采用的统一坐标系进行配准，其控制点的坐标和标识码贮存于 TIC 文件中。

(2) ARC（弧段坐标和拓扑关系）它表示线状要素的空间位置。每段弧的信息包括：内部识别码、该段的坐标点数及其坐标数据。

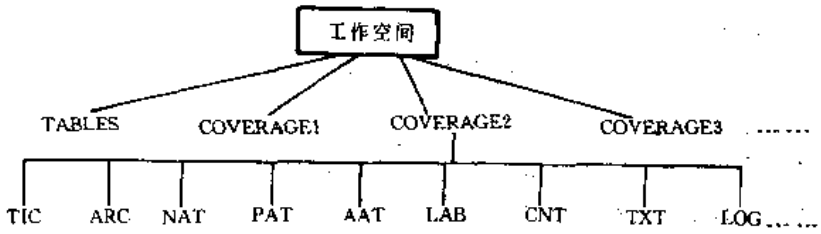
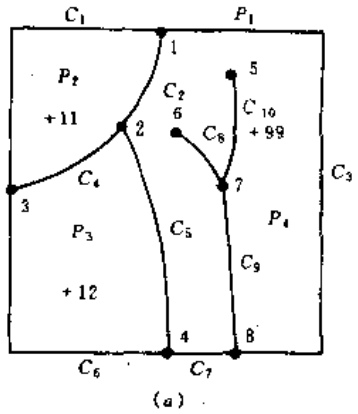


图 5-2 图层工作单元的文件组织

(3) NAT（结点属性）结点为弧段的端点，其空间位置在弧段文件中确定，这里只定义其属性值，包括用户标识码、内部码和关联的弧段（图 5—3*a*）。



结点拓扑

NODE#	ARC#
1	1, 2, 3
2	2, 4, 5
3	1, 4, 6
4	5, 6, 7
5	10
6	8
7	8, 9, 10
8	3, 7, 9

多边形拓扑

POLY#	USER-ID	ARC#
1	0	1, 3, 6, 7
2	11	1, 2, 4
3	12	4, 5, 6
4	99	2, 3, 5, 7

(b)

图 5—3 结点和多边形的拓扑表示

(4) PAL（多边形的弧段表）多边形的空间位置由弧段确定，该文件只存储组成该多边形的所有弧段的内部标识码、该多边形所属结点的内部标识码和该多边形的用户标识码。PAL 文件的记录号是多边形的内部标识码，而该记录号也同时存在于 CNT 和 PAT 文件中，通过查询 PAL 文件，可以提取与多边形有关的所有信息，如图 5—3*b* 所示。

(5) AAT（弧段属性表）它包括每个弧段的一系列属性。其文件形式如表 5—1 所示。

表 5—1 弧 段 属 性 表

PNODE	TNODE	LPOLY	RPOLY	RPOLY	COVER#	LENGTH	COVER-ID
3	1	1	2	1	1	2.00	121
1	4	1	3	2	2	3.00	14
1	2	3	2	3	3	0.75	144
2	3	4	2	4	4	0.75	82

(6) LAB（标识点的坐标及其拓扑关系）每个标识点为一个记录，存储该点的坐标与拓扑关系（多边形的内部标识码，即在 PAL、PAT 和 CNT 文件中该多边形的记录号）。如果该值为零，表示是点要素的标识码。

(7) BND（工作区范围坐标）该文件由一个记录构成，它定义分析或操作时的工作区范围，同时也作为绘图文件的“窗口”。其文件形式为

X_{min}	Y_{min}	X_{max}	Y_{max}
0.5	0.5	15.0	15.0

（8）CNT（多边形标识点表）该文件记录多边形的几何中心及位于该多边形范围内的多边形标识码。每个多边形占一个记录，其记录号与 PAT 和 PAL 文件中的记录号相同。

（9）PAT（多边形或点的属性）对于多边形，存储其面积、周长、内部标识码和用户标识码；对于点状要素，其面积和周长为哑值。多边形的记录号即为该文件中的内部标识码，与 PAL 和 CNT 文件的记录号相同。点的记录号则与 LAB 文件中的记录号同。其文件形式如图 5—4 所示。

AREA	PERIMETER	COVER #	COVER-ID
−10.23	7.5	1	—
2.76	3.04	2	101
4.18	4.82	3	102
2.58	2.39	4	103

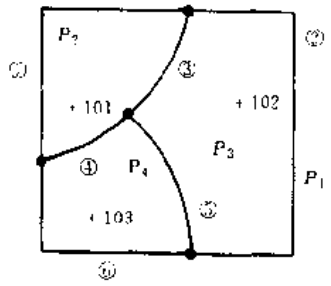


图 5—4　多边形的属性

（10）TXT（注记的位置和内容）该文件存储地图的注记要素，包括注记的位置形状、符号及内容（字符串）。该文件只供显示和绘图使用。

上述文件之间是互相联系的，文件之间的连接操作是通过作为数据项的标识码（USER-ID）进行的。ARC/INFO 用四个不同的 INFO 数据文件存储 COVERAGE 的特征属性数据，这四个数据文件分别是：TIC、BND、PAT 和 AAT。在一个 COVERAGE 中，用户标识码既存放在位置数据中，又存放在 INFO 的属性数据文件中（PAT 或 AAT）。要素属性文件的管理与位置数据的管理是平行的，即当位置数据改变时，要素属性文件相应也跟着变化，同样，属性文件变化时，位置数据也变。这种“平行控制”是由系统软件自动完成的。ARC/INFO 常用数据文件之间的内部关系如图 5—5 所示。

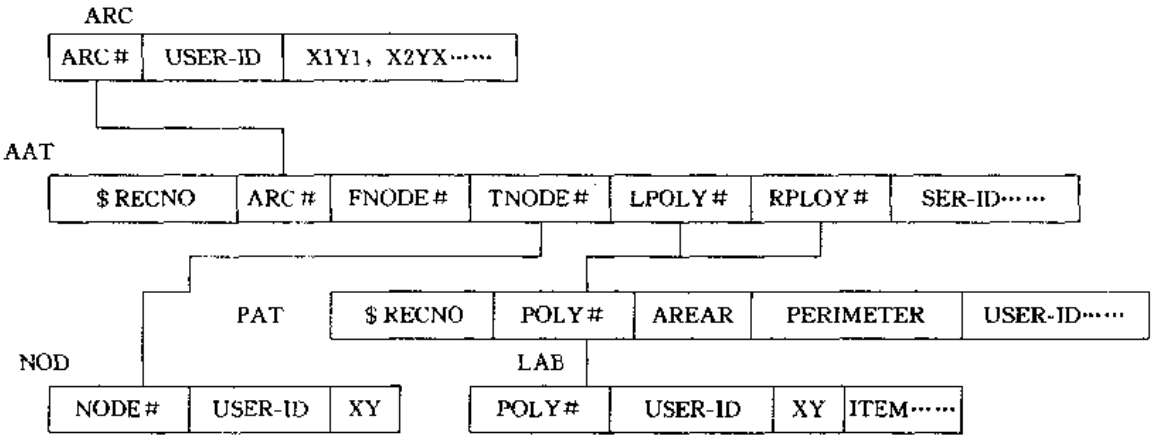


图 5—5　ARC/INFO 文件内部关系图

三、实现汉字注记的技术方案

PC ARC/INFO 软件在空间信息的管理、分析、评价及自动制图等功能上有不少优势。

突出的问题之一是非汉化版本在输出图件上不能实现汉字注记，影响了图件的实用性。要实现整个系统的汉化，难度和工作量很大。笔者在系统消化该软件系统的功能后，开发了汉字注记的功能，方便了使用。

该系统中的 ARCPLOT 子系统具有图形显示和交互式地图编制功能。利用 ARC PLOT 能方便地设置线型、线宽和颜色；其西文文本注记允许用户定义注记符的旋转角度、间隔及字符的大小等。还有线型编辑（Lineedit）和字符编辑（Foutedit）的模块，给用户提供定义多种特殊的线型和字符。据此，实现汉字注记可有如下三种方案[28]：

（1）利用 Pattern 实现汉字注记。

（2）利用 Gridpoly 实现点阵汉字矢量化，尔后用图形叠加功能注记汉字。

（3）利用 SML 编制写汉字模块。

上述三种方案的优缺点比较如下。

第一、二方案的优点：

（1）第一方案把汉字存贮于 IGLFNT 文件中，把汉字作为绘图符号与 ARCPLOT 兼容，能充分利用 ARCPLOT 绘制符号的所有命令。

（2）第二方案不需现存的矢量汉字库，根据普遍使用的点阵汉字就能直接生成汉字 Coverage。

第一、二方案的缺点：

（1）第一方案中一个 IGLFNT 文件只能存贮 570 个汉字，注记一幅图所需的汉字很可能存于几个 IGLFNT 文件中。因此，注记时往往得变换目录或拷贝，深感不便。

（2）第二方案所需的存贮空间大。一个“百”字 Coverage 占 4 千多字节，生成 map 文件后，仍还有 800 多字节。

此外，这两种方案用户都不能自己选定字体线划的宽度和长宽比等，自由度较小。

第三方案的优缺点：

（1）用户自由度大，线划宽、长宽比、角度、大小等可自由定义。

（2）存储空间小，本方案增加的存储空间只有几个不同字体的矢量汉字库。

（3）扩充性好，若要增加一种新字体，修改程序中的几个参数即可调用新增字体的矢量汉字库。

（4）SML 语言为解释型语言，执行速度相对较慢。因此，本方案注记汉字的速度要比前二种方案慢。

权衡利弊可见，第三种方案为最优方案。因此，利用 SML 编制写汉字的模块，实现汉字注记是切实可行的较优方案，并已在实践中应用。

四、PC ARC/INFO3.4.2 的特点和功能

ESRI 于 1994 年 11 月正式推出 PC ARC/INFO3.4.2 版本。它集成于 Microsoft Windows 平台上；继承了 Windows 的汉化功能，即具有一些汉化功能。

1. 集成于 Windows 平台

Rev3.4.2 允许用户在 Windows 环境中运行 ARC/INFO，主要特点和改进的功能包括：图形窗口和命令窗口分开，在图形窗口上可直接完成漫游、放大、缩小、剪切、打印、绘图等功能；扩充了 SML 的一些新命令，用于 Windows 的环境控制和界面设计，如 WIN CB——读写剪切板（Clipboard），WIN DB——建立 Windows 风格的对话框，WIN EX-

EC——控制图形窗口的缩放、漫游、拷贝等功能，WIN MB——建立 Windows 风格的信息框，WIN MENU——建立 Windows 风格的下拉式菜单。此外，交互式图形界面工具，允许用户进行二次开发和界面设计。利用 Windows 的 Notepad、Clipboard、Write、Paint Brush 等工具，可作图形剪切、整饰、制作“幻灯片”式的演示，打印报告等。

2. 汉化功能

该新版本继承了 Windows 的矢量汉字，可在 Chinese star 上实现：制作汉化的菜单、汉化的按钮、汉化的对话框；在数据库中有汉字字段；可有汉字的图形显示，在 PC ARCPLOT 中可用 TEXTFONT 命令调用 Windows 的汉字字体。

五、ARC/INFO7.0.2 版本的特色

较之以前的版本，7.0.2 版本增加了一些新命令和新功能，增加了 ArcStorm，Arcscan 和 ArcExpress 等新模块。

1. 新的命名方式

该新版本开始采用 DOS 的 8.3 命名标准（ISO 9660），其优点是：方便工作站平台和 DOS 平台的数据交换，便于 PC ArcView 与 Arc/Info 之间配合使用；使每个工作区的 INFO 文件数从 1000 增加到 10,000 个。

2. Arc Storm 模块

Arc Storm（Arc Storage Manager）采用客户/服务器计算模式，实现几种要素级（如弧段级、结点级、多边形级）地理数据库管理，并保持数据库的完整性。它不但支持 Arc/Info7.0.2，而且支持 Arc View2.0 和 Arc CAD2.0 版的使用，还能够与 DataBase Integrater 所支持的其它关系数据库系统紧密结合。

3. Arc Scan 模块

该模块集成于 ARC EDIT 环境中，可进行扫描图像的编辑和矢量化，可把多种标准格式的图像转换为 GRID，可进行几何校正、空间配准及投影转换，以校正扫描图纸的变形。“噪声清除”可消除扫描图纸上的斑点，“滤波”能针对某一区域，以其中占主要地位的值代替（如删除图像上的文字）。栅格矢量一体化的编辑环境允许用户对扫描图像进行旋转、镜像、移动、复制、粘贴、填充、绘制等编辑。自动矢量化和交互式矢量化可以跨越间隙，判别虚线，自动取直，自动生成交线以及自动细化和粗化。

4. Region（区域）

Region 是一新的数据模型，可把多个不相邻的或相互重叠的多边形作为一个图层中的一个图形要素，而不是相同属性的多个要素，在数据库中只使用一条记录，而不是多条记录。同一个 Coverage 可以有多个子级（Subclass）的 region，多个子级之间的空间分析，如 Union、Intersect、Clip、Erase 等都操作简便。

5. ARCEDIT

可在 ARCEDIT 下直接生成和维护多边形的拓扑关系，Region 的编辑，溪流方式（Stream mode）的数字化，可使用 Arcplot 命令。在更换编辑要素时，保持原来选择的要素。可作图形的动态平滑漫游，而不必全屏幕刷新。

6. ARCPLOT

可在 ARCPLOT 下自动绘制坐标格网，注记的自动布设（调整注记的密度、位置），可用复合的扇形或方形表示各要素的多种属性，可执行 ARC 命令。

7. GRID

GRID 模块有许多改进，具有较强的栅格 GIS 分析、编辑、显示功能。增加了一些一元和多元分析模型，如一般回归，逻辑回归，多元回归，建立样区，ISO 聚类，非监督分类，扩散分析，水模型分析，邻近分析和区域分析。还有交互式栅格数据编辑，Nodata 的透明显示等。

8. Network

在 Network 功能下对路线、网络的建立和编辑更为方便。在编辑 Route 和 Section 时，能自动检查路线属性的错误，提供多种网络选址和分配模型，包括最短距离（Mindistance），有限制的最短距离（Constrained Mindistance），有权重的最短路径，最大出席数（Maxattend），最大覆盖面积（Maxcover），有限制的最大覆盖面积（Constrained maxcover）等。

9. Arc Express

它是图形显示加速模块，能使图形显示速度提高十倍以上。

10. 其它

此外，Arc/Info7.0.2 版本可在应用程序之间通讯（Inter-Application Communication，IAC）时，与本机或网络上其它结点的应用程序相互通讯，包括发出申请，执行请求，发送或接收数据，返回分析结果。IAC 可在多种协议标准下实现，如 Unix RPC，PC Microsoft Windows DDE，Macintosh Apple Event。

该版本可在流行的 UNIX 工作站上实现全面的汉化，如 SUN、DEC、SGI、HP、IBM 等。

在图形输出方面增加了新的输出格式，如 VCGL2 和 RTL，RTL 支持的 HP650C 喷墨绘图仪和 HP7600 静电绘图仪的图形/图像混合输出。

第 5—3 节　桌面式 GIS 软件 ARCVIEW2.0

最近，ESRI 在 ARCVIEW1.0 的基础上又推出了 ARCVIEW2.0 版本。它集空间图形、关系数据库、统计图形、空间分析、网络通讯、面向对象的程序设计于一体，图形用户界面直观，借助鼠标器即可完成许多 GIS 任务。

（一）概况

ARC VIEW2.0 以 Project 为基本应用单元。它由多种 Document 组成：Views，地图显示、信息查询、空间分析；Tables，数据库管理、电子表格；Charts，统计图形制作；Layouts，组合设计，地图制作；Scripts（Avenue），面向对象的程序设计、调制、开发工具。各 Document 之间为动态联接，即任一 Document 的信息被操作或改动，其余的都自动更新，以反映其结果。每一种 Document 都有其相应的图形用户界面，包括 menu Bar，Button Bar，Tool Bar，并且它们都是可编辑的。

ARCVIEW2.0 可在 ESRI 支持的所有 Unix 平台，以及 Open VMS，Apple Macintosh，Microsoft Windows 上运行，并保持界面完全相同。此外，可在 PC 机平台上借助 Windows 的汉化（如使用中文之星等）来实现 Arcview2.0 的汉化。

（二）VIEW

1. 数据格式

该软件支持多种空间数据格式，如 ARC/INFO 的 Point，Polygon，Line，Route，Region，

Grid 以及 TIFF，ERDAS，BSQ，BIL BIT，SunRaster，DXF，Run-length 压缩的文件。该软件还引入另一种数据格式——Shape，它是一种无拓扑关系的矢量数据，但有属性数据，可与 Table 相连。Shape 格式是开放的，用户可直接生成、编辑或删除 Shape 文件。

此外，点的位置信息可直接显示和编辑，其坐标可由 ASII 文件、数据库文件、外部数据源（如 GPS 接收机）输入。借此以实现对动态目标的实时跟踪显示。

2. 图形显示和信息查询

Coverage：可按任一属性项分级显示。分级的方式有等距、等样本数和自定义方式。并可用随机、过渡（ramp）或自定义方式为每一级指定合适的 Marker、Color、Shade、Line 等。

Grid：可做线性拉伸，等积拉伸，灰度显示，假彩色制作和颜色表定义等功能。

Image：有单波段显示，多波段彩色合成，透明与不透明两种叠加方式。

注记：可交互式加入，也可把任一属性项的全部或部分注记到图上。

支持空间查询，逻辑查询。查询结果在 View 和 Table 中显示，在 Chart 中自动生成统计图。

3. 投影变换

支持 10 多种投影变换。一旦一个 VIEW 的长度单位（Unit）、投影方式、参数（经纬线等）改变，即能以新的投影方式显示出来，而无需对原始数据做投影转换。

4. 制图综合

该版本能较好地解决地图比例尺与信息容载量问题，能指定任一 Theme 显示的比例范围，使图形随放大而详细，随缩小而概括，以保持地图的清晰性和逻辑性。

5. 热连接（Hot Link）

所谓热连接就是把某一要素和某种图像、文本文件、Document、Project 或 Avenue 程序连接起来，当鼠标点中该要素时立即显示这些数据或执行某个程序。

（三）Table

1）数据源　该版本在 Table 下可直接读取 INFO、DBF、ASCII，并通过 SQL Connect 和任何 SQL DBMS（如 ORACLE，SYBASE，INFORMIX，INGRES 等）相联接。

2）功能　具有数据建立、增删、修改、派生、电子表格形式显示、字段选择，支持假名等功能。可进行一对一、一对多、多对一、多对多的数据库联结，支持复合条件查询。可对多个字段进行多种统计，计算平均值、求和、最小值、最大值、标准差、极差等，并以表格（Table）形式表示。具有动态连接能力，数据库操作结果瞬时改变，并显示在 View，Chart，Layout 上。

（四）Chart

Chart 以直观、动态的统计图方式表示 View 和 Table 当前所选择的内容，使数据分析简单化、视觉化。它提供了六类统计图：Area，Bar，Column，Line，Pie，XY 散点图。Line 适于表现变化趋势（率），Area 适于表示变化量，Bar 和 Column 适于表示不同样的比较，Pie 适于表示结构组成，XY 散点图适于表示趋势或相关程度。

（五）Layout

它用于对多种 Document 或地图的各种 Views，Charts，Tables，Pictures，Graphics（Lines，Circles 等），以及图例、指北针、比例尺、标题等，每一元素的大小和位置进行调

整。每当地图设计完毕，可通过 Print 选择项直接由打印机或绘图仪输出。也可以转换成多种格式，如 Postscript，Adobe Illustrator，CGM (Binary，Character，Clear Text)，Bitmap，Metafile，PICT。

（六）Script（Avenue）

Avenue 是一个面向对象的程序设计语言和开发环境，可用以重新组织 Arcview 的界面和功能，执行其它的多媒体程序。具有数据类型，变量，List，支持代数运算和逻辑运算，支持程序调用和参数传递。可移植到其它平台的 Arcview 上执行。

（七）空间分析

1）theme-on-theme selection　完成叠加、缓冲区、临近分析等空间操作。它有四个参数：被选图层（Target Theme）；选择图层（selector Theme）；空间关系（Spatial relationship），如落入，包含，中心线落入，包含中心线，切割，在一定距离内，距离（Distance）。

2）空间联结（Spatial Join）　通过 Table 中的 Shape 项，把不同的主题或同一主题的要素按位置联结起来，从而产生诸如 Point-near-line，line-in-line，Point-in-Polygon 等多种空间分析。

（八）地址匹配

所谓地址匹配就是根据报案地址找出案发点位置，然后用 Arcview 分析周围环境，确定行动方案。还具有地址编码和地址匹配功能，允许设置匹配的严格度（Spelling sensitivity）和要求的最小得分，实现了模糊匹配，增强了地址匹配的灵活性和实用性。

（九）应用程序间通讯（Inter-Application Communication，IAC）

Arcview2.0 可与本机或网络上其它结点的其它应用程序之间相互通讯，包括发送或接收数据，执行请求，返回分析结果。IAC 可在多种协议标准下实现，如 PC Microsoft Windows DDE，Unix RPC，Macintosh，Apple Event。Areview 既可作 Client，又可作 Server。

第 5—4 节　GENAMAP GIS 软件

GENAMAP 是澳大利亚 GENASYS Ⅱ 公司开发的 GIS 软件产品，它基于 UNIX、XWINDOW 操作系统。它的功能较强，具有较好的一致性、开放性和易操作些。可应用于自然资源清查和管理、地质勘测、城市规划和管理、环境监测和管理、交通管理，以及社会经济分析和管理等许多方面。

（一）系统的运行平台和基本结构

GENAMAP 的最新版本是 V6.1，运行的软件平台是 UNIX 和 Xwindow，硬件平台可以是微机和工作站，举例如下：

工作站或微机	操作系统
HP9000/700 系列	HP—UX9.01
HP9000/800 系列	HP—UX9.00
HP9000/300/400 系列	HP—UX9.00
IBM RS6000	AIX3.2.5
SUN OS	SUN OS4.1.3

Sun Solaris2	Solaris2.3
ICL DRS 6000	7.3.5
Digital DEC 工作站	Ultrix4.3a
Digital Alpha	OSF 1V2.0
Silicon Graphics	IRIX5.1
IBMPC 兼容机（486 以上）	SCO Open Desktop3.0 或 Unix Ware

该系统共由十个模块组成，其中 Genamap 和 Genacell 是核心模块。此外，Genarave 和 Genacivil 两个模块既可独立使用，也可与 Genamap 组合使用。十个模块的基本功能是：①Genamap 是一个完整的矢量式 GIS 系统，是整个系统的核心模块；②Genacell 是实现 Genamap 栅格图形数据处理、管理和分析的模块；③Genaref 是从 Genamap 向其它系统或其它系统向 Genamap 输送图形信息的格式转换和接口模块；④Genacomm 是 Genamap 向外部商用关系数据库输送数据或调用外部数据库的接口模块；⑤Genadoc 是文档图像处理和管理模块；⑥Genarave 是栅格式图形数据自动矢量化模块；⑦Genacivil 土木工程辅助设计和测量数据处理模块；⑧Genius 用户菜单生成和管理模块；⑨Genacarto 综合制图生成管理模块；⑩Genavive 矢量图、影像图集成显示、处理模块。

（二）各功能模块简介

1. Genamap、Genaref、Genacomm、Genius、Genacarto 及 Genavive 模块

Genamap 既是一个独立的 GIS 系统软件，也是整个系统中的核心模块。它包含有空间型关系数据库，用于座标存贮、空间图形管理、映射变换、检索查询和图形显示。该数据库可自动建立库中每一层图内点、线、面以及各层图之间的拓扑逻辑关系，实现库内图形连续式的无接缝完整图形。通过各图形与属性数据库中的属性数据建立关系，可生成动态空间图形视图。其关系型属性数据库（Attribute）用作属性数据管理，以及与各层图形建立相关关系。其图示图例符号库用于存贮和管理各类专用符号、图形，用户还可自造符号。各种图形编辑工具可用于数字化、校正、修改、拼接和更新各类图形。此外，还有多种空间分析工具，如网络分析、叠加分析、缓冲区分析、重新赋值分类分析、相邻分析等。使用其过程化程序设计语言 Script，可将 GENAMAP 系统中十个模块的所有功能组合成一个可执行程序，满足用户的各种需求。

Genaref 是 Genamap 与其它 GIS 或图形软件系统交换数据时，对图形数据进行格式转换处理的接口。它可以转换和接受 DXF、DLG、DGN、ARC Export、TIGER、Etak、SPOT/Landsat、USGS DEM、VPF、IGES 等系统的不同格式的图形数据文件。

Genacomm 是 Genamap 与各类关系型数据库的接口。它可以连接 Oracle、Ingres、Informix、Sybase 和 IBM'S SQL/400 等关系型数据库系统。

Genius 是 Genamap 提供用户屏幕菜单生成和管理功能的模块。它将 Genamap 系统中的图形库、命令过程和数据库组织在统一的菜单管理之下，具有下拉、弹出、覆盖、滚动等功能菜单。

Genacarto 是一个具有图形用户界面的综合制图工具模块，可快速生成、放置、编辑地图的各种元素和组合成分（图形的或非图形的），并作硬拷贝输出。

Genavive 能提供仿真图像显示环境，可显示和操作多种图像数据。并可以影像为背景底图，在屏幕上用鼠标跟踪数字化，形成矢量图。显示格式直接支持 TIFF、JFIF 和 Genamap

栅格图。

2. Genacell 模块

该模块是栅格图形数据处理、管理和分析的模块，可生成栅格单值图形，用插值法实现从矢量图生成或转换为栅格图。可生成区域报表，产生三维 DTM 模型。具有分类、叠加、距离分析和邻近分析、容积计算等功能。

3. Genacivil 模块

该模块为土木工程设计和测绘模块，具有工程制图、测量处理（COGO）、三维模型建立、管线设计等功能。Surveying 和 drafting 工具提供野外电子手簿数据输入、测量平差、COGO 坐标几何、图形操作等功能。TIN 地表模型生成工具提供等高线生成、三维立体图、高程计算、面积和容积计算等功能。道路设计工具（Road design）与 TIN 地表模型相结合，进行道路中心线、横纵断面设计和显示、输出，计算面积和容积量。水文和地下管网辅助设计工具（Hydrology）能提供水流向、洪水淹没和地下管线等工程辅助设计的功能。

4. Genarave 模块

它既可与 Genamap 构成完整的系统，也可在 UNIX、Window 平台上独立使用，是一图形自动矢量化模块。它有许多功能特点，例如细化图形线，等高线自动生成，点、线拓扑关系生成，线形识别，地图定位管理等。

第 5—5 节　Intergraph MGE GIS 软件

英特格拉夫（Intergraph）公司是生产交互式图形计算机系统的一家大公司，总部设在美国阿拉巴马州的汉茨维尔市，在 50 多个国家设有分公司。该公司于 1984 年首先在北京设立了代表处，之后在上海和深圳设立了办事处和企业。在中国的主要业务领域涉及勘测、土木工程、地理信息系统、机械设计、建筑设计和石化工程等。

Intergraph 的典型 GIS 软件是模块化、一体化的产品 MGE（the Modular GIS Environment）。它能够运行在多种操作系统平台——DOS/Windows、UNIX、Windows95、Windows NT 上，能够对不同的数据类型如矢量、栅格、二进制、Voxels 等进行逻辑操作和综合处理，绘图输出也不限于矢量数据。

（一）基本 GIS 模块和综合性的运行于 PC 机上的模块

MGE 有三个基本模块，利用它们可以完成 GIS 中的项目管理，图形数据的录入，图形数据编辑修改，数据库的建立和查询，坐标系统的定义及制图输出等。对于 PC 机而言，有一个综合性的模块 MGE-PC。

1. MGE Basic Nucleus（MGNUC）模块

MGNUC 是 MGE 家族中所有软件产品的基础，它提供 GIS 制图应用的基本平台。其基本功能包括项目管理、数据查询与检索、投影坐标系统的应用。它能支持单机/网络的运行方式，图形数据能支持 Auto CAD 和 Microstation 图形文件，并支持 SDTS 标准。

2. MGE Basic Adminsitrator（MGAD）模块

该模块提供单机或网络上用户的 MGE 系统和数据库的管理工具。能进行数据库表的定义和表间操作。对数据库操作时，可利用 Intergraph 提供的一个工具 RIS（Relational Interface System）。通过 RIS，MGAD 能够将图形元素和各种支持 SQL 的数据库产品连接起来，这些数据库系统包括 Oracle，Sybase，Informix，Ingres，Xbase，MSSQL Server 等。

只要这些数据库系统运行在网上，而不管它是运行在何种硬件平台，何种操作系统，则MGE 通过 RIS 就能够完成图形和数据表的联系。

3. MGE Base Mapper（MGMAP）模块

MGMAP 模块提供对数据的采集、操作及归整处理，可同时进行交互式和批命令处理。功能包括：数据收采、整理，数据库交互式查询、编辑，GIS 图形特征的建立和符号定义，还能进行面积、周长、坐标的计算等。

4. MGE-PC 模块

该模块是运行于 PC 机上，在 MS-Windows 95，Windows NT，DOS/Windows 环境下的一个 GIS 系统，具有图形数据的采集、查询、显示和修改的功能，还能进行 MGE 与 MAP-Info GIS 软件数据间的相互转换。MGE-PC 也能通过 RIS 支持网上的数据库，还具有空间查询、分析，缓冲区生成（bufferzone），多边形图形彩色充填等功能。

（二）几种主要分析、处理功能模块

1. 拓扑空间分析模块 MGA

需要作拓扑分析时在 MGE 基本模块的基础上，利用 MGE Analyst（MGA）模块就可完成这一工作。其功能包括：拓扑结构的生成，缓冲区生成，空间图形及数据库的查询、分析、逻辑（AND、OR、NOT）和空间关系操作，报表和图形输出，计算长度和面积，利用查询结果生成新的图形文件，进行多边形之间的拼接和融合等。

2. 线性网络分析模块 MGNA 和 MGSM

在进行交通和管线的管理、分析时可利用 MGE Network Analyst（MGNA）和 MGE Segment Manager（MGSM）两个模块。MGNA 提供线性网络的建立、管理和分析，即利用数据库中的图形和属性数据进行网络连通性分析，路径评价及费用分析，还能定义道路的限制条件如障碍、单行线、不能转弯的交叉路口等。分析结果可存储于数据库或显示、生成图形文件。

MGSM 模块提供线段、符号化线性特征（道路、管线）的分析评价等，如路面状况、事故率、交通堵塞情况的分析及结果制图。

3. 地形模型分析模块 MGM 和 MTA

两块三维地形分析模块 MGE Modeler（MGM）和 MGE Terrain Analyst（MTA）的功能强弱不一样，用户可根据工作需要选用。MGM 提供点高程、数字化等高线、边缘线和流水线等的输入，生成三角网或格网的 DTM（数字地面模型），并以多种图形表达，生成工程图或等值线图。

MTA 除包括 MGM 的基本功能之外，还能提供更强的分析工具，包括地形坡角、坡向及高程的生成，交互式视图分析，栅格影像的叠加分析等。

4. 格网分析模块 MGGA

MGE Grid Analyst（MGGA）模块是基于格网数据的 GIS 分析工具，主要分析功能包括：叠加（Overlay）和代数运算（Map Algebra）。Overlay 是利用 GOAL 语言（GIS Oriented Analysis Language）来建立分析函数。MGGA 的其它功能有矢量/格网方式的相互转换，多元回归分析，相关及线性回归分析，直方图统计，缓冲区及优化分析，裁剪、采样、拼接、加密操作及统计等。

5. 地下三维分析模块 MGVA

MGVA（MGE Voxel Analyst）提供三维模型的建立与分析。它接收三维的格网数据，这些数据可包含一个或多个属性值，格网数据可以是规则的或不规则的，可作多种显示操作，如垂直/水平断面，透视图等。数据可以用 ASCII 码或 GRD 文件输入。

6. 图像处理与分析

对航空和卫星图像的处理和分析主要由 MBI（MGE Basic Imager）和 MAI（MGE Advanced Imager）两个模块来完成。前者提供基本的图像处理和管理，主要包括图像制图、地图更新，矢量图形与栅格影像的综合处理，影像压缩，动态增强处理，假彩色、直方图和色调调整，影像几何运算（影像到影像、影像到图形的校正，旋转、裁剪、镜像及镶嵌处理，边缘提取，卷积运算等）。MAI 模块则能提供多波段、多功能的影像处理与分析，专题信息提取，栅格到矢量的转换等。具体功能有：光谱增强与分析（比值变换、主成分变换等），富里叶分析，监督与非监督分类，平滑、彩色编码、错误分析、编辑修改等。

7. 制图分析

图形的整饰和输出主要由 MGFN（MGE Map Finisher）和 MAPPUB（MGE MAP Publisher）两模块（系统）完成。MGFN 主要用于地图的整饰、制图符号的定义，多种图的合成，图形的裁剪、拼接，加网格、边框、注记和图例等，对地物进行颜色、符号填充、优化处理等，使输出的图形逼真、优美。

MAPPUB 是一地图制版系统，包括一台 Intergraph 的扫描/胶片记录仪（Mapsetter）。该系统可将地图/影像图形进行分色，合成输出到胶片记录仪上，并且记录在正片或负片上的分色胶片可直接印刷。

8. 数据交换

MGE 除接受和输出标准格式的图形/图像和数据库之外，还支持 SDTS 标准，并通过 MGAL（MGE ASCII Loader）模块来完成 MGE 与其它 GIS 系统之间的数据交换。这种 MGE 的 ASCII 文件中可包含图形的坐标及其相关的属性信息，能支持 ARC/INFO、EDS-GDS、Genasys 等系统输出的 ASCII 文件。MGAL 还可以处理地面测量、GPS 测量的点数据等。

此外，该 GIS 系统还能提供 AM/FM 管理系统 FRAMME，使各种工程信息与地理信息综合在一起提供各类用户使用。多媒体 GIS 软件——MGE Vistamap 能将抽象的点、线、面等图形符号用声音、动画、相片、图表、文字等加以包装，多角度地再现地理特征，使提供的信息更直观、形象，易于接受。

第 5—6 节　GRASS GIS 软件

地理资源分析支持系统（Geographical Resources Analysis Support System，简称 GRASS）是美国军用工程公司建筑工程研究所与美国土壤保持部、地壤管理局、环境保护局等 11 个部门联合研制的一种多功能地理空间信息系统软件工具。它基于微型机工作平台，并将栅格式 GIS、矢量式 GIS 和遥感影像处理系统溶为一体，包括资源数据库管理系统（RDBMS）、专家系统和遥感影像处理系统，以满足资源与环境等领域中多方面的需要。

GRASS 以 Unix 为操作系统，C 语言编程，包括 250 个用户程序命令，300,000 行的 C 语言程序语句。它是命令式驱动软件，易于扩充，利于软件的进一步开发。GRASS 约占用 120MB 的空间。该系统的基本硬软件配置如下：PC386 或 486、DEC 工作站、SUN 工作站

(内存 8MB 以上)、IBM PS-2 (AIX) 等；硬盘 120MB 以上，单软或双软驱动器；高分辨率监视器 (VGA 或 EGA 卡)；大中型数字化仪，彩色图形打印机，矢量绘图仪；Unix3.20 以上版本。

(一) 总体结构与技术特点

GRASS 采用模块化结构，由移植性较好的 Unix C 语言编写而成。系统由基本数据开发部分、影像处理模块和分析、叠加、显示部分组成。在基本数据开发部分中，图形数据在数字化命令支持下数字化，交互式编辑，自动生成拓扑关系，进入系统数据库。由矢量数据转换为栅格数据后经转化命令后进入系统栅格数据库。点数据、数字高程模型 (DEM)、数字化线划图 (DLG) 和 ASCII 文件数据等数据源也可转换成栅格数据。影像处理部分可由磁带或磁盘读入 TM、MSS 和 SPOT 卫星图像数据或航空像片数据等，经编辑、校正、增强、分类或判读后进入数据库。叠加、分析、显示部分包括逻辑推理，加权叠加和各种显示操作，如三维图像显示、彩色图表设计等。系统支持多种空间或属性动态的交互查询和信息提取，结果可直接输出，也可作进一步模型分析。分析与查询的结果由用户采用人机对话方式进行图符编辑整饰，由绘图机、打印机或屏幕拍照方式输出。

该系统的主要技术特点有：将 GIS 和图像处理相结合，两者在软件系统中运用相同的子程序，并且数据库可以相互转化，减少了更新 GIS 的中间环节，方便判读操作；采用矢量结构、栅格结构、点结构并存方式，点数据可转化成矢量图或栅格图，反之亦然，转换灵活；采用模块化设计，便于改进模块的功能。在 10 个基本模块中有上百个子程序；系统在运行过程中可依据所选定的分辨率重新采集数据，并输入到用户选择的区域中。而大多数其它系统在重新采集数据时，必须以原固定的栅格单元大小输入；可与多种软件接口及数据交换。数据可转化成多种格式，如 ARC/INFO 格式，DXF、DLG-3、ASCII 等。

(二) 数据结构

GRASS 采用矢量数据结构、栅格数据结构、点结构及影像像元结构并存，并且彼此可以相互转换，因此便于存储、管理不同类型的空间数据。属性数据库与图形数据库紧密联系，管理方便、灵活。图形符号库由底层和上层组成，底层由系统提供，包括各种基本线型要素；上层符号由用户在系统支持下灵活地建立，由底层的要素组合而成。

(三) 模块功能

GRASS 程序命令同 Unix 操作系统命令相接应，用户通过程序可知道操作什么样的数据类型。如表 5—2 所示为 GRASS 的程序结构。下面对主要功能模块作些介绍。

1. 输入编辑模块

该模块包括专题地图数字化输入编辑，遥感影像预处理，外部数据输入，点数据输入和属性数据库建立。

(1) 专题地图是最重要的数据源，利用以 *v.* 开头的程序命令完成专题地图的数字化、增删、修改、格式转换、开窗缩放、栅格图自动生成、坐标网格生成，长度、周长、面积等基本空间参数的量算等。

(2) 遥感影像是另一种重要的数据源，

表 5—2 GRASS 程序结构

命令前缀	命令描述
g	数据文件管理
d	图像显示
p	打印图像
r	栅格图操作和分析
v	矢量图操作和分析
s	点数据操作和分析
i	影像数据操作和分析
m	数据格式转换，外来数据输入

操作命令以 *i.* 开头，可对 SPOT、TM、MSS 遥感影像和航片进行预处理、校正、编辑、增强和信息复合。在专题信息支持下进行人机交互判读和分类，自动生成属性数据库、矢量和栅格数据库，实现 GIS 专题信息与遥感影像的覆盖、综合分析及管理。

(3) 在接受其它外部数据输入时，以 *m.* 命令提供对不同格式的数据进行采集。可接受的外部数据包括 ARC/INFO 数据格式、DLG 格式、DXF 格式、ASCII 格式、数字高程模型（DEM）数据格式。还可进行坐标系统转换，如经纬度坐标转换成 UTM 投影坐标等。

(4) 点数据输入，命令以 *s.* 开头。点数据可进行表面生成、统计比较，并转换成栅格和矢量图重新分类。

(5) 属性数据库管理系统采用 Unix C 语言编制，用以实现对属性数据的定义、修改、编辑和逻辑查询。

2. 模型分析模块

模型分析是 GIS 区别于其它信息系统的重要标志之一。在模型分析时可对数据库中的数据进行提取、筛选或重新组织，得到分析、决策信息。用户借助以 *r.* 开头的程序命令可进行多种运算、分析评价操作，如布尔（Boolean）叠加，加权叠加，Cellophane 叠加，聚类分析，近似分析，视野分析，网络分析，最短路径查询等。

3. 查询检索模块

该模块以图幅、属性、空间位置及空间关系为索引，对空间数据库进行检索，生成数据子集，实施模型分析和输出表示。其程序命令以 *g.* 开头。常用的查检内容包括两类：特定空间位置或空间范围的属性信息；特定属性组合的空间实体或空间分析。

4. 图形图像显示模块

程序命令以 *d.* 开头。该模块的功能包括矢量、栅格图显示、不同视角的三维立体显示，明暗模拟立体显示，地形信息提取，比例尺改变，图形颜色改变，直方图显示等。

5. 整饰输出模块

其程序命令以 *p.* 开头。该功能模块在系统的图形符号库支持下，通过对结果的符号编辑和图面整饰之后再绘图输出。还可打印制表。彩色图形打印机和矢量绘图仪分别支持栅格方式和矢量方式的数据输出。影像处理结果也能直接打印输出。

第 5—7 节　城市之星（Citystar）GIS 软件

城市之星（Citystar）GIS 软件是由北京大学遥感所、城市与环境学系与北京三秦信息技术有限公司联合研制的系统，它是原 Spaceman 系统的升级换代产品，属高档微机型系统。

（一）硬件配置和软件环境

主机是 486/586 微机，RAM 8M～128M，硬盘 340M～2G，外设包括 A_3～A_0 幅面的各种型号的数字化仪，A_4～A_0 幅面的各种型号的扫描仪，以及喷墨或笔式绘图仪，还可以扩展包括多媒体设备，各种扩充存贮设备。

系统的软件环境是：中文 Windows3.1，3.2 或西文 Windows3.1，95，Windows NT+汉化环境。

（二）主要软件配置和功能

该系统面向城市管理、地学及环境设计、管理。提供中文操作界面及应用界面辅助生

成系统，中国用户不需要软件编程，可直接建立空间模型和独立的管理系统，实现管理系统的二次开发。该系统同时具有栅格/矢量数据方式，可互相转换，可处理遥感图像，管理GPS数据。主要功能模块包括：

1. Citystar GIS 部分

Citystar-Vector（Edit）是矢量型GIS，主要功能有：矢量图输入、编辑、分层管理、注记，多媒体、遥感图像的连结，GPS定位数据更新，最佳路径选择。

Citystar-Scanner 扫描处理系统，具有图像图形扫描输入、网格图直接对像索引建库、矢量化与属性赋值等功能。

Citystar-DEM 数字高程模型系统，具有DEM插值、剖面线、坡度、坡向、阴影、汇流计算等功能。

Citystar-3D 三维立体系统，具有立体图制作，遥感影像三维显示，三维模拟及量测，基于3D的多媒体查询等功能。

Citystar-dbs 属性数据库查询管理系统，具有统计制图、制表等功能，可与通用商业数据库连接。

Citystar-Thematic Mapper（Mapper Ⅱ）专题图制图系统，有较齐全的制图功能。

2. Citystar RS 遥感图像处理系统

该系统的功能包括：图像几何纠正，配准、增强、漫游，多层图像运算，非监督分类与监督分类，与矢量数据叠加、统计，以及遥感分析结果多元信息表示等功能。

3. Citystar-GPS GPS 信息管理系统

该系统可通过多类微机接收GPS接收机的信号，并通过差分校正与投影变换将信息显示到电子地图上，具有动态跟踪、轨迹矢量化、目标点信号查询、搜索及量测等功能。

4. Citystar-Officer 办公管理系统

该系统是为办公管理无纸张化而设计的GIS工具。能在图形图像空间数据库的基础上建立文档管理系统，具有彩色地图制作，多媒体音像信息演示与发布等功能。

5. Citystar-3S 一体化部分

该部分通过用户界面自动生成，数据转换接口程序、多维信息分析等子系统将GIS、遥感分析系统与GPS系统综合起来，为资源与环境信息管理、城市规划、市政工程及社会经济管理提供基础工具。

第5—8节　吉奥之星（Geostar）GIS软件

吉奥之星（Geostar）GIS软件系统是由武汉测绘科技大学地理信息系统研究中心研制开发的，于1996年正式推出。该软件系统是基于矢量栅格一体化的数据结构和面向对象的数据模型，采用面向对象的空间数据库设计原理和面向对象的系统分析与设计方法，以C++作为程序设计语言。该系统有三种版本：在PC Windows支持下运行的单一个人机或网络多用户系统；其二是在微机Windows NT和工作站支持下运行的系统；以及在Unix操作系统和X-Window环境下的版本。

该系列软件包括面向对象的数据库管理系统Geo DB，地图扫描数字化识别软件GeoScan，数字高程模型及其应用软件Geo DEM。该系统各软件包之间的联系如图5—6所示。

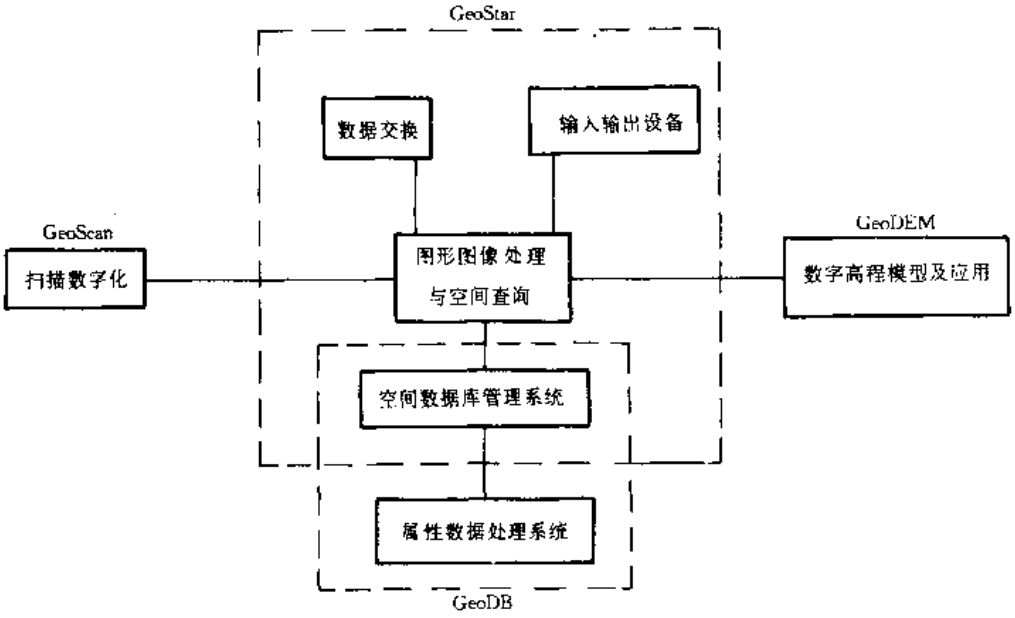


图 5-6 吉奥之星系列软件及其联系

吉奥之星具有图形数据与属性数据结合紧密，GIS 与遥感图像处理系统一体化性能较好的优点，体现了整体空间信息系统的特征。它的主要功能包括：

(1) 工程管理，如定义工程名和工程范围，打开老工作区或创建新工作区，定义地物类型，设计空间地物的属性数据结构，设置专题层，选择地物类型，工程交付等。

(2) 图形数字化。可数字化点、弧段、曲线及独立的点状地物，可作结点自动匹配、自动直接平差，可利用游标输入地物的编码或等高线的高程等。

(3) 图形编辑，包括增加、删除或移动结点、弧段，点状、线状和面状地物，注记等。

(4) 建立拓扑关系。通过一个开关设置能自动建立结点和多边形的拓扑关系，可以通过鼠标将若干条弧段装配成一个线状地物，将若干个简单地物装配成复杂地物。也可删除不正确的结点、线状地物、面状地物或复杂地物的拓扑关系。

(5) 属性数据输入与查询。除了在 Geo DB 中有属性数据的结构设计、编辑、查询与报表输出功能之外，在 Geostar 中也设有属性数据的处理功能，如在图形状态下直接输入属性数据，在建立图形的拓扑关系时即时输入属性数据，在图形屏幕下直接编辑、修改或查询属性，或者根据给定的属性条件查询相应的图形，实现图形与属性的无缝结合。

(6) 符号设计。系统提供了国内基本图图例的符号，可以满足一般需要。当需要添加符号时可选择“符号设计”菜单，进行符号的修改、设计，组装，以产生新的符号。

(7) 地图制图，包括图幅的定义，图框设计，图名，图幅说明，注记，图幅内容的选择，图例设计，地图文件生成，绘图仪的安装与设置，地图的输出。

(8) 图像处理与遥感分类，包括图像的纠正、滤波、增强、变换、边缘提取等。在图像处理基础上作遥感分类，如监督分类、非监督分类或人机交互分类。之后，自动建立栅格矢量化与拓扑关系。并可输入属性或产生专题图。

(9) 图形显示，包括图形的放大、缩小、漫游、开窗放大，地物的分层显示，各种要

素的选择显示，坐标显示等。

(10) 空间分析，包括在栅格结构下进行叠置分析，矢量拓扑结构下进行最佳路径分析。在栅矢结合条件下的缓冲区分析，地形分析，空间分析等。

(11) 数据交换功能。Geostar 既能接受多种设备采集的原始数据，如解析测图仪，全站仪，DXF 文件，扫描数字化处理的矢量文件，又可以与一些知名 GIS 软件系统的数据库直接交换，如与 ARC/INFO、Genamap、System 9、MGE、GRASS 等系统。

(12) 输入输出设备的安装与设置。

第六章　空间数据的处理与管理

空间数据有多种来源，不同的数据源其输入方法不同。不论采用什么方法输入数据都会有一些问题，如输入过程中意外的错误，输入数据与使用格式不一致，各种来源数据的比例尺、投影不统一，图幅间不匹配等。因此，必须对空间数据进行处理和管理，才能得到纯净统一的数据文件，使存储的空间数据符合规范、标准，满足使用和分析的需要。

第 6—1 节　空间数据的输入与编辑

一、图形数据的输入

图形数据的输入过程实际上是图形数字化处理过程。对于不同来源的空间数据，很难找到一种统一而简单的输入方法，只能从几种普遍适合的方法中选用。

（一）手工键盘输入

1. 手工键盘输入矢量数据

手工输入矢量图形数据就是把点、线、面实体的地理位置（坐标），通过键盘输入到数据文件或程序中去。实体坐标可从地图上的坐标网或其它覆盖的透明网格上量取。

2. 手工键盘输入栅格数据

栅格图形数据是以一系列像元表示点、线、面实体。这种数据的手工输入过程为：首先选择适当的像元大小和形状（一般为正方形网格）并绘制透明格网；然后确定地物的分类标准，划分并确定每一类别的编码；最后将透明格网覆盖在待输入图件上，依格网的行、列顺序用键盘输入每个像元的属性值即各类别的编码值。

手工键盘输入方法简单，不用任何特殊设备，但输入效率低，需要做十分繁琐的坐标取点或编码工作。这种方法在缺少资金或输入图形要素不复杂时可以应用。

（二）手扶跟踪数字化仪输入

这是目前最常用的图形数据输入方式。

1. 手扶跟踪数字化仪简介

这类数字化仪从工作原理和设计结构上属直角坐标式数字化仪。根据数据采集方式可分为机械式、超声波式和全电子式三种，目前应用得最多的是全电子式数字化仪（图 1—4）。

数字化作业时，随着鼠标器在图形输入板上位置的变化，输入交流信号的线圈由于电磁感应产生电场，并引起输入板内正交线栅网导线相应位置上电场的变化。当数字化点位的坐标时，只要把鼠标器十字丝精确放在待输入点上，并按压相应的按钮，便可记录该点应有的电信号。此信号通过设备的自动转换，就可得到图形输入板上的物理坐标（X，Y）值，该坐标值通过接口装置以实现计算机与数字化仪之间的数据通讯和控制功能。

2. 数字化过程

把待数字化的资料——地图、航片等固定在图形输入板上，首先用鼠标器输入至少四

个控制点的坐标和图幅范围，随后即可输入图幅内各点、曲线的坐标。

（三）自动扫描输入

自动扫描输入方式输入速度快，不受人为因素的影响，操作简单；缺点是硬件设备昂贵．图形识别技术尚不完全成熟。这种方法是图形自动输入的发展方向。有关扫描仪的工作原理、种类和精度见第1—3节。

（四）解析测图仪法空间数据输入

解析测图仪利用航空或航天影像像对，建立空间立体模型，直接测得地面三维坐标（X，Y，Z），并直接输入计算机，形成空间数据库。这种仪器不仅能记录三维坐标，还能通过联网的计算机处理比例尺变形和其它制图变形。

这种输入方法在建立数字地面模型（DTM）时尤其重要。

（五）已有数字形式空间数据的输入

这种方式用来接收已是数字化形式的信息数据，包括机载或星载多光谱扫描磁带等系统引进的信息。这些外来信息数据虽然已是数字格式，但其格式不一定与GIS数据库一致，还需作某些必要的预处理，如调整影像分辨率和象元形状、地图投影变换、数据记录格式转换等。对于星载多光谱图像，由于其空间分辨率较低，存在着各类边界的匹配和定向问题。虽然可以从卫星图像上推知它的定向和定位元素，但因像元过于粗略而不能精确定位。另外，多光谱扫描图像数据量太大应进行合理的压缩简化处理，例如把几个波段合成或其它基本变换使多波段数据组合在一起，进行专题要素的分类，把分类结果输入到空间数据库。

二、属性数据的输入

属性数据即空间实体的特征数据。例如，土地产权图斑既包括一组连续的像素或矢量表示的面实体，还包括属性数据如土地使用者、地块面积、产权沿革、土地使用类型等。显然，这些属性都与该面域空间实体相关，也应有效地存储和处理。一般，对一个空间实体及其属性赋予一个或多个关键字进行连接。

属性数据一般采用键盘输入。当数据量较大时，属性数据与空间数据分别输入并分别存储。把属性数据首先输入一个顺序文件，经编辑、检查无误后转存到数据库的相应文件或表格中。这是属性数据输入的一种常用方法。

三、GIS数据的编辑

数据编辑的目的在于消除数据输入过程中引入的错误或误差。它们包括空间点位不正确，变形，空间点位和线段的丢失或重复，线段过长或过短，面域不封闭，区域中心识别码的遗漏，结点代码和区域码不能符合拓扑一致性，属性的错误分类，错误编码或误输入等等。

D. F. Marble教授设计了一种层次与模块式相结合的数据编辑方法。图6—1所示为这种数据编辑方法的流程。

在GIS数据编辑前，应进行多工序的数据检核。检核方法有：

1）目标检核　将图形实体显示在屏幕上，检查一些明显的错误，如丢失了线段、图斑不闭合、线段过长等。

2）机器检核　这种检核主要是对数字化数据的拓扑一致性进行逻辑检核，把弧段连接成多边形以进行数字化限差的检查等。

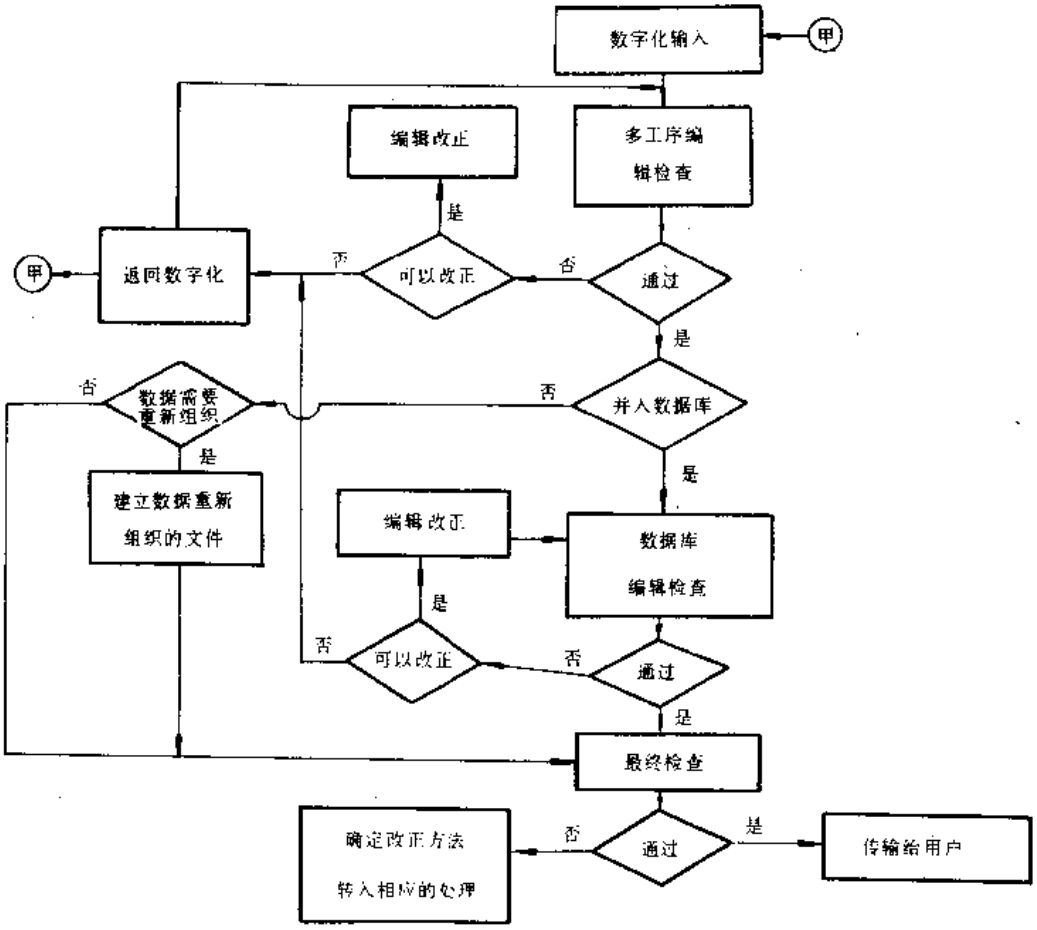


图 6—1　GIS 空间数据编辑过程

3）图形叠合比较法　即按与原图相同的比例尺用数据输出模块把输入的图形及其相应的属性绘到透明材料上，然后与原图精确套叠，在透光桌上仔细地观察和比较，查找遗漏、位置错误等，并做好相应的标记。

4）属性数据检核　属性数据的检核方法很多，常用且简单的方法是用打印机输出属性文件，逐行检核。另一种方法是编制检核程序，用程序扫描数据文件，看有无文字代替了数字或数字超过了允许范围等粗差，该程序还应有标出误差或者错误的能力。

GIS 中数据的编辑是一种非常耗时、耗精力的交互处理工作。它所花费的时间和精力甚至比数据输入更多。除逐一修改图形与属性输入的误差和错误外，图形的分割与合并，数据更新等许多工作也是在数据编辑模块中完成的。全部编辑工作是把数据显示在屏幕上，由键盘和编辑菜单来控制数据编辑的各种操作活动得以实现。

修改栅格数据中像元的属性值和空间位置差错时，通常用命令调出有差错像元的行列号，然后按键输入新像元值来取代有误的像元值。若差错的像元数太多，则宁可全部重新数字化，而把以前数字化的数据覆盖掉。

矢量数据库中的错误修改，可以在数字化仪上用鼠标器指向有错误的点、线，然后重新输入；也可以用计算机编辑命令来删除、插入、伸长、截取、移动、旋转图形实体等，使其达到正确校正的目的。矢量编辑不能孤立地进行，应当同时检查、编辑和修改相应的属

性，以保证空间实体及其属性数据的一致性。

表 6—1 表示一种空间数据编辑系统的功能。

表 6—1 GIS 数据编辑系统功能汇集

编辑				图形编辑	建立拓扑关系	属性编辑
地物类型				删除元素：点 / 一段弧 / 一条弧	创建线状目标	输入属性
图形编辑					人工建立多边形	属性人工连接
建立拓扑关系					自动建立多边形	切断属性联接
属性编辑				删除实体：点 / 线 / 面	拓扑关系	属性自动联接
创建点目标					装配复杂地物	删除属性
创建直线				修改图形	删除一类地物	修改属性
创建折线				拷贝：一条弧 / 一个目标		拷贝属性
创建光滑曲线						
创建多边形				移动：一条弧 / 一个目标		
创建椭圆						
创建圆				旋转：一个目标 / 一条弧 / 窗口内内容		
创建圆弧						
创建矩形						
注记				平行线		
显示所有类	漫游	开窗切割	关闭窗口	镜面反射		
整体放大	整体缩小	窗口放大	窗口缩小	联接元素		
				断开联接		

第 6—2 节　遥感数据与 GIS 的综合

GIS 是一种管理与分析空间数据的有效工具，而遥感是一种对空间数据采集和分类的有效工具，它们的工作对象都是空间实体，彼此之间存在着密切的联系。

(1) 遥感是 GIS 的一个重要数据源和强有力的数据更新手段。遥感使 GIS 的活力和应用大大加强，可作为 GIS 的数据采集子系统，为它提供及时、准确、综合和大范围的各种资源与环境数据；同时，遥感所具有的动态多时相特点，又为 GIS 空间数据库的更新或建立动态系统提供动态数据源。

(2)GIS 可为遥感提供各种有用的辅助信息和分析手段。GIS 有助于提高遥感数据自动分类的精度，为遥感应用分析提供一个良好的环境，从而大大提高遥感信息识别的精度和效率。它不仅是图像处理系统的合理“延长”，而且使遥感应用的深度和广度达到一个新的水平。GIS 的多数据源可使遥感数据的分类精度大大提高。目前，GIS 的一个重要发展趋势是加强空间信息管理系统与遥感图像处理系统之间的联系或结合，以提高资源与环境信息系统在动态分析、监测与预报方面的能力，并改善遥感数据分析的精度。

一、GIS 与遥感图像处理系统的结合

GIS 与遥感的结合，可以采取下列三种可能的技术途径：

第一种途径是发展一个能够综合处理矢量数据和栅格数据（遥感图像属栅格数据格式）的新的 GIS，这就要求有更有效的数据模型和空间数据的管理系统，从而使遥感数据能

够纳入到该系统中去。这时，地理信息系统除本身应有的数据处理功能外，还应配置必要的图像处理功能。如美国喷气推进实验室（JPL）的VICAR/IBIS地理信息系统，兼有遥感影像和GIS双重功能。该系统把各种数据源分成三种类型：表格数据（table），矢量数据（Vector），图像数据（Image）。它以影像为基础，把各种图形数字化、网格化，按它们的行号、列号进行行编码，使所有数据叠加在一起，归入统一坐标，建立起栅格结构数据库，如图6—2所示。

第二种途径是发展一种标准的空间数据交换格式，作为GIS与遥感图像处理系统，以

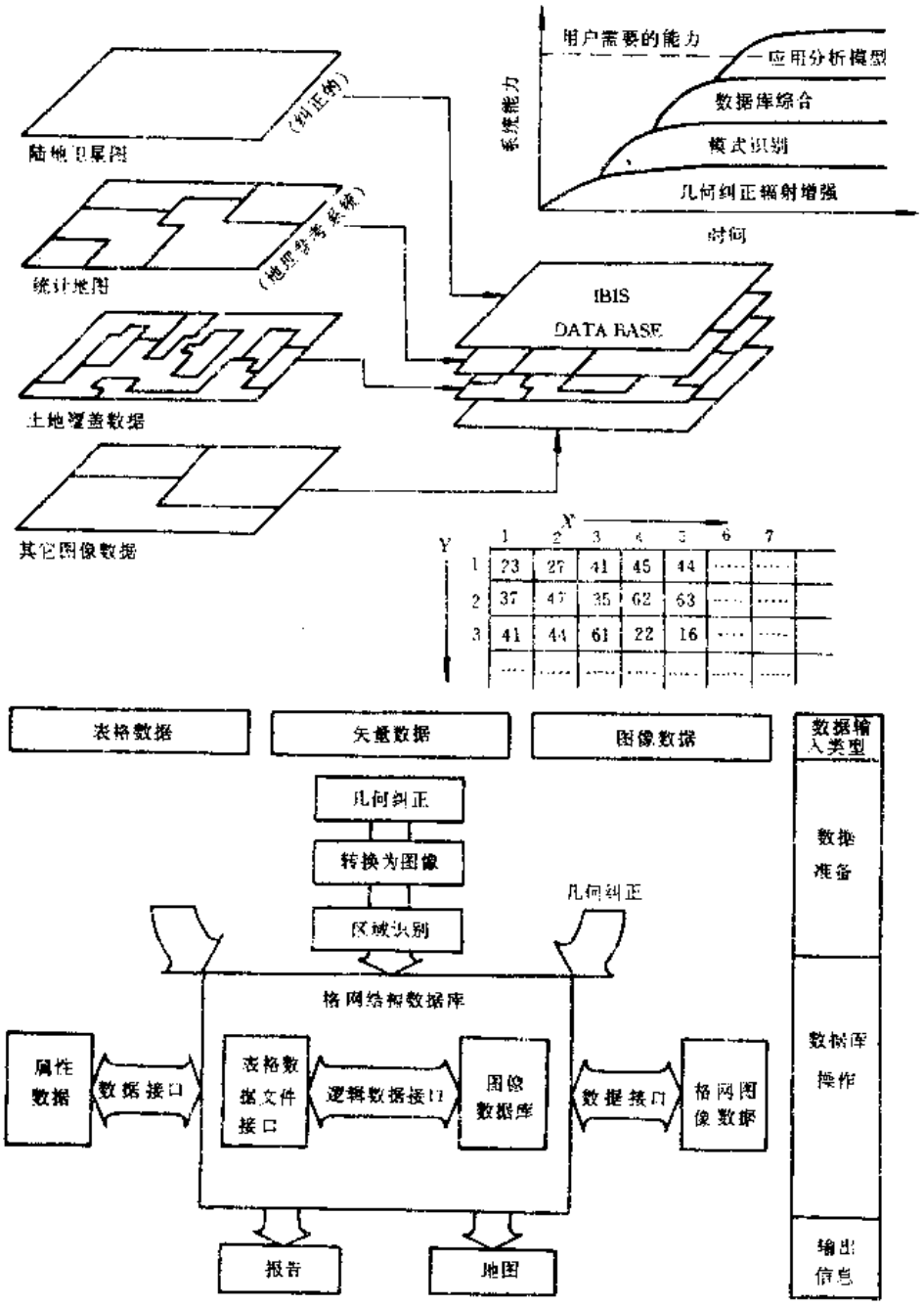


图6—2 VICAR/IBIS地理信息系统

及不同类型的 GIS 之间相互转换的中间格式。70 年代末，加拿大空间数据转换委员会（SDTC）设计的磁带格式即属此类。空间数据的标准格式一直被公认为是必需的，从长远来看，这有利于环境信息的共享，对全球气候变化的监测十分有用。但要设计这样一种能被普遍接受的标准是很困难的。

第三种途径是在现有的 GIS 和遥感图像处理系统之间建立接口，以实现数据格式的转换和数据的传送，这是目前比较现实且经济可行的一种途径（图 6—3）。

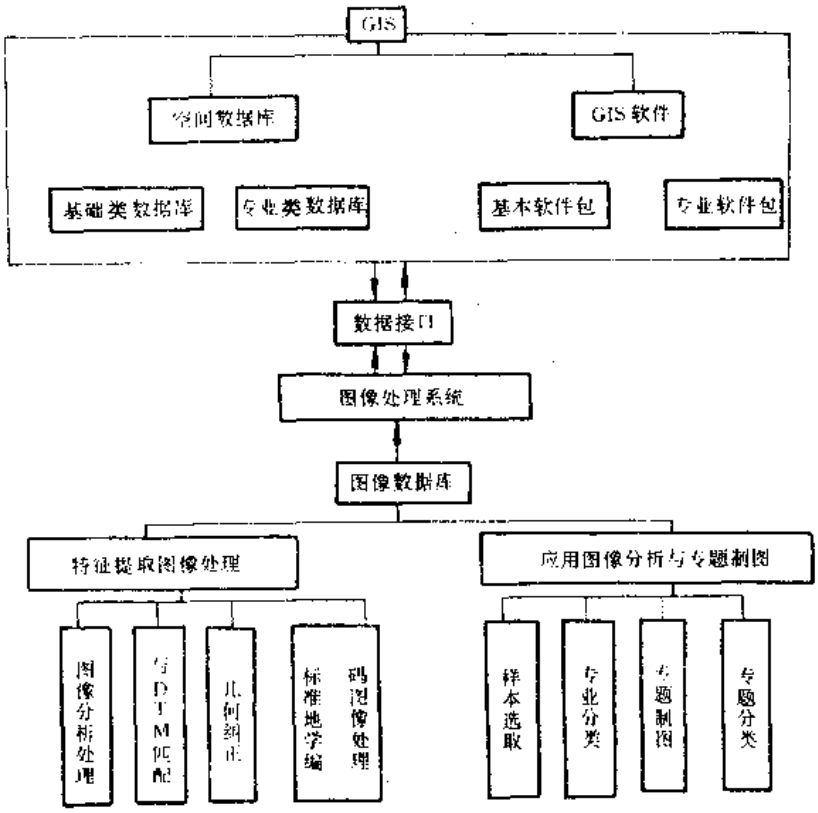


图 6　3　GIS 数据库的管理与更新

遥感与 GIS 技术是地理信息产业的两大支柱，应保持密切的关系。不论采用什么途径，首先应保证数据共享和标准化，以实现地学的动态分析、决策等功能。从地理信息学（Geomatics）的角度考虑，信息的流程可用图 6—4 来描述。

现代 GIS 的设计日益表现出明显的综合性特征。信息系统的综合性包括系统数据的完整性，系统的遥感影像数据与地图矢量数据的结合（称为综合地理信息系统 IGIS）以及系统使用的矢量数据与栅格数据的相互兼容。综合地理信息系统一般以功能较强的图像系统为基础，通过对遥感影像进行专题分类，然后对分类后的栅格影像（像元）进行矢量化，最后以弧段的格式存入数据库，以便与其它来源的数据进行匹配，提供分析使用。

二、遥感图像信息的解译

遥感图像的数据量十分庞大，例如 Landsat 卫星的 TM 图像，一幅图像覆盖 185km×185km 的地面面积，像元分辨率为 30m，共 75 个波段，每个像元的光谱分辨率为 2^8bit，则一幅图像数据量达 230MB 字节数。如此庞大的数据量若都存储于系统中，则系统无法正常工作。因此，对图像数据应进行适当的处理，主要是专题信息提取。

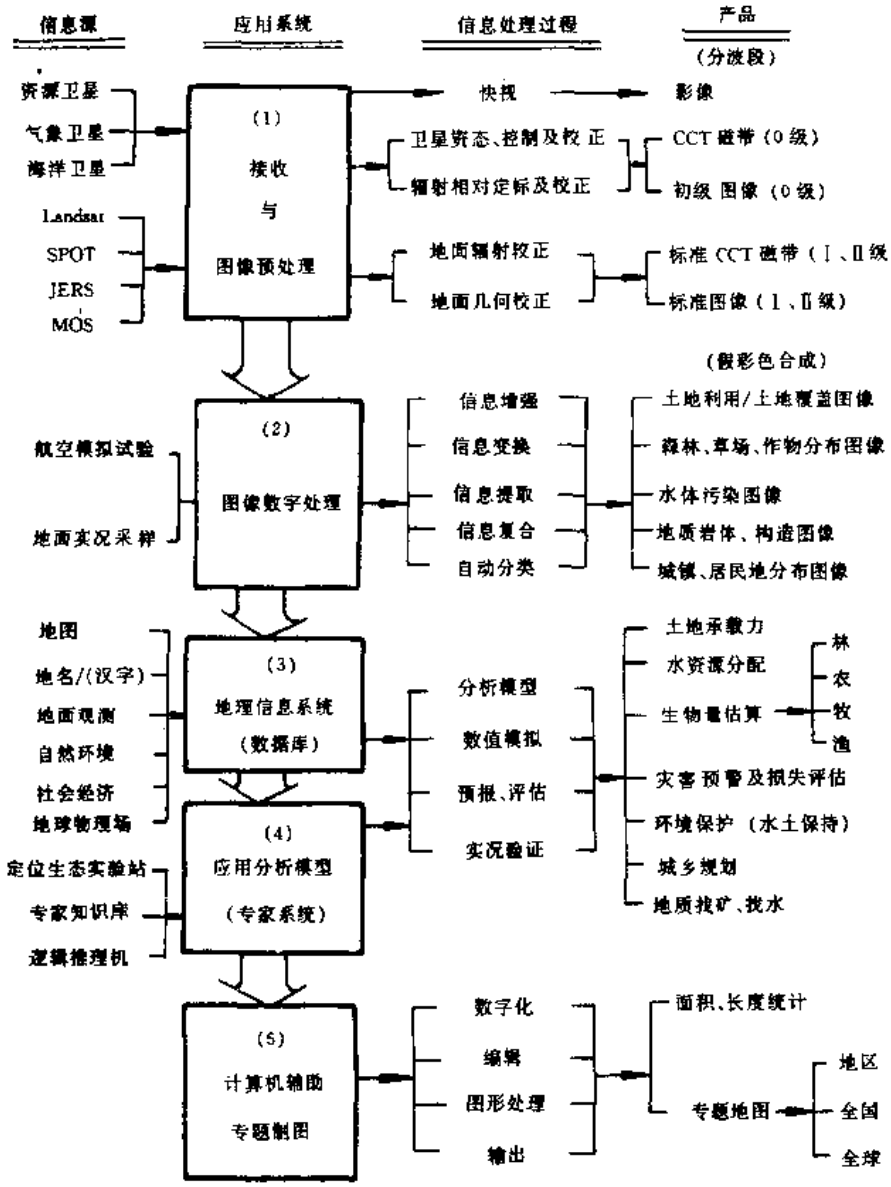


图 6—4　地理信息流程图

遥感图像专题信息提取的方法是进行图像的自动分类或目视解译。为了获得清晰且容易分类的图像，在分类前可进行适当的图像增强。

（一）遥感图像的数字增强处理

1. 密度分割

遥感图像数据所表示的灰度值是采样和量化的灰阶（如 TM 图像为 256 级），把灰阶顺序归纳成几个级别，并用不同的彩色表示，则得到密度分割图像。

2. 反差增强

遥感探测器的光谱灵敏度范围很宽，但因受大气的干扰，使图像上所反映的亮度集中在某个区段。为了充分利用图像的全部亮度范围，常采用反差增强的方法来改善地物的反差，以识别地物。

最简单的方法是线性反差扩展。每个像元的亮度值按线性内插法，使之分布在整个亮

度区间，线性扩展公式是

$$K(I,J)=\frac{k(I,J)-a}{b-a}(B-A)+A \qquad (6—1)$$

式中 I、J——像元所在的行、列号；

a、b——原始图像的最低和最高亮度值；

A、B——扩展后图像的最低和最高亮度值；

$K(I,J)$、$k(I,J)$——扩展后和扩展前的像元亮度值。

此外，还有其它反差增强法，如直方图均衡化、指数扩展、对数扩展、任意曲线扩展、分段线性扩展等。

3. 边缘增强

在那些图像像元之间的亮度值变化超过预先设定的阈值处，用线划或一种亮度在图像上表示出来，得到明显的边界线。

4. 滤波

一种方法是用空间模板在图像上移动，对像元进行卷积运算，适当选择模板的大小和参数，可使原图像的亮度值平滑或锐化；另一种方法是采用数字滤波器截止某种频率的波，让其它频率的波通过，从而使某类目标得到增强，而其它目标被抑制。滤波还可有方向性滤波。

5. 图像运算

图像处理系统可对多波段影像数据进行加、减、乘、除运算，以突出某种信息。例如作（MSS7—MSS5）/（MSS7+MSS5）运算得到的影像对识别作物的长势更为有效，而不同相的遥感图像的差值运算，对动态研究十分有效。

（二）图像识别与分类

图像识别与分类的目的是提取专题要素，参与分类的是多波段图像。基本的分类方法有监督分类和非监督分类两种。

监督分类法是选择有代表性的场地，以其中地物的光谱特性作为分类标准并输入计算机，尔后计算机检测待分类图像上的每个像元，根据最大相似或最小距离判别原则分别归入预先设立的类别。

非监督分类法不需要已知样本，仅按各像元亮度的相似性合并或分割成若干集群，从而达到分类的目的。所分的每一类别需经过实地调查或与已知类型的数据比较后才能确定。

（三）遥感图像的目视解译

图像的目视解译是在研究地物性质、电磁波特性及影像特征三者关系的基础上，凭借人眼及一定的工具，从影像的几何特征和光谱特征判断地物的电磁波性质，从而确定地物的属性，即从影像特征来识别地物。

三、遥感图像信息与其它空间信息的综合

各种单一遥感手段获取的图像，在几何、光谱和时间分辨率方面都存在着局限性。如果通过匹配处理将不同遥感手段得到的图像综合起来分析，并加入其它空间信息，则有利于专题信息分类和定量评价。除遥感图像本身的多光谱信息复合、多时相遥感信息复合、多平台遥感信息复合外，在空间多元信息分析中，经常使遥感图像信息与其它空间信息复合，实现多元信息的综合，以表达和揭示这些信息要素复杂的空间关系和形成机理。

为了实现多元信息的综合，需要采用统一的空间坐标和投影，才能得到定位配准的空间信息层。

1. 遥感图像与地图的综合

这种综合的产物是影像地图，它具有一定的数学基础，有丰富的影像信息和一定的地理要素，使得地图形象生动、正确详细，而且有利于更新地图。

2. 遥感数据与 DTM 的综合

DTM 与遥感图像的综合，可用 DTM 来校正遥感图像中因地形起伏所造成的像元位移，提高遥感图像的定位精度，同时 DTM 可参与遥感图像的分类，改善分类精度。

3. 遥感与地质、地球物理、地球化学信息的综合

地质、地球物理、地球化学信息反映了不同深度地壳物质的地球物理特征和化学性质。这些特征和性质在遥感图像上反映不出或十分微弱。对这些地学数据与遥感信息进行空间匹配复合、相关研究、综合分析，不仅可以清晰直观地反映出各种数据间的相互关系和各种地球物理场和地球化学场的空间分布，研究自然作用的内在规律，揭示图像的成像机理，而且可以准确而有效地提取地质信息，有助于地质构造分析，成矿预测和水文勘探等（图 6—5）。例如我国在安徽铜陵地区应用卫星图像、地质、断裂、矿产、航磁和航空放射性数据进行配准叠合分析，取得了成矿的构造地质条件分析及成矿预测成果。

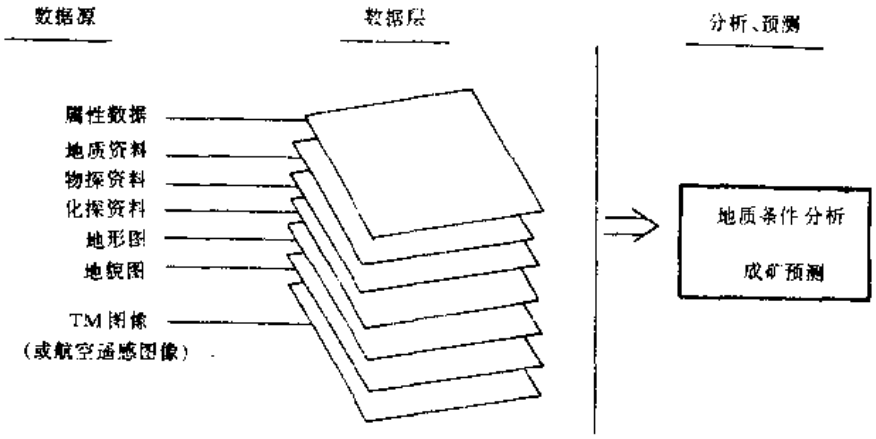


图 6—5　遥感与多源信息综合及分析

第 6—3 节　空间数据管理

空间数据经编辑处理后，一般只能消除或减少在数字化过程中因操作产生的局部误差或明显错误，还可能存在着：各种数据来源的比例尺和投影不一致，各幅地图数据之间的不匹配，数据的冗余度等问题，必须作进一步的预处理——数据管理过程，才能得到可以满足用户要求的数据文件。空间数据管理包括：图幅数据处理，数据格式转换，空间数据的内插、定位或定性检索，数据之间的边缘匹配，数据更新等。

一、图幅数据的处理

地理信息系统的数据来源十分广泛，包括遥感数据，各种实地测量数据和统计资料。目前，各种地图的数字化数据，特别是手扶跟踪的数字化数据，仍是地理信息系统输入数据的主要来源。由于数字化仪桌面坐标系与地图坐标系不一致，以及地图图纸的伸缩变形等

原因，数字化仪读出的坐标（x'，y'）可能与地图的直角坐标（x，y）不一致。有时，不同来源的地图还存在地图投影与地理坐标的差异，以及地图比例尺之间、地图比例尺与数字化仪的长度单位之间的种种不一致。因此，为了获得与地图直角坐标系一致的数据，使不同地图的投影坐标取得统一，必须进行图幅数据的处理。可见图幅数据的处理就是图幅内 x，y 坐标数据的变换，它包括数字化坐标数据比例尺的变换，变形误差的消除，坐标的旋转和平移，以及投影类型的转换（图 6—6）。

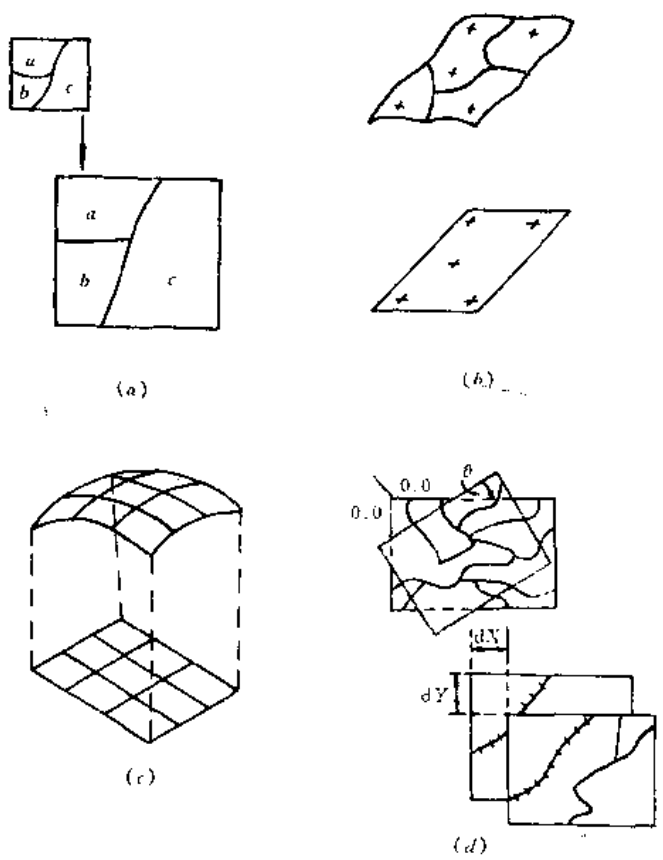


图 6-6　图幅数据的处理

a—比例尺变换；*b*—变形误差消除；*c*—投影类型转换；*d*—坐标旋转和平移

1. 数字化仪坐标的变换

如图 6—7 所示：设 xoy 为地图直角坐标系，$x'o'y'$ 为数字化仪坐标系，两坐标系夹角为 θ，o' 相对于 o 点平移了 a、b，地图的实际比例尺横向、纵向分别为 Q_1、Q_2，则根据图形变换原理，可得出变换式

$$\begin{cases} x=Q_1\ (x'\cos\theta-y'\sin\theta)\ +a \\ y=Q_2\ (x'\sin\theta+y'\cos\theta)\ +b \end{cases} \tag{6—2}$$

令 $A_1=Q_1\cos\theta$，$A_2=-Q_1\sin\theta$，$B_1=Q_2\sin\theta$，$B_2=Q_2\cos\theta$，则上式可以简化为

$$\begin{cases} x=A_1x'+A_2y'+a \\ y=B_1x'+B_2y'+b \end{cases} \tag{6—3}$$

式（6—3）含有 6 个参数 A_1、A_2、a、B_1、B_2、b，要建立坐标变换关系式，须建立 6 个方程，即需在地图上数字化 3 对坐标已知的点，但由于数字化过程中存在着误差，实际数字化的点应大于 4 个。下面按最小二乘法原理，求解转换参数，并评定转换精度。

设 V_x、V_y 表示转换坐标与理论坐标之差，则有

$$\left.\begin{aligned} V_x&=x-\ (A_1x'+A_2y'+a) \\ V_y&=y-\ (B_1x'+B_2y'+b) \end{aligned}\right\} \tag{6—4}$$

按 $V_x^2+V_y^2=\min$ 原理，可建立法方程

$$\left.\begin{aligned} &a\cdot n+A_1\Sigma x'+A_2\Sigma y'=\Sigma x \\ &a\cdot\Sigma x'+A_1\Sigma\ (x')^2+A_2\Sigma x'y'=\Sigma x'x \\ &a\cdot\Sigma y'+A_1\Sigma x'y'+A_2\Sigma\ (y')^2=\Sigma y'x \\ &b\cdot n+B_1\Sigma x'+B_2\Sigma y'=\Sigma y \\ &b\cdot\Sigma x'+B_1\Sigma\ (x')^2+B_2\Sigma x'y'=\Sigma x'y \\ &b\cdot\Sigma y'+B_1\Sigma x'y'+B_2\Sigma\ (y')^2=\Sigma y'y \end{aligned}\right\} \tag{6—5}$$

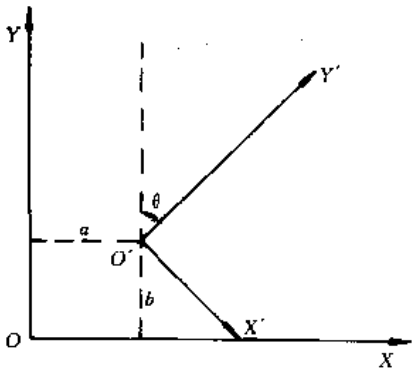


图 6—7　坐标变换原理

式中 n 为数字化已知坐标点的个数。解算该法方程，可求得各转换参数。转换精度以转换点位中误差表示，即

$$m=\sqrt{\frac{V_x^2+V_y^2}{2n-1}} \qquad (6—6)$$

在建立变换方程后，地图上各数字化点的坐标均用式（6—3）从数字化仪坐标转换为地图坐标。该公式考虑到 x、y 两个方向上因图纸变形而引起实际比例尺的不同，因此称为仿射变换公式。

2. 地图投影的变换

进行空间数据分析时，要求所有地理信息必须是同种投影。因此，当系统使用的数据取自不同地图投影的图幅时，需要将一种投影的数字化数据转换为所需要的投影的坐标数据。投影转换可由下列三种途径实现：

1）正解变换　通过建立一种投影变换为另一种投影的严密或近似的解析关系式，直接由一种投影的数字化坐标（x，y）变换为另一种投影的直角坐标 X、Y。

2）反解变换　即由一种投影的坐标反解出地理坐标：$B=F_1(x, y)$，$L=F_2(x, y)$，然后再把地理坐标代入另一种投影的坐标公式中 $X=T_1(B, L)$，$Y=T_2(B, L)$，从而实现两种投影坐标间的变换（x，$y\rightarrow X$，Y）。例如：研究区域恰好横跨两个高斯-克吕格投影带，则应将两个投影带坐标统一到同一个投影带上才能实现图幅的拼接，这时就需用反解变换法。

3）数值变换　根据两种投影在变换区内的若干同名数字化点，采用插值法，或有限差分法、有限元法、待定系数最小二乘法，实现两种投影坐标之间的变换。这种变换公式为

$$X=\sum_{i=0}^{m}\sum_{j=0}^{m}a_{ij}x^iy^j$$

$$Y=\sum_{i=0}^{m}\sum_{j=0}^{m}b_{ij}x^iy^j \qquad (6—7)$$

如取 $m=3$ 时，有

$$X=a_{00}+a_{10}x+a_{01}y+a_{20}x^2+a_{11}xy+a_{02}y^2+a_{30}x^3+a_{21}x^2y+a_{12}xy^2+a_{03}y^3$$

$$Y=b_{00}+b_{10}x+b_{01}y+b_{20}x^2+b_{11}xy+b_{02}y^2+b_{30}x^3+b_{21}x^2y+b_{12}xy^2+b_{03}y^3$$

为了解算以上三次多项式，需要在两投影间选定相应的 10 对点，其坐标分别为 x_i，y_i 和 X_i，Y_i。由此可组成线性方程组，解算该方程组，得系数 a_{ij}、b_{ij}，这样就可确定一个坐标变换方程，由该方程对其它待变换点进行坐标转换。

这种坐标转换法因此也称作待定系数法。

二、数据格式的转换

空间数据一般有栅格数据和矢量数据两种类型。它们的相互转换方法已在第二章中介绍。此外，还会遇到不同系统的不同数据格式，如 ARC/INFO 系统采用 ARC/INFO 格式，SICAD 系统采用 INFOMIX 格式。为了实现数据共享，要求 GIS 能把其它系统的格式转换为本系统适用的数据格式。转换的方式有两种：一是建立两种系统的数据格式转换模块，直接转换，这种方式要求知道其它系统所采用的数据格式，如 ARC/INFO 系统提供了 20 多

种数据格式相互转换的命令；二是把不同系统的图形文件和属性文件转换成 DXF（图形交换文件）和 DBF 格式，然后再由 DXF 和 DBF 格式转换成其它系统能识别的数据格式。其中 DXF 文件是标准的 ASCII 文本文件，很容易转换成其它图形文件格式。DXF 文件的一般文件结构如下：

（1）标题段。有关图形的一般信息可在 DXF 文件的此段中找出，每个参数具有一个变量名和一个相关值。

（2）表段。此段包含几种命令项的定义、线型（LINETYPE）表、层表、字体表和视图表。

（3）块（BLOCKS）段。此段含有块定义的实体，以描述组成图的每个块。

（4）实体（Entities）段。含有图形实体，供任何段调用。

（5）文件结束符（EOF）。

即使一个 DXF 文件缺少某些段、表，只要有实体段和结尾符号，则可向其它系统转换。

三、空间数据与属性数据的连接

空间数据输入时虽然可以直接在图形实体上附加一个特征编码或识别符，但是这种交互式的编辑方法需输入大量复杂的属性数据，工作效率低。空间和属性数据连接的较好方法是利用专用程序自动地把属性数据与空间实体数据连接起来，此时，只要求空间实体带有唯一的标识符即可。标识符可以用手工输入，也可以由程序自动生成并与图形实体的坐标存储在一起。

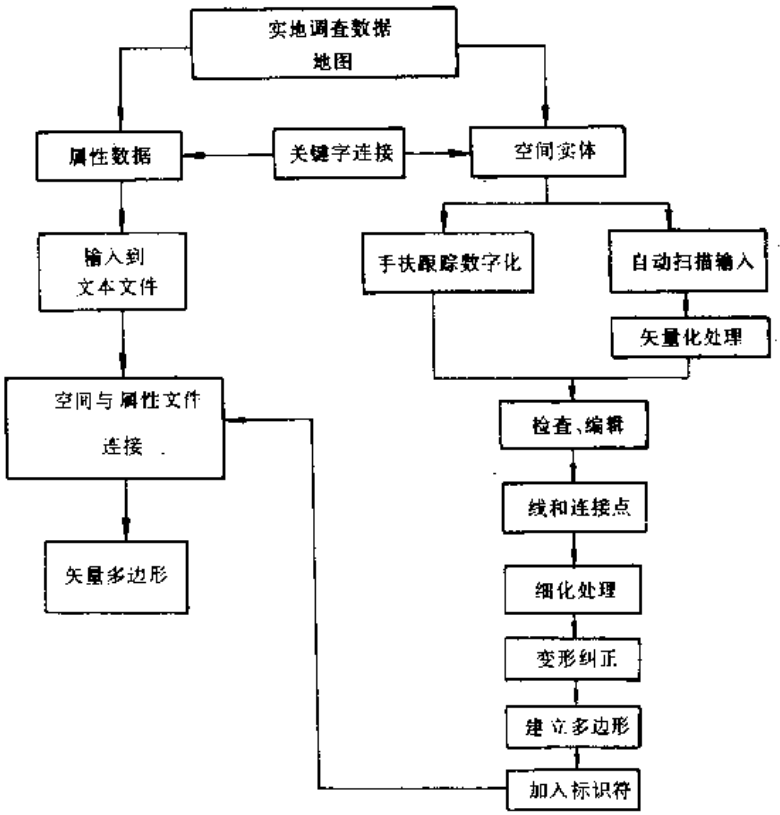
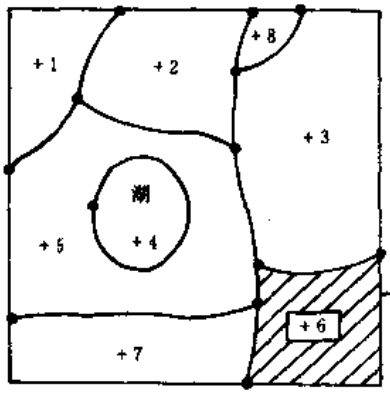


图 6—8　建立矢量多边形数据库的过程

空间实体的属性项目很多，把属于同一实体的所有数据项放在同一个记录中，记录的顺序号或某一个特征数据项作为该记录的标识符或关键字。它和图形的标识符都是空间和属性数据连接和检索的钮带。唯一性的标识符只有在特殊的计算系统中才能直接附加到图形实体中去。在多边形网络模型中，首先必须建立多边形，然后才能附加标识符。图 6—8 说明了按拓扑结构建立具有多边形边界的图形数据与属性数据完整的矢量多边形数据库的整个过程。

在 ARC/INFO 地理信息系统中，空间实体与属性的连结是通过机内代码和用户代码（USER-ID）来建立联系的，见图 6—9。



机内码	用户标识码	周 长	面 积	土壤类型	适应性
1	0	34.0	-71.6	—	—
2	1	9.0	8.0	113	HIGH
3	2	8.5	12.5	95	LOW
4	3	11.3	14.3	212	MIDDOW
5	4	4.5	4.3	201	MIDDOW
6	5	12.5	11.2	85	LOW
7	6	11.6	10.5	77	HIGH
8	7	8.5	8.7	117	LOW
9	8	2.7	2.1	514	MIDDOW

图 6—9　ARC/INFO 中多边形与属性的连接

四、空间数据开窗处理

开窗处理指用户以给定的空间范围（窗口），进行空间数据的提取，包括提取窗口内的空间实体及其属性。当只分析研究整个空间数据库的一部分，或对某一局部区域进行数据显示与转贮时，往往要进行开窗处理。开窗方式有正开窗和负开窗两种。正开窗是选取整个空间数据在窗口内的子集，而负开窗是选取窗口外空间数据的子集（图 6—10）。

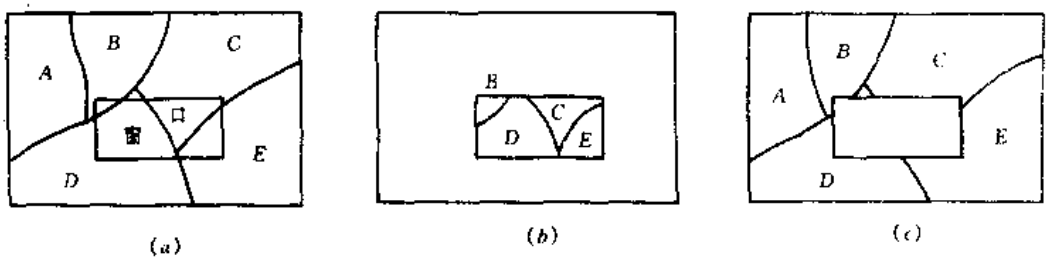


图 6—10　正开窗和负开窗示意图

a—原图；*b*—正开窗；*c*—负开窗

通常正开窗用得更多。窗口可以是矩形或其它任意多边形，由用户根据需要选定。窗口的范围即为轮廓点坐标，可由键盘输入或用光标在屏幕上确定。

窗口确定后，还要考虑如何切掉窗口以外（正开窗时）或窗口内（负开窗时）的空间实体，这一工作称为剪辑。现以矩形窗口的正开窗为例，讨论空间实体的剪切过程。

（1）对每一条需显示的线段都要进行坐标检验，判断其是完全在窗口内还是在窗口外，

或是与窗口相交。

(2) 根据实际情况分别处理。如果线段全部在窗口内部，则显示此线段；如果线段全部在窗口外，则不显示；若线条与窗口相交，则必须求出交点坐标，截断线条，仅显示窗口内的部分，如图 6—11 所示。

下面分别讨论三种基本空间要素：点、线、面的开窗剪切算法。

1. 点的开窗剪切

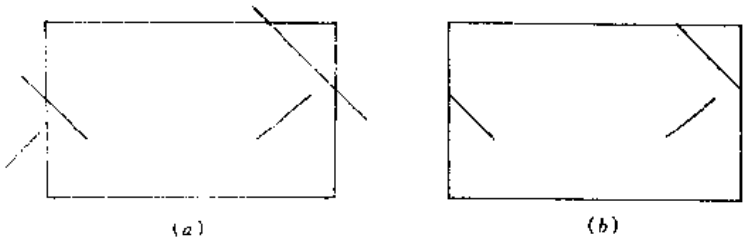


图 6—11 线段的正开窗处理

a—剪切前；b—剪切后

设窗口的左下角和右上角坐标为 (x_{min}, y_{min}) 和 (x_{max}, y_{max})，任意点 P 的坐标为 (x_p, y_t)，显然只要下式：

$$x_{min}<x_p<x_{max}, \text{ 且 } y_{min}<y_p<y_{max} \tag{6—8}$$

成立，则 P 点在窗口内，否则舍去。

2. 线状要素的开窗剪切

线状要素是由有序线段组成的折线逼近的，因此，对线状要素的开窗剪切，可简化为对线段的剪切。对线段的开窗剪切算法很多，在此仅介绍 Sutherland-Cohen 算法。

为了迅速判断线段是否完全可见或完全不可见，将空间依据窗口的方位划分为 9 个子区（图 6—12），并用一个 4 位码来表示线段端点位于哪个区中。编码顺序从右向左，4 位编码中每一位对应线段端点的定义是：第 1 位为“1”，表示点在左边框 $w_{x左}$ 的左边；第 2 位为“1”，表示点在右边框 $w_{x右}$ 的右边；第 3 位为“1”，表示点在下底边 $w_{y下}$ 的下边；第 4 位为“1”，表示点在上顶边 $w_{y上}$ 的上边；否则，相应位置上为“0”。

$w_{y上}$	1001	1000	1010
	0001	0000	0010
$w_{y下}$	0101	0100	0110
		$w_{x左}$	$w_{x右}$

图 6—12 空间被分成 9 个区，其中窗口所在区为 0000

由此可见：①如果线段的两个端点的 4 位码都是“0”，则两端点都在窗内，线段完全可见，接受此线段；②如果两端点码对应位的逻辑“与”不为“0”，则整条线位于窗口的同一侧，线段完全不可见，舍去该线段；③当两端点码不都是“0”，但各位的逻辑“与”都为“0”时，则线段可能部分可见，也可能完全不可见，此时，总存在着一个端点（设为 P_1）在窗口外，并由它的端点码知它处于哪一侧。进行该线段的剪切，则要进行线段与窗口边界交点的计算。

设线段的两个端点为 P_1 (x_1, y_1) 和 P_2 (x_2, y_2)，过此两点建立直线方程

$$y=\frac{y_2-y_1}{x_2-x_1}(x-x_1)+y_1 \tag{6—9}$$

令 $m=\frac{y_2-y_1}{x_2-x_1}$，该直线与窗口边界 4 条直线的交点坐标是

窗左边：$x_l=x_{min}$，$y_l=m\cdot(x_{min}-x_1)+y_1$　$m\neq\infty$

窗右边：$x_r=x_{max}$，$y_r=m\cdot(x_{max}-x_1)+y_1$　$m\neq\infty$

窗上边：$x_t=x_1+\frac{1}{m}(y_{max}-y_1)$，$y_t=y_{max}$　$m\neq0$

窗下边：$x_b=x_1+\frac{1}{m}(y_{min}-y_1)$，$y_b=y_{min}$　$m\neq0$

求出交点后，舍去 P_1 到该交点的部分线段。其它线段按上述方法逐一剪切，直到对剩余线段能利用端点码判断是全部接受或舍去。

3. 面状要素（多边形）的开窗剪切

多边形的边界是由若干条有序折线封闭而成的，因此，对多边形的剪切与对线段的处理基本相同，但剪切后，应把窗口边界上所有线段与被裁剪图形组成封闭多边形（图 6—13），而其属性可从原空间数据库中拷贝。

4. 任意形状窗口剪切

对于非矩形窗口，如圆形，小比例尺图上由曲线形经纬线所围成的区域，或由用户任意划定的不规则多边形，剪切的算法要复杂一些。

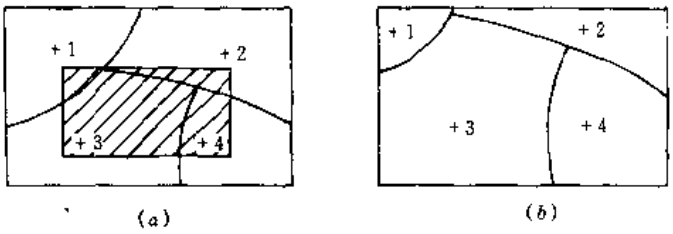


图 6—13　面状要素的开窗剪切
a—剪切前；*b*—剪切后窗口放大

在非矩形窗口的情况下，空间数据库中任何顺序上相继的两个点，相对于窗口的位置关系可归纳成以下 9 类（图 6—14）：

（1）均在窗口内且相连（如 A_1，A_2）；

（2）均在窗口内，但不相连（如 B_1，B_2）；

（3）由内到外且相连（如 C_1，C_2）；

（4）由外至内且相连（如 D_1，D_2）；

（5）先内后外不相连（如 E_1，E_2）；

（6）先外后内不相连（如 F_1，F_2）；

（7）均在窗外不相连（如 G_1，G_2）；

（8）均在窗外相连，但跨窗口边界（如 H_1，H_2）；

（9）均在窗外相连，但离窗口边界（如 I_1，I_2）。

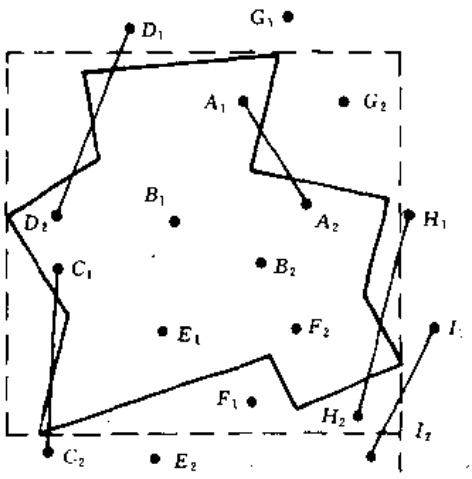


图 6—14　相继两点与窗口的各种关系

对于任何一个离散点可利用铅垂线内点匹配法判别该点是在多边形外还是在多边形内。铅垂线算法的思路是：从任一离散点向下引铅垂线，计算它与窗口边界各边的交点个数，若为偶数，则该点位于窗口内，否则在窗口外。这种方法基于拓扑学原理，即任一条直线与一个封闭多边形的交点个数为偶数。这种方法对一些特殊情况（如铅垂线恰好穿过多边形的某个顶点），应做特殊处理。

为了加快判断速度，可分两步进行。首先判断某点是否位于窗口多边形的最小外接矩形内，如果在其内再用铅垂线内点匹配法进一步判定是否在窗口内，如在图 6—14 中，D_1，G_1，H_1，I_1，I_2 均位于最小外接矩形之外，则不可能在窗口内。

对于任一条折线，要确定哪一部分在窗口内，应根据各线段两端点相对于窗口位置的情况来确定。通常从折线的始点开始，按顺序逐段判别各线段与窗口多边形各边有无交点。设其中一线段的端点为 A_1 和 A_2，多边形某一边的端点为 B_1，B_2，如果判别条件：$\max(x_{A_1}, x_{A_2}) < \min(x_{B_1}, x_{B_2})$ 或 $\min(x_{A_1}, x_{A_2}) > \max(x_{B_1}, x_{B_2})$ 或 $\max(y_{A_1}, y_{A_2}) < \min(y_{B_1}, y_{B_2})$ 或 $\min(y_{A_1}, y_{A_2}) > \max(y_{B_1}, y_{B_2})$ 成立，则这两条线段不可能有交点。对于每一线段来说，首先均需按此条件判断它与多边形各边有无交点的可能。若有，则计算交点坐标（在线段或窗口边上的交点除外），然后连同线段端点按 x（或 y）坐标依次排队（从小到大或从大到小），最后按下面两种情况配对作剪切选取：

(1) 若点 1 在窗内，可按$\overline{12}$、$\overline{34}$，$\overline{56}$…顺序配对，这里 1，2，3…为排队后交点的序号；

(2) 若点 1 在窗外，可按$\overline{23}$、$\overline{45}$，$\overline{67}$…顺序配对。

这种运用于非矩形窗口的算法也适用于矩形窗口。

五、空间数据的压缩处理

把数据组织到数据库中，是建立 GIS 的一个重要环节。利用现代的数据采集手段对空间要素进行量化，若以微米级的分辨率，将产生很大的坐标数据量。若不采用数据压缩技术，系统存储空间和处理时间上都会受到很大压力，甚至导致系统的失败。此外，当空间数据的比例尺变化时，也往往要求压缩数据。

所谓数据压缩，是从取得的数据集合 S 中抽出一个子集 A，这个子集作为一个新的信息源，在规定的精度范围内最好地逼近原集合，而又取得尽可能大的压缩比。

对于栅格数据结构方式存储的数据库，其压缩技术有游程编码、四叉树法等。这已在第二章中作了详细的论述。

在矢量数据结构中，数据压缩的主要对象是线性要素中心轴线和面状要素的边界数据，常用的数据压缩方法有：

1. 间隔取点法

设曲线由折线点序列 $\{P_1, P_2 \cdots P_R\}$ 构成，给定反映其坐标值的两数组 $(x_1, x_2 \cdots x_R)$ 和 $(y_1, y_2 \cdots y_R)$，则任意两点间的距离为

$$D=\sqrt{(x_M-x_N)^2+(y_M-y_N)^2} \qquad (6—10)$$

并给定临界距离 $D_{临}$。首先，保留曲线始点，然后计算 P_2 点与 P_1 点之间的距离 D_{21}，若 $D_{21} \geqslant D_{临}$，则保留第 2 点，否则舍去 P_2 点。

依此方法，逐一比较 P_3 与 P_2 点…以确定舍去那些离已选点比规定距离 $D_{临}$ 更近的点（图 6-15），但曲线的末尾点一定予以保留。

采用数字化仪时，这种方法可大量压缩取样点列中的点，但不一定能恰当地保留方向上曲率显著变化的点。

2. 垂距法和偏角法

这两种方法是按垂距或偏角的限差选取符合或超过限差的点。即利用曲线点序列中顺序的 3 点，P_{n-1}，P_n，P_{n+1}，把 P_{n-1} 与 P_{n+1} 点相连，计算 P_n 点到 $P_{n-1}P_{n+1}$ 线的垂距（垂距

法）或计算 $P_{n-1}P_n$ 与 $P_{n-1}P_{n+1}$ 直线的夹角（偏角法），并与规定的限差比较，以确定 P_n 点是否取舍，其过程见图 6—16。

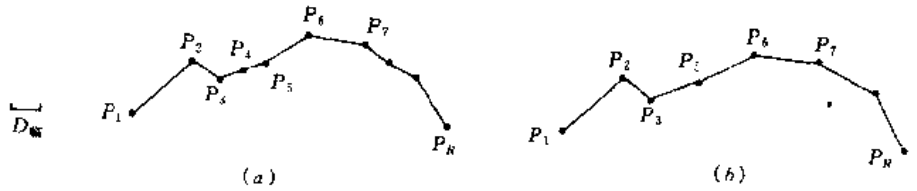


图 6—15　间隔取点法取点示意图
a—原曲线；*b*—经压缩后曲线

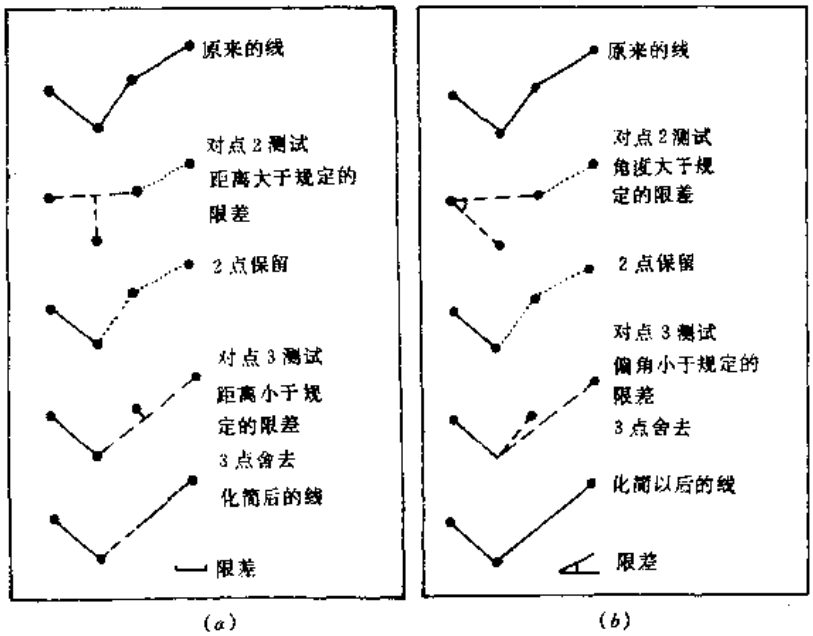


图 6—16　按垂距法和偏角法曲线压缩过程
a—垂距算法；*b*—偏角算法

这两种方法虽然不能同时考虑相邻点间的方向和距离，且有可能舍去不该舍去的点，但比前一种方法好。

3. Douglas-Peucker 方法

该方法试图保持曲线走向并允许用户规定合理的限差，其执行过程见图 6—17。

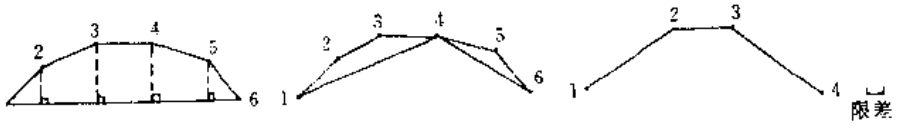


图 6—17　Douglas-Peucker 法过程示意图

首先，把一条曲线首末两点连成一条直线，其直线方程为

$$Ax+By+C=0 \qquad (6—11)$$

曲线上各点到该直线的距离为

$$d_i=|Ax_i+By_i+C| \tag{6—12}$$

选取距离中最大者 max (d_i) 与规定的限差比较，若大于限差，则离该直线距离最大的点保留，否则将直线两端点间各点全部舍去。显然图 6—17 中 4 号点应该保留。

然后，将已知点列分成两部分处理，计算 2、3 点到 1、4 点连线的距离，选距离大者与限差比较，结果 2、3 点均应舍去；计算 5 号点到 4、6 两点连线的距离，经比较，应保留。依次类推，最后保留的点在原数据库中的编号为 1、4、5、6 点，重新排序后得到经压缩的点列 1、2、3、4。

六、图幅接边与分割

在对空间数据管理时，按图幅存储的数据不能进行大区域的信息分析，需要把相邻图幅拼接成大型数据库；或是数据覆盖区域太大，不利于输出，又需要把大数据库分割成分析或输出的小区域。这就是图幅的接边与分割（图 6—18）。

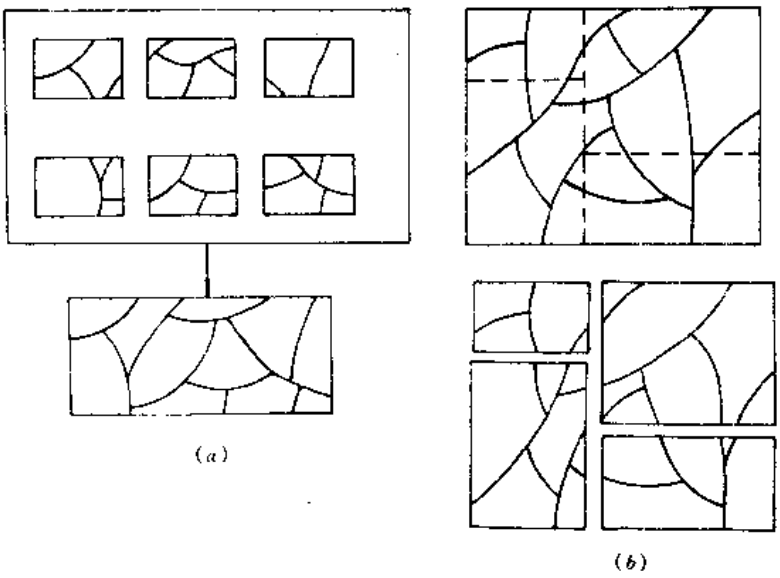


图 6- 18　图幅数据的拼接与分割示意图

a—图幅的拼接；*b*—图幅的分割

（一）图幅的拼接

图幅的拼接总是在相邻两图幅之间进行的。要将相邻两图幅之间的数据集中起来，就要求相同实体的线段或弧的坐标数据相互衔接，也要求同一实体的属性码相同，因此必须进行图幅数据边缘匹配处理。

1. 逻辑一致性的处理

由于人工操作的失误，两个相邻图幅的空间数据库在接合处可能出现逻辑裂隙，如一个多边形在一幅图层中具有属性 A，而在另一幅图层中属性为 B。此时，必须使用交互编辑的方法，使两相邻图斑的属性相同，取得逻辑一致性。

2. 识别和检索相邻图幅

将待拼接的图幅数据按图幅进行编号，编号有 2 位，其中十位数指示图幅的横向顺序，个位数指示纵向顺序（图 6—19），并记录图幅的长宽标准尺寸。因此，当进行横向图幅拼

接时，总是将十位数编号相同的图幅数据收集在一起；进行纵向图幅拼接时，是将个位数编号相同的图幅数据收集在一起。其次，图幅数据的边缘匹配处理主要是针对跨越相邻图幅的线段或弧而言的，为了减少数据容量，提高处理速度，一般只提取图幅边界 2cm 范围内的数据作为匹配和处理的目标。同时要求，图幅内空间实体的坐标数据已经进行过投影转换。

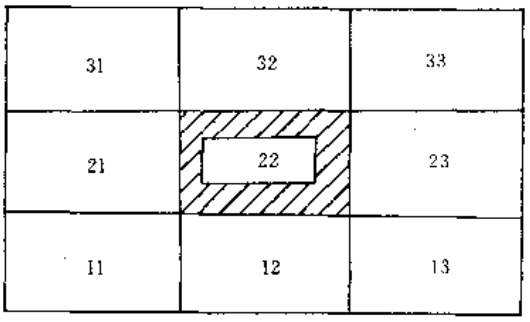


图 6—19　图幅编号及图幅边缘数据提取范围

3. 相邻图幅边界点坐标数据的匹配

相邻图幅边界点坐标数据的匹配采用追踪拼接法。追踪拼接有四种情况（图 6—20），只要符合下列条件，两条线段或弧即可匹配衔接：相邻图幅边界两条线段或弧的左右码各自相同或相反；相邻图幅同名边界点坐标在某一允许值范围内（如±0.5mm）。

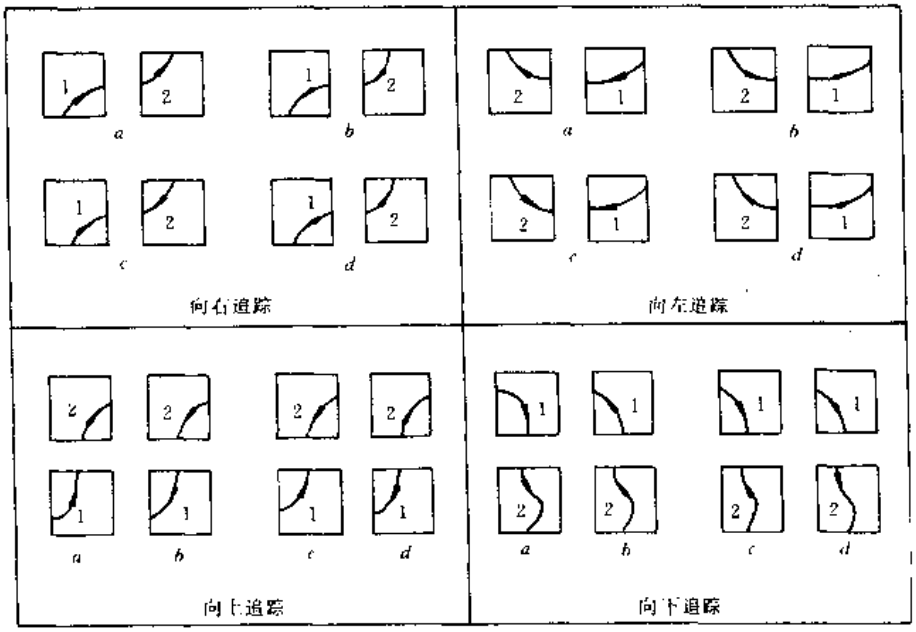


图 6—20　追踪拼接法

匹配衔接时是以一条弧或线段链作为处理的单元，因此，当边界点位于两个结点之间时，须分别取出相关的两个结点，然后按照结点之间线段链方向一致性的原则进行数据的记录和存储。

4. 相同属性多边形公共边界的删除

当图幅内图形数据完成拼接后，相邻图斑会有相同属性。此时，应将相同属性的两个或多个相邻图斑组合成一个图斑，即消除公共边界，并对共同属性进行合并（图 6—21）

多边形公共界线的删除，可以通过构成每一面域的线段坐标链，删去其中共同的线段，然后重新建立合并多边形的线段链表（图 6—22）。

对于多边形的属性表，除多边形的面积和周长需重新计算外，其余属性保留其中之一图斑的属性即可。

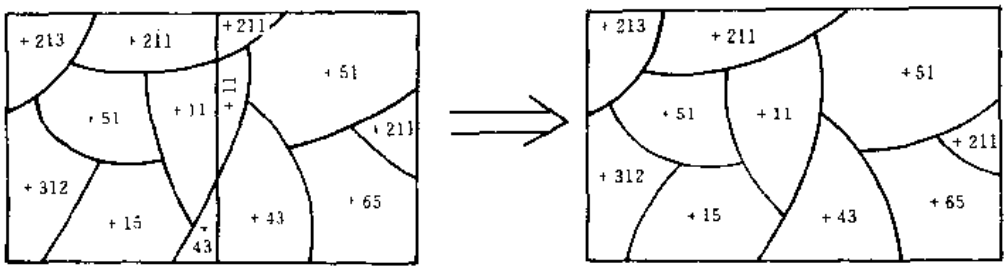


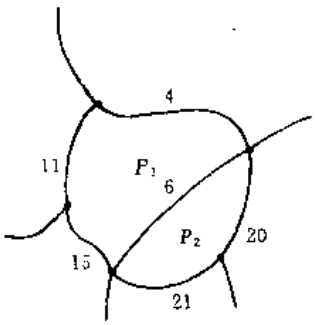
图 6－21　相同属性多边形公共边界的删除与属性合并

（二）图幅的分割

图幅分割实际上就是对整幅图层的多个相邻窗口处理，其算法原理与开窗运算相同。

七、空间数据的统计综合

空间数据的统计综合是一种基于属性数据的空间图形处理操作。如属性数据重新合并归类后空间图形的化简，其操作步骤和结果见图6—23。



P_1		P_2		$P_1 + P_2$
4		6		4
6	+	20	=	20
15				21
11		21		15
				11

图 6—22　多边形公共边界的自动删除

八、数据更新

许多地理数据会随着时间的变迁而发生变化。如行政区界的改变、土地使用权的再分配使土地界线改变等造成图形边界的坐标变化，土地使用权的转让等会造成属性数据的改变。为了保持GIS中数据的时效性，应对GIS数据进行更新，即修改陈旧数据以新数据代替或补充。此时，必须进行多方面的调查与分析，了解现存于数据库中哪些空间数据发生了变化，变化程度多大，以确定更新的方法。

空间实体更新的方法是：若变化程度不大，则直接使用图形编辑的方法进行；若变化区域面积大，则采用下列方法和步骤（图6—24）：

（1）重新数字化已变化区域的空间实体；

（2）显示整个区域的空间实体，对变化区域作负开窗处理；

（3）对以上两种空间实体进行叠加处理。

不论哪种方法，均需重新建立空间实体图形与属性数据的一一对应关系（拓扑结构）。

属性数据的更新，相对来说要简单一些。因为目前许多GIS软件采用关系数据库管理系统（RDBMS）来管理属性数据，且为用户提供了较强的数据管理功能，如表6—2所示。

因此，可直接利用属性数据库管理模块提供的功能，提取属性已变化的空间实体，对相应的属性进行编辑处理。

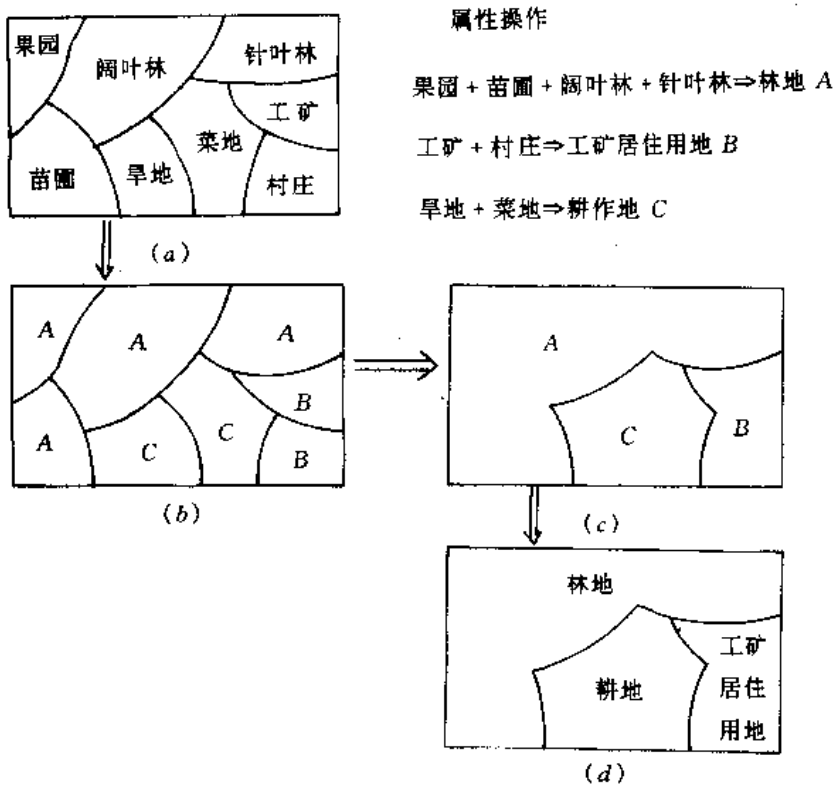


图 6—23　空间数据的统计综合

a—原始空间数据；*b*—属性重新分类；*c*—邻接线段删除；*d*—空间数据综合结果

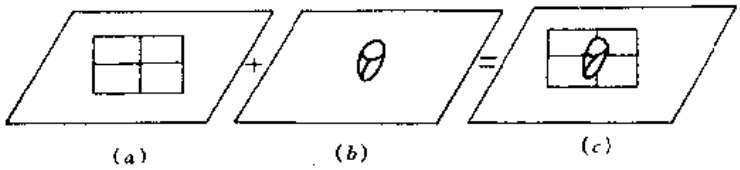


图 6—24　空间实体数据的更新

a—原始数据库；*b*—变化区域；*c*—更新结果

表 6—2　属性数据库管理模块

文件管理	结构设计	数据编辑	数据查询	统计报表
创建文件	设计新结构	添加数据	浏览数据	统计总数
打开文件	修改老结构	插入数据	显示一条记录	计算平均数
添加文件	拷贝结构	删除数据	显示几行属性	统计最大，最小值
存储文件	删除结构	修改数据	显示几列属性	报表设计
打印文件	合并结构	合并数据库	条件查询	图示分析
退出	查询结构信息	移动数据	多表连接查询	打印报表
		行拷贝数据		
		列拷贝数据		

九、空间数据的索引与检索

(一) 空间数据的检索

在对空间信息进行各种管理操作过程中需要经常提取和检索空间数据，如：根据图形检索其相应的属性信息；按照属性特征检索对应的空间实体；或者依据空间实体的相互关系检索其它实体，这种关系可以是关联关系和位置关系，举例如下（图 6—25）：

1）面—面关系　检索与某个面状地物相邻的所有多边形（图 *a*）。

2）线—线关系　检索所有与某一主河流关联的支流（图 *b*）。

3）点—点关系　检索到某点的距离小于 3km 的所有点状地物（图 *c*）。

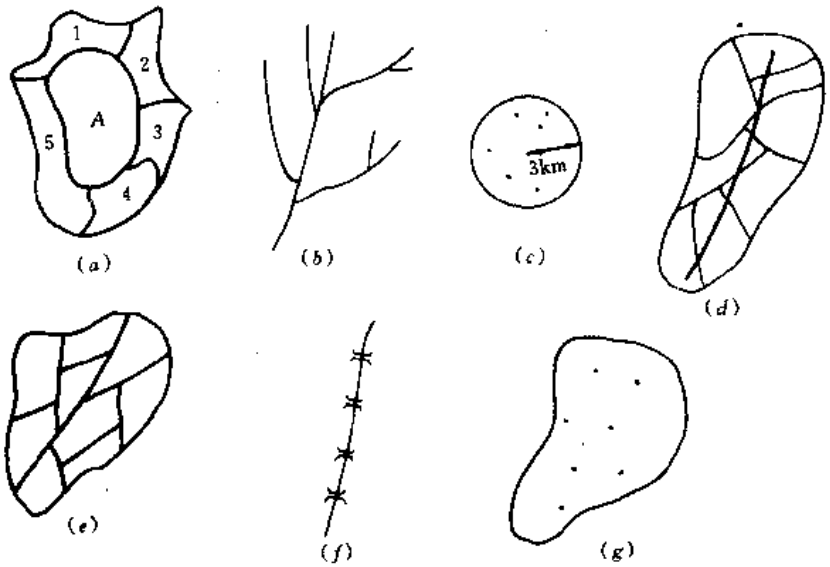


图 6—25　拓扑空间检索示意图

4）线—面关系　检索某道路经过的行政区（图 *d*）或检索某县内的所有公路（图 *e*）。

5）点—线关系　检索某河流上的桥梁（图 *f*）。

6）点—面关系　检索某区内邮筒的个数（图 *g*）。

(二) 空间数据的索引

GIS 数据库是一个庞大的数据库。要完成上述内容的检索，对空间数据进行存取是一件很费机时的工作。为了提高空间信息检索效率，应建立良好的索引机制——索引文件。

1. 索引文件

在讨论空间信息索引之前，首先对索引的背景知识作些介绍。

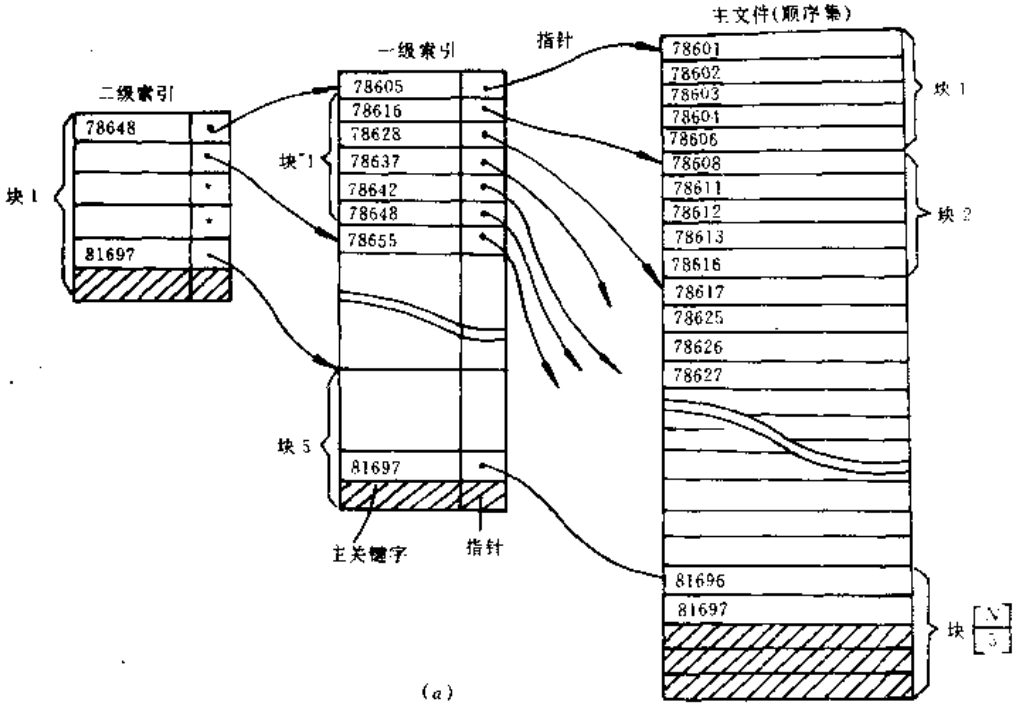
索引文件为组织数据提供了另外一种途径。其特点是，除了存储记录本身（主文件）之外，还建立了若干索引表。这种带有索引表的文件叫做索引文件。索引表中列出记录关键字和记录在文件中的位置（指针），读取记录时，只要提供主关键字，系统通过查找索引表获得记录在随机存取介质（磁盘）上的物理位置，然后取出该记录。索引文件既可以是有序的，也可以是无序的，可以是单级索引，也可以是多级索引。多级索引可以提高查找速度。但占用的存储空间较大（图 6—26）。

2. 空间数据的索引

空间索引的目的是为了加快空间定位数据的存取速度，尤其对大型空间数据库而言更

是如此。可以按空间实体的名称或一连串的实体元素（坐标）来建立空间索引。然而，空间索引应当针对位置或空间块，而不是对某个具体的目标提供存取路径。

在此使用图层这一概念来表示要进行空间索引的区域。那么索引就是与图层内某块空间实体相联系的一种数字符号。建立空间索引的方法之一是在图层上叠置一个正方形或矩形的格网（图 6—27），并以某实体的坐标作为进入数据库提取该实体的索引。但此时，需要建立以 x 方向和 y 方向的两种不同的索引，是十分不利的，也就是说，在二维空间对数



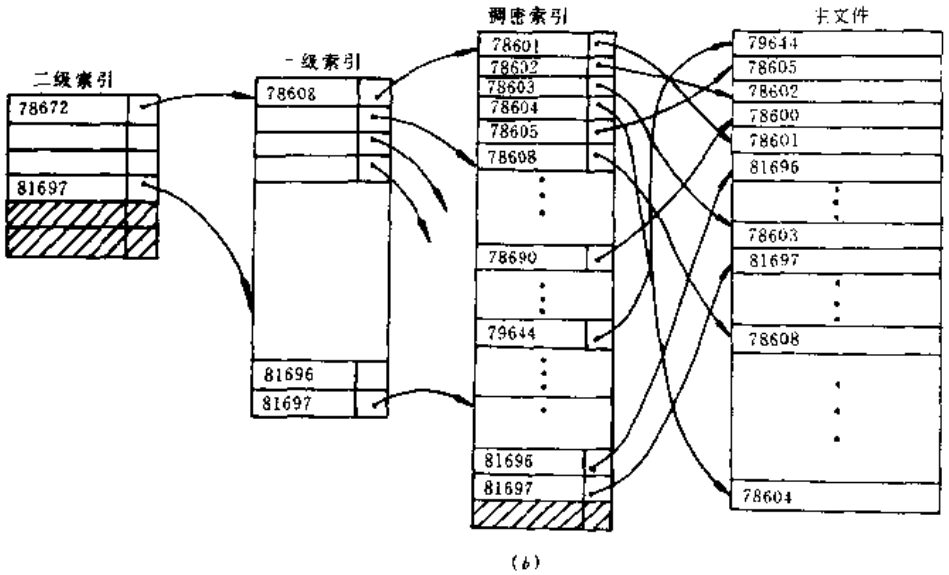


图 6—26　索引文件的结构

a—索引有序文件；*b*—索引无序文件

据记录进行存取时，必须搜索这两种索引，然后把结果联系起来。同理，对三维空间进行索引时，需建立三种索引（x，y，z），也很不便。

如果采用如前所述的经典索引方法，即通过标识符来存取数据，那么象里程碑、邮政编码、地址等属性，可当作存取标志。此时，可以通过如“人民路”的街道地址来提取某个实体或是提取发生在“245km”处的事件。但是，以二维或三维表示的定位点（x，y）或（x，y，z）则不能用此法直接处理。事实上，即使多维问题可用综合索引来解决，还仍然存在着需处理大量标志的问题。

第二种方法是用单个索引来表示空间排列的实体。假定有一如图 6—28 所示的区域，包含有 3 个点状实体 A、B 和 G，2 个线状实体 C 和 D 以及一个面状实体 F。如果采用空间曲线中具有游程编码的 Peano 码作为实体的索引，则在此图中，给定了实体的起、终点 Peano 码。

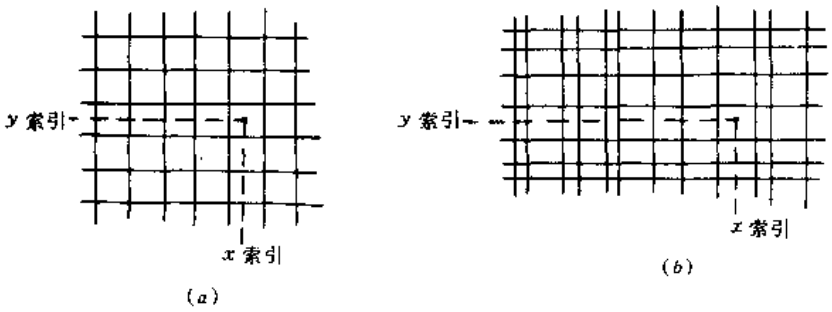
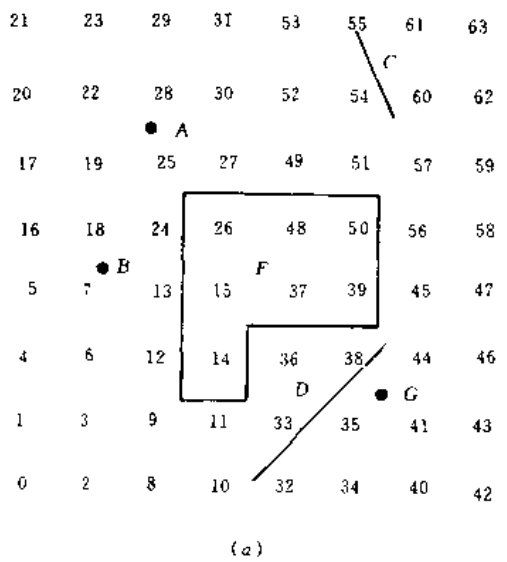


图 6—27　用方形或矩形格网的双索引

a—固定格网索引；*b*—可变格网索引



Peano 码	实体
7	B
14	F
15	F
25	A
26	F
32	D
33	D
35	D、G
37	F
38	D
39	F
48	F
50	F
54	C
55	C
60	C

(b)

实体	Peano 码
A	25 - 25
B	7 - 7
C	54 - 55
C	60 - 60
D	32 - 33
D	35 - 35
D	38 - 38
F	14 - 15
F	26 - 26
F	37 - 37
F	39 - 39
F	48 - 48
F	50 - 50
G	35 - 35

(c)

图 6—28　空间索引的实例

a—地图中的空间实体与 Peano 编码法；*b*—空间索引表；*c*—实体索引表

在此例中，是用一维索引作引导，以提取图层中的某些空间实体。如以一个单一的 Peano 码 35，可以确定几个实体 D 和 G。

为了索引的完整性，也可以同时用实体编码和简单的字母数字式标识符，以检索实体的位置及其属性（此图中未绘出）。

可以用下列几种方法来建立空间索引（图 6—29）：若把实体中的点看作分形点，并用空间充填曲线法对它们排序，则可用分形几何来确定它们的分辨率大小；为了建立某种范围区，如最小矩形边界、圆等，并把实体按某种分割规则分层，可利用欧氏几何原理从中选择特殊的点；使用综合的欧氏几何及分形几何四叉树法。

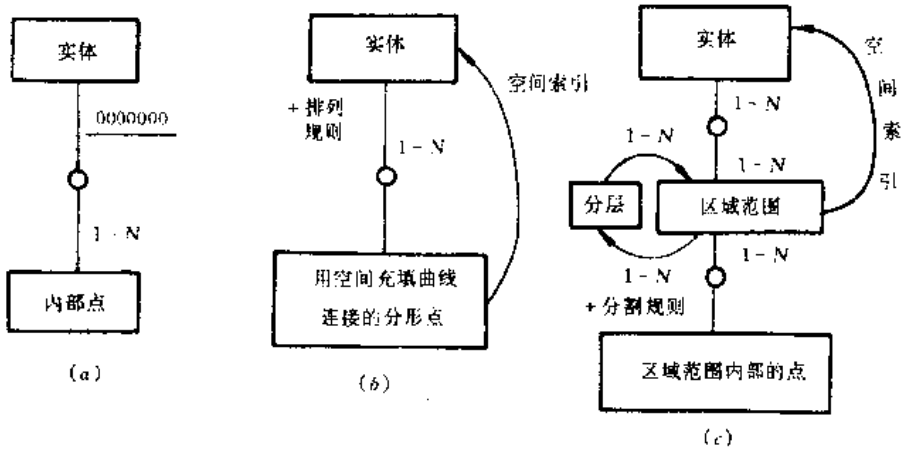


图 6—29 空间实体与其实用索引的理想模式

a—不能实现的理论模型；*b*—基于分形几何的实用等价模型；*c*—基于区域分层的实用模型

1）空间充填曲线索引法

空间充填曲线，如 Peano 曲线（N 型）和 Hilbert 型（π 型），可对图层内所有点经过一维编码来排序。但它们不可能适合于所有情况，而且，不同曲线的排序特性、稳定性以及计算的难易性都不同。因此可以有几种选择标准（影响因素）：

评价不同曲线的第一个标准是对实体进行空间查询的能力。对此，这两种曲线都不存在问题。

第二个标准是曲线从一点穿过相邻点的性能。由于 Hilbert 曲线中的相邻点在空间上是邻接的。因此 II 型排序较好，而 N 型排序则有时得不到保证。

第三个标准是以坐标计算关键码或其逆运算的速度衡量，此时 N 型排序比 II 型排序快。

第四个标准是借助四叉树作空间分层索引的实用性。

另外，空间索引应能够有效地组织点、线、面类及三维实体。对于点来说，Peano 码较好，并用四叉树很容易扩展到面域。然而，对于很长的线段，用这种空间充填曲线法作空间索引则有些不足。

图 6—30 是用 Peano 码表示点状实体的一级和二级索引的一个例子。

这种索引是一种分级目录，它比普通的顺序索引有较高的效率，尤其是处理大量实体数据时更为明显。为简明起见，在此，关键码采用十进制码，而实际应用中是二进制码。

对于面类实体的空间索引可用同样方法建立。要注意的是，因为面类实体所覆盖的位

置可能会在曲线的不同部位，所以必须同时使用低位和高位值（图 6—31）。因此用二级索引表示 Peano 码的低位和高位值。一般情况下，若想知道位于 (x, y) 处的是什么目标，则可把该坐标转换成 Peano 码，并从索引表中查找。

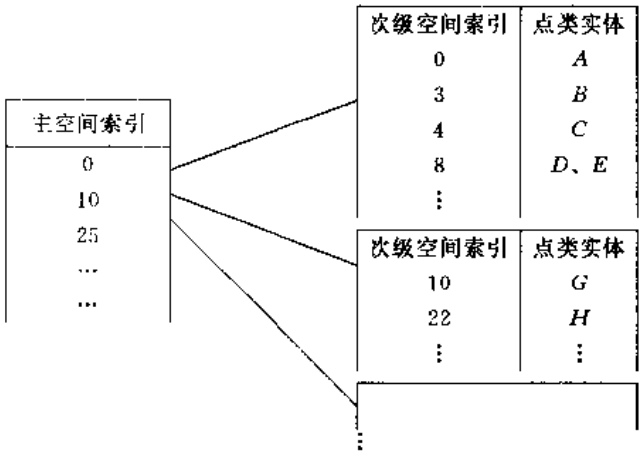


图 6—30 Peano 码空间索引

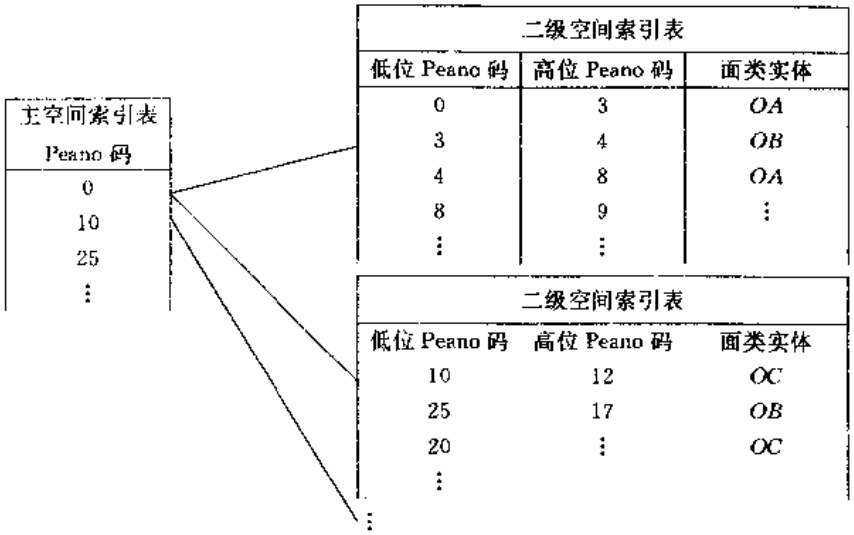


图 6—31 面类实体 Peano 码空间索引

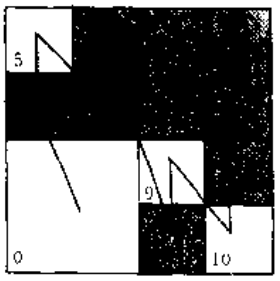
空间充填曲线法的主要缺点是：该方法的代码对方位和原始空间坐标十分敏感。当拼接两个不同来源和方向的图层时，其中的一幅图层必须在拼接前重构方位和坐标。因为 Peano 码需重新计算，其索引结构也相应地发生了变化，使在拼接不同图幅后需重新组织索引表。

2）四叉树索引

四叉树的应用为空间实体的检索提供了另一种途径。图 6—32 分别使用 Peano 码和 Hilbert 码来表示其概念。表中列出的数据包括了四叉树块代码，块的大小即树的层次以及属性状况，此处是树结点的颜色。

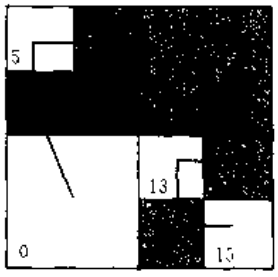
图 6—33 和图 6—34 分别描述使用线性四叉树和分层四叉树表示一个图层中包含多个不同实体的空间索引方法。图 6—34 显示了图 6—33 中处于不同水平（深度）的实体和标号。

在图 6—33 中，线实体 E 跨居边长为 2 的 3 个方块。因此它必须以边长为 4 的整个大方块为基准，其编码值是位于左下角的“0”。图 6—34 揭示了一定空间分辨率的不同实体的可检索性。因为空间充填曲线索引法需要大量索引，所以为了能使用四叉树法，最好对分形点按四等分重新分组，这样可以用不同大小的块来存储不同的实体。因此，大面域地理实体的位置更靠近树的根部，小目标位于叶端。



(a)

Peano 码	方块边长	颜色
0	2	白
4	1	黑
5	1	白
6	1	黑
7	1	黑
8	1	黑
9	1	白
10	1	白
11	1	黑
12	2	黑

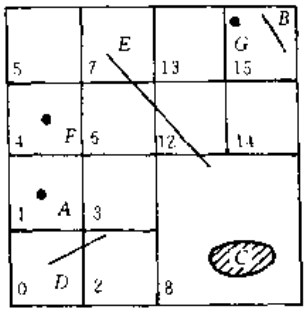


(b)

Hilbert 码	方块边长	颜色
0	2	白
4	1	黑
5	1	白
6	1	黑
7	1	黑
8	2	黑
12	1	黑
13	1	白
14	1	黑
15	1	白

图 6—32　Peano 码和 Hilbert 码的空间索引编码方法

a—Peano 曲线；*b*—Hilbert 曲线



Peano 码	边长	实体
0	4	E
0	2	D
1	1	A
4	1	F
8	2	C
15	1	B、G

图 6—33　用线性四叉树组织的空间索引

3）R 树与 R⁺树索引

此外，还可以使用包围空间实体的区域范围来实现空间索引。使用 R 树或 R⁺树方式组织的最小边界矩形法就是其中的一种。

就 R 树而言，认为有几个实体被 N 个矩形（Rectangles，R）所包围，现欲寻找某一特定的矩形，或是检索一个矩形中某一特定的点，若对数据不进行适当组织的话，那么测试的次数与矩形的个数成正比。假如有成千上万个矩形，则所需的机时量将无法接受，因此需要研究更快的检索方法。

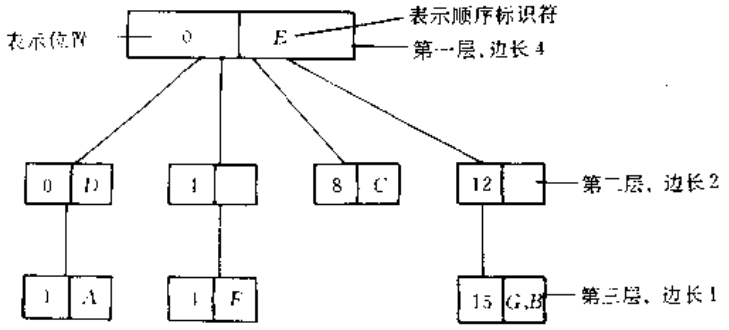


图 6 34 用分层四叉树组织的空间索引

一种方法是把相邻的矩形重新组合成一个大的伪矩形。反复进行这种操作可建立分层矩形结构。对于该结构的测试次数与实体个数成对数增长而不是成线性增长。例如，若不进行组织，则对于 10000 个矩形的测试次数为 5000，而用 R 树结构时，测试次数仅为 100 次，这种分层结构见图 6 35。

为了简单起见，本例中仅有 2 层，内层矩形为空间实体，外层矩形为伪矩形。如 B 伪矩形中包含了实体矩形 H、I、J、K。

这种矩形方向与坐标方位轴一致，以便：包含尽可能多的空间实体；矩形间的重叠率尽可能少；允许在每个矩形内再划分小矩形。

R 树空间索引就是按包含实体的矩形来确定的，树的层次表达了分辨率信息，每个实体与 R 树的结点相联系，这点与四叉树相同。使用这种关系模型，可以很方便地对 R 树进行编码。在关系模型中，对矩形可作如下描述：

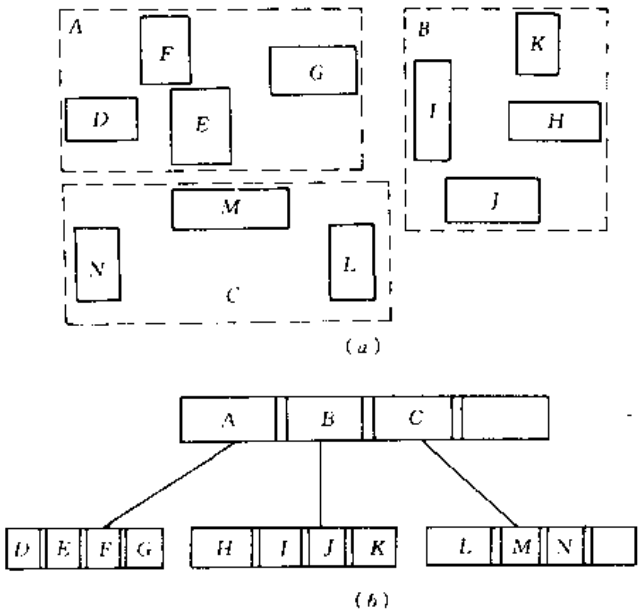


图 6—35 R 树结构示意图
a—二层不重叠矩形；*b*—层状结构

RECT （Rectangle-ID， Type，Min-X，Max-X，Min-Y，Max-Y）

其中 Rectangle-ID 为矩形号，Type 是矩形的类别（真实矩形或伪矩形）Min-X，Min-Y 为对应矩形左下角坐标；Max-X，Max-Y 为对应矩形右上角坐标。

在重叠时，伪矩形与真实最小包围矩形的编号相同。此时，还必须建立不同层次矩形的相互关系表达式：

PS（高层伪矩形标识号，低层伪矩形标识号）

由于矩形空间分布的不均匀性，使得有时不得不对它们进行分割并使 R 树保持平衡，这样，就产生了 R^+ 树。例如，为了使一棵树保持平衡，使每一层上都有 4 个真实矩形或伪矩形，这是对任一实体具有相同的存取时间的必要条件（图 6—36）。

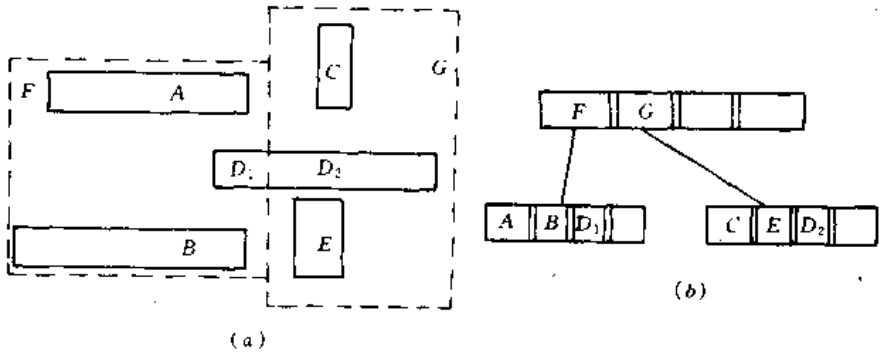


图 6—36 R^+ 树结构示意图

a—矩形被高层矩形分割；*b*—层状结构

在数据的关系模型中，R^+ 树的编码规则与 R 树相同，但此时，对于被分割的矩形，还应有关系表达式：

DECOMP（原矩形标识号，分割后矩形 1 的标识号，分割后矩形 2 的标识号）

4）用其它形式树索引

最小边界矩形索引的主要缺点是其空间索引的方法受实体方位的限制。最近，有人提出了基于球（圆）形（图 6—37）和任意多边形（图 6—38）的索引方法。如 Van Oosteron 和 Classen[55]提出了用圆形来包围实体（对三维实体用球），尽管使用圆形在计算上增加了难度，但圆形不受实体方位的限制，并且当坐标轴旋转时，使用圆形包络面空间索引也不受影响，其主要难点是自动确定包络实体的圆（球）的算法。

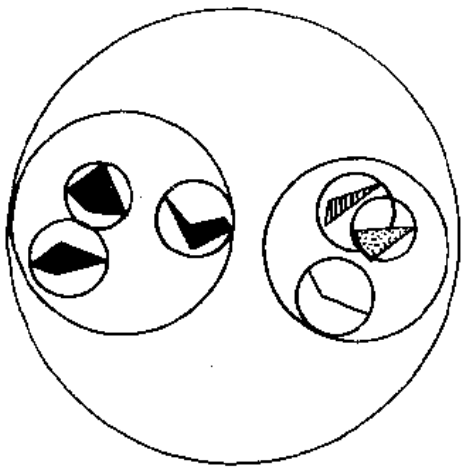


图 6—37 球形树空间索引

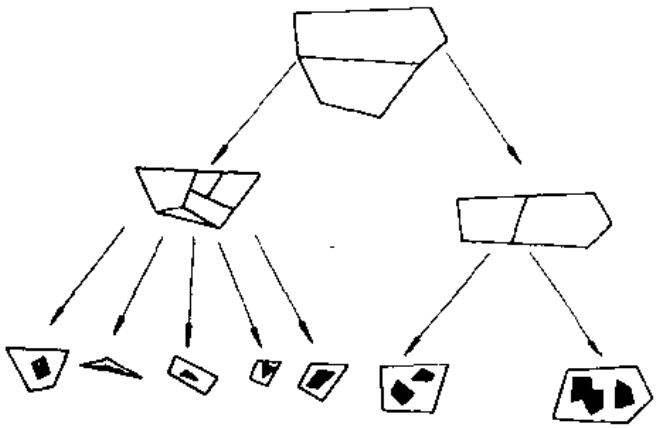


图 6- 38 任意多边形空间索引

还可使用任意多边形的空间索引法，一般为凸状多边形。其技术关键是快速地确定包

围实体的多边形，特别是它的边数。

此外，美国马里兰大学的 Faloutsos 和 Rong 于 1989 年提出了综合考虑 R 树与分形特性的所谓双转换法（图 6—39）：用最大或最小坐标值定义的矩形可以用四维空间（Min-X，Min-Y，Max-X，Max-Y）中的一个点来表示，这是第一次转换；然后，所有表示矩形的四维点用四维 Hilbert 码或 Peano 码排序，这是第二次转换。研究结果表明，4 维 Hilbert 码较好些。

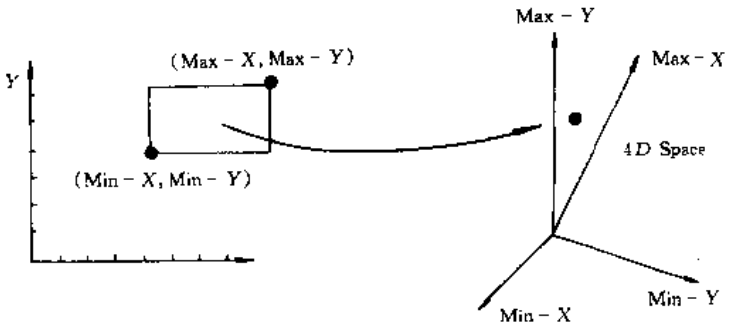


图 6—39　以 4 维空间中的一个点表示一个矩形

第 6—4 节　空间数据的质量与表示

为了对交换的地理数据的特征、属性及空间几何位置进行评价，使它们不含糊地综合到地理数据库中，必须对 GIS 中的数据质量进行评述。如果进入数据库数据的质量过高，超过实际需要，则是浪费，很不经济；反之，质量偏低，则达不到要求，可能造成严重的后果。

由于地理信息学（Geoinformatics 或 Geomatics）所研究的各种地理对象普遍地存在着某种不确定（uncertainty），而绝大多数 GIS 的实验者却假定，世界是由那些有着清晰确定的分界和绝对确定的属性的地理对象构成的。这一认识上的理想化必然带来偏差或数据质量问题。对数据质量的评述指标不象传统测量数据那样主要是精度或误差这一项，而通常是七种指标：数据情况说明（source 或 lineage），精度（包括位置精度和属性精度）、现势性（currency），逻辑一致性，完整性（completeness）、可靠性（reliability）和安全等级（security classification）。这些数据质量和精度问题来源于人们对客观世界认识的不足和局限性，人类自身条件（如人眼的分辨能力等）的限制以及技术水平和仪器设备的限制等。它们是不可避免的，但应该认识它们，设法控制它们。因此，随着 GIS 技术的深入发展，应当研究：进入 GIS 数据库的各类数据质量和精度的确定方法；GIS 中所用各种模型对数据误差的传播规律，并在模型中增加误差控制机制；根据用户对输出结果的要求，帮助用户确定各种输入数据的误差限。总之，对 GIS 数据质量的度量、分析评价及可视化的研究已成为 GIS 学科的重要内容，并且与 GIS 用户密切相关。

由于 GIS 空间数据的基本特征是与时间有关的空间位置特征和专题属性特征，因此，空间数据误差一般分为位置误差、属性误差和时间误差，也有人把它分为空间误差、专题误差和时间误差。前面提到的数据情况说明是指数据的来源、数据的处理、编码方法和数据所进行的变换；位置精度和属性精度分别指精度的空间因素和专题因素；逻辑一致性指数据编码关系的可靠性，包括拓扑、空间位置及专题属性的一致性；完整性是描述数据库

中目标之间和目标的抽象概括之间的关系。

空间数据的质量通常用误差来衡量，而误差定义为空间数据与其真值的差别。空间数据误差的来源是多方面的，例如，GIS 的原始录入数据本身包含着数据采集过程中引入的源误差。另外，在原始数据录入到空间数据库以及随后的数据分析处理和结果输出过程中，每一步都会引入新误差。根据 P. A. Burrough (1986) 的建议，可以将 GIS 数据误差的来源归纳为三类，如表 6—3 所示。

表 6 3 空间数据误差的来源

误差类型	误差来源	误差特征
源误差	数据年代；数据的空间覆盖范围；地图比例尺；观测密度数据的可访问性；数据格式；数据与用途的一致性；数据的采集处理费用	明显、易探测
由自然变化或原始测量引起的误差	位置误差；属性误差；质量和数量方面的误差 数据偏差；输入输出错误，观测者偏差，自然变化	不明显 难测定
GIS 处理过程引起的误差	计算机字长引起的误差 拓扑分析引起的误差；逻辑错误、地图叠置操作 分类与综合引起的误差；分类方法，分类间隔，内插方法	复杂 难探测

一、源误差

GIS 数据的来源主要有，直接从现场利用 GPS 或全站仪采集的数字数据；现有纸质地图的数字化；航空影像和遥感数字数据或统计调查数据等。这些数据受表 6—3 中前两类误差来源的部分或全部影响。

(一) 地面测量数字数据的误差

来源于地面测量的数字数据中含有控制测量和碎部测量误差。其中控制点误差又受控制网的参考基准、网形和观测精度以及观测费用等因素的影响。碎部点误差除继承了控制点的误差外，还受自身的观测方法、观测精度和地界的人为判断，以及地物地貌的取舍等因素的影响。当然，原始观测误差受观测仪器、观测者和外界环境三种因素影响。地面测量数据中的误差可以表现为随机误差、系统误差或粗差。一般而言，粗差采用可靠性理论探测剔除，系统误差采用实验方法校正或建立系统误差模型处理，随机误差采用随机模型，如最小二乘法平差处理。利用误差传播律可求得点坐标的方差-协方差阵或误差椭圆(球)元素来表达点坐标数据的精度。

(二) 地图数字化数据的误差

目前 GIS 数据的主要来源之一是现有地图数字化。原图固有误差和数字化过程误差是地图数字化数据误差的主要来源。原图固有误差除含有上述地面控制测量和碎部测量的全部误差外，还含有制图误差。

1. 制图误差

1) 控制点展绘误差　展绘控制点是成图的第一步。对于与立体测图仪相连的直角坐标展点仪而言，该项误差的均方根为 0.17～0.32mm，这在地图精度要求不高时可不予考虑。

2)编绘误差　地形图的编绘一般用摄影测量手段把各种原始资料化为同一比例尺的过程，其中所引入的误差大约在 0.30～0.33mm 之间。

3) 绘图误差　绘图误差是在绘图过程中产生的。误差的范围为 0.06～0.18mm 之间。

4) 综合误差　地图综合误差的大小取决于特征的类型与复杂程度，故该项误差极难量

化。象曲线这类特征，其综合误差可能很大，而直线的综合误差又近似为零。当两个或两个以上的特征无法在图上适当表示时，需要人为地在地图上将特征表示符移位，使地图清晰易读。这种移位也会引起综合误差，其大小主要取决于特征的接近程度和成图比例尺的大小。比例尺越小，移位越大，误差也越大。

5）地图复制误差　地图复制误差的均方差为 0.1～0.2mm。

6）分色板套合误差　彩色地图是由一系列金属分色板套合生成的，每一块分色板在图上印出一种颜色。该项误差的均方差为 0.17～0.30mm。

7）绘图材料的变形误差　地图一般印刷于纸质载体上，随着温度和湿度的变化，纸的尺寸也会改变。如果温度不变，湿度从 0%增至 25%，则纸的尺寸改变 1.6%。由于纸的膨胀系数和收缩系数并不相等，因此，即使湿度又降至原来大小，图纸也不可能恢复到原来尺寸。由于纸张在印刷时温度升高，纸张长度会伸长 1.5%，宽度伸长 2.9%，而当纸干燥或冷却后，其长度和宽度又分别收缩 0.25%和 0.75%。因此，印刷冷却后，图纸在长、宽方向上的净伸长分别为 1.25%和 2.15%。

8）归化到同一比例尺所引起的误差　地图的比例尺是指真实的主比例尺。例如，对兰勃特正形投影而言，主比例尺只有沿着标准纬线时才是不变的，而标准纬线之内的比例尺变小，之外的比例尺变大。因此，在将地图数字化或者从地图量距时，必须利用适当的比例尺因子进行修正。

9）特征的定义　自然界的许多特征并无明确的界限。例如，海岸线的实际位置在哪里？森林和草地间的界线如何确定？因此，特征定义会引起特征位置的某些不确定性，但并非所有的特征都有此误差。

10）特征夸大误差　为了增强图的可读性，有时需要夸大某些特征，因为它们在图上难以按其真实情况表示。例如，分界线应是无宽线、但表示在图上时，却占有一定宽度。有时根据地图的用途，需要将某些特征夸大，以区别于其它特征。如，在道路上将道路特征夸大以示醒目。特征夸大误差取决于地图的比例尺、用途和所涉及特征的类型。同样地，并非所有的特征都具有此类误差。

由于很难知道制图过程中各种误差间的关系以及图纸尺寸的不稳定性，因此，很难准确地评价原图固有误差。

2. 数字化误差

数字化方式主要有手工数字化和扫描数字化。目前在生产实践中多采用点方式的手工数字化。线划跟踪与扫描数字化所引起的平面误差较小，只是在扫描数字化时，要素结合处出现的误差较大。跟踪和扫描数字化引起的误差主要与被数字化的要素对象、作业员和数字化仪三者有关。

1）要素对象　要素本身的宽度、密度和复杂程度对数字化误差有显著影响。例如，数字化曲线时，不可能总是在线的中央采样，从而引起数字化误差，将地图上的连续曲线转换成数字表示时，包括了将曲线分解为小段直线和采样两个过程，组成曲线的无穷多个点在数字化过程中仅有很少一部分被采样，这是产生数字化误差的另一个原因，这类误差随采样点数目的减少而增大。此外，当数字化地图要素密集的区域时，例如山区的等高线时，作业员易产生误差和差错。不同的要素，数字化误差也是不同的，例如，点状要素出现的误差不同于线状要素出现的误差。

2）作业员　作业员的经验与技能主要表现在选择最佳点位、跟踪曲线和判断十字丝与目标重合程度的能力等方面。作业员疲劳时的数字化质量要比正常情况下差很多。因此，要保持连续一致的精度，每天的数字化工作最好不超过 4 小时。另外，数字化速度对数字化误差也有影响。

3）数字化仪　数字化仪的分辨率对数字化误差有决定性影响。一般在选择数字化仪时采用经济实用的原则。但应注意到，数字化仪的实际分辨率与标称分辨率往往不一致，一般都低 1 至 2 个分辨单位。因此，选择数字化仪时应有一定的精度储备。

（三）遥感数据误差

遥感数据的误差累积过程可以区分为：数据获取误差、数据处理误差、数据分析误差、数据转换误差和人工判断误差。

1. 数据获取误差

遥感资料获取时，存在着多种误差。其中有些误差源是可控的，例如，几何误差或辐射误差；有些则是不可控的，如气候条件和景观的自然变化。可控的数据获取误差又包括几何误差和遥感平台误差。

2. 数据预处理误差

数据预处理包括利用地面控制进行校正，几何校正、图像增强和分类等。这些预处理会引入各种各样的误差。例如，几何校正会引起平面位置和专题误差。位置误差取决于校正模型的逼近和校正模型参数求解时所用控制数据中的误差。专题误差来自重新抽样。分类本身就是一个引入显著误差的过程。潜在的分类误差源可能有遥感数据测量中的随机误差，类型清晰度中的模糊度，混合类型分类系统，以及不同的分类系统缺少相容性。

3. 数据转换误差

在矢量-栅格数据转换过程中，会出现显著误差。引起矢量数据栅格化误差的主要原因有两个：属性概括误差；拓扑匹配错误。栅格化误差的大小与下列因素有关：转换中所用的算法，格网的尺寸，栅格表达时的定向，以及地图图形的复杂性等。

4. 人工判读误差

在数据分析和判读阶段会引入人工判读误差。在 GIS 中使用遥感数据时，人工判读误差会影响随后的 GIS 分析结果。这种误差是主观的，因此难以量化，它与判读员从遥感影像中提取信息的能力和技能有关，还与影像的尺度和所需的概括程度有关。

二、操作误差

除了 GIS 原始录入数据本身带有的源误差外，空间数据在 GIS 的模型分析和数据处理等操作中还会引入新误差。如，由计算机字长引起的误差，拓扑分析引入的误差和叠置中引入的误差等。

1. 由计算机字长引起的误差

在计算机中，数据是由一定字长的编码数字表示的。舍入误差是由计算机字长引起的一种误差。这种误差出现在 GIS 的各种数值运算和模型分析过程中。由舍入误差引起的问题很多，例如 GIS 空间数据库中整数编码对面积和周长计算的影响，比例尺变换和旋转变换对拓扑关系的影响等。削弱舍入误差影响的方法主要有：增加字长并用带小数的实数表示数据；改变数据在计算机中的表达方式；采用合适的算法等。

除了数据处理精度外，数据存储精度也与计算机字长有关。16 位的计算机在存储低分

辨率的栅格图像时不会出现问题，但存储高精度的控制点坐标或点位精度要求高的地理数据时，则不能胜任。减少存储误差影响的方法有：采用双精度字长存储数据，使用有效位数多的数据记录控制点坐标；使用32位整数运算法处理数据，即把坐标数据转换成以厘米或毫米为单位的整数，然后用整数存储和运算；采用嵌套法，即先把研究区域划分成若干块，并用粗略的分辨率来定义它们，每块再以有限精度建立局部原点，以此为基准确定块中各细节的坐标。这种方法可以用单精度字长记录非控制点坐标，用双精度字长记录控制点坐标，并可将它们联接起来。嵌套可以是多级嵌套。

2. 由拓扑分析引起的误差

GIS中的拓扑分析会产生大量的误差。如矢量数据栅格化引起的拓扑匹配误差，多层叠置过程中产生的无意义多边形等。这是由于GIS中进行的大多数拓扑空间操作都隐含地假设：原始数据是均匀分布的，数字化过程没有错误；拓扑网络的叠置仅仅是网络多边形边界相交和重新构成网络问题；边界线能明确定义和描绘；所有的算法都假定为完全确定的操作；对某类型或其它自然因素所确定的分类区间是最合适的。

3. 数据分类和内插引起的误差

GIS空间数据库中数据的许多不规则性是由所使用的分类方法和从点数据到面数据的内插方法引起的，即它们作为GIS中的一种操作会引入许多误差。

以上讨论了GIS空间数据库中原始录用数据本身含有的源误差和随后空间操作中引入的新误差。一般来说，源误差远大于操作误差，因此，要想控制GIS产品的质量，良好的原始录用数据是首要的。

三、GIS空间操作中误差的传播

GIS产品是利用含有源误差的空间数据，通过GIS分析操作产生的。源误差和操作误差通过GIS操作经累积后传播到GIS的产品中。考虑如下的GIS空间操作：

$$y=f(x_1, x_2, \cdots x_n) \tag{6—13}$$

其中 x_i $(i=1, 2\cdots n)$ 为描述空间数据的自变量，它带有源误差；y 为描述GIS产品的因变量，$f(\cdot)$ 为描述GIS空间操作过程的数学函数，用以计算操作误差。根据 $f(\cdot)$ 的特征，可以分成两类运算：算术运算和逻辑运算。下面讨论这两种运算关系下的误差传播。

（一）算术关系下的误差传播

如果 $f(\cdot)$ 为算术关系，例如独立变量的和差关系、倍数关系或线性关系，则其中误差传播规律是众所周知的；若为相关或一般非线性函数时，其误差传播规律也在经典测量误差理论中已有详细介绍。

（二）逻辑关系下的误差传播

除了上述算术关系操作外，在GIS中还存在着叠置、推理等大量逻辑运算，如布尔逻辑运算（AND、OR，NOT）和专家系统中的不精确推理。

1. 布尔逻辑运算下的误差传播

布尔逻辑运算是GIS中的一类典型操作。如，空间集合分析就是按照逻辑运算合成的。举例来说，现有一幅土地利用现状图和一幅土壤类型图，需查询土层厚度大于50cm的麦地。在分析操作时，首先，从土地利用现状图中找出小麦地的子集 A_1，再从土壤类型图中获取土层厚度大于50cm的子集 A_2；然后求两个子集的交集 $A=A_1 \cap A_2$ 即为查询结果。由于子集 A_1 和 A_2 实际上都含有误差，如 A_1 子集中可能也包含了其它地物，麦地只是其中心

类别。设 A_1 的误差为 5%，A_2 的误差为 10%，则 A 的误差应是多少？这是不易计算的。

2. 不精确推理关系下的误差传播

GIS 作为辅助决策工具常常需要进行综合分析、评判和基于知识的推理。在利用含有误差的知识时，经不精确推理所得结论的精度和可信度如何？这是推理关系下的传播定律所要解决的问题。

逻辑关系下的误差传播律正处于研究中，借用信息论、模糊数学、人工智能和专家系统的基础理论可望解决这一问题。

通过上述讨论不难理解，由于 GIS 还是一门比较年轻的学科，对 GIS 数据的质量和精度问题还不大为人们所了解，还有许多问题有待于深入研究。例如，由于用户的需求多式多样，对数据质量和精度还没有一致的意见；缺乏度量空间数据和 GIS 输出结果不确定性的方法；在 GIS 功能中没有标准方法来建立误差模型。对如何处理误差，没有成熟的规范可循；现有的 GIS 能提供一个派生信息的工具，但却没能提供一个关于系统可靠性的工具，即交出的最终结果没有附上一份相应的关于其可靠性的分析报告。

第七章　GIS 中的数据分析与输出

第 7—1 节　空 间 数 据 分 析

地理信息系统（GIS）与一般的计算机辅助制图（CAC/CAD）系统的主要区别在于 GIS 具有空间数据的分析、变换能力。除一些基本的变换功能如数据更新、比例尺变换，投影变换外，主要的空间分析和变换功能为地理数据的拓扑和空间状况运算，属性综合运算，几何要素与属性的联合运算等。为了完成这些运算，GIS 一般都以用户和系统交互的形式提供以上分析处理能力。

图 7—1 表示以分级结构形式概括的各种空间分析类型和方法。其中一些是数据库的数据恢复——单变量或多变量统计分析等简单分析，有的则是涉及邻域分析、综合几何与属性的复杂分析。空间模拟技术能用来建立几乎无限的数据分析能力，有些复杂分析方法则是一系列简单方法经组合后形成特殊模型而产生的。应指出，栅格数据结构与矢量数据结构的空间分析方法有所不同。一般来说，栅格结构组织数据的空间分析方法要简单一些。

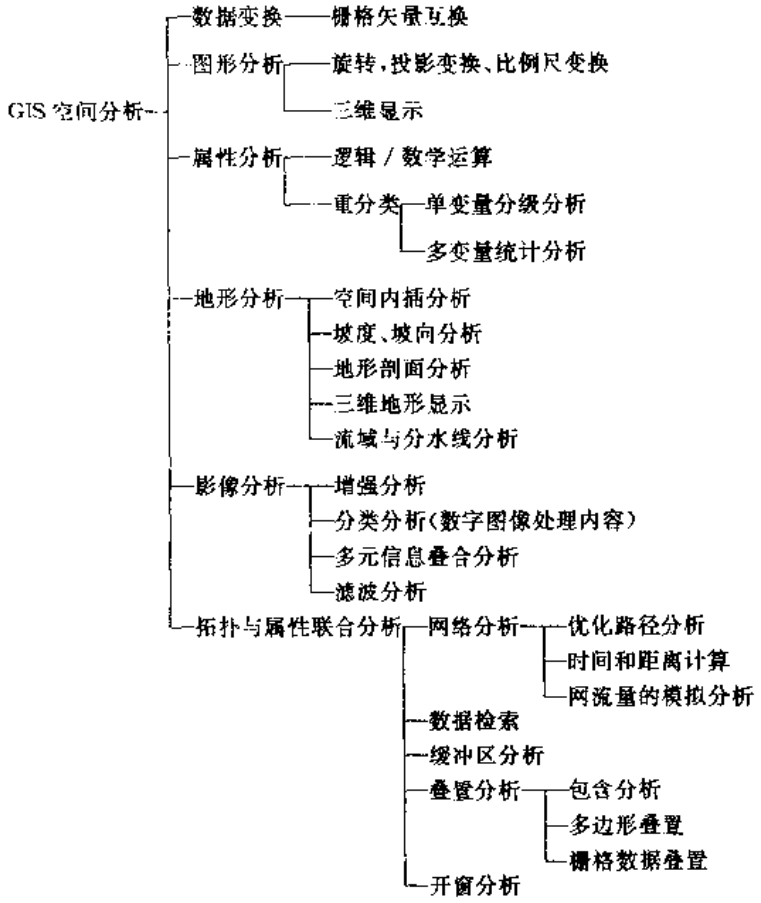


图 7—1　GIS 的空间分析方法

表 7—1 逻辑运算的真或假

A	B	NOT · A	A · AND · B	A · OR · B	A · XOR · B	B · NOT · A
1	1	0	1	1	0	1
1	0	0	0	1	1	0
0	1	1	0	1	1	1
0	0	1	0	0	0	0

注：表中1 表示“真”，0 表示“假”。

一、综合属性数据分析

GIS 中属性数据一般采用关系型数据库管理，因此，关系数据库中各种分析功能都可以对属性数据进行分析。

（一）数学计算

属性数据中的数字型数据可以进行“加”、“减”、“乘”、“除”、“乘方”等数学运算，以产生新的属性值，如人口数/图斑面积（km）＝人口密度。

（二）逻辑运算

逻辑运算的基本原理是布尔代数，这种逻辑分析几乎可以在所有的空间分析中得到应用。它按属性数据的组合条件来检索其他属性项目或图形数据，以及进行空间聚类。

布尔代数的基本运算符号是 AND，OR，XOR，NOT，逻辑运算的结果为“真”或“假”，如表 7—1 所示。

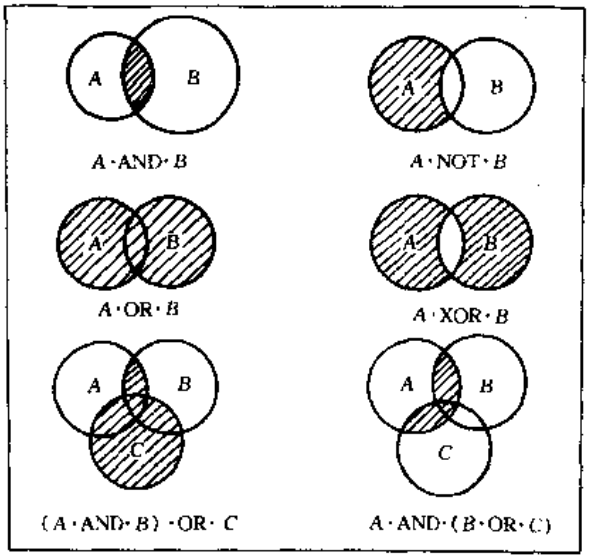


图 7—2 布尔逻辑运算的结果

假设 A、B、C 分别为具有属性 a、b、c 的集合，那么，布尔逻辑运算的结果组成了新的属性集合（图 7—2）。

例如在利用 GIS 进行土地规划时，不同图斑具有不同属性，其中 A 为土层厚度＞50cm 的土壤单元结合，B 为土壤类别为红砂壤的单元集合，C 为 pH 值大于 7.0 所有单元集合，则

A · AND · B 检索出全部土层厚＞50cm 且土壤类别为红砂壤的土壤单元；

A · OR · B 检索出全部土层厚＞50cm 以及土壤类别为红砂壤的土壤单元；

(A · AND · B) · OR · C 检索出全部满足土层厚＞50cm 且土壤类别为红砂壤这两个条件以及 pH 值大于 7.0 的所有土壤单元；

值得注意的是，布尔逻辑运算不满足交换律，(A · AND · B) · OR · $C \neq A$ · AND · (B · OR · C)，因此，在应用时，必须依“AND”或“OR”的优先权，用圆括号来明确运算顺序。

以上四个基本逻辑运算符号通过组合可以组成复杂的综合属性检索条件。

布尔逻辑运算表达式除完成检索功能外，还可以进行再分类分析。如，当（pH 值>7.0 · AND · 土壤厚度>50cm）· OR · 土壤类别为红砂壤时，土壤适宜种茶叶。

当（pH 值>7.0 · AND · 土壤厚度>50cm）· AND · 土壤类别为红砂壤时，土壤适宜种苹果。

此时，复杂的布尔逻辑表达式检索出满足上述条件的图斑并显示出结果，就可以得到种植适宜性评价图。

（三）单变量分级分析

属性的单变量分级分析是把单个属性作为变量，依据布尔逻辑方法分成若干个类别。例如，属性表中土壤厚度从 0 到 1m，依据表 7—2 中的标准划分为五个类别。

表 7—2 土壤厚度分级标准

土壤厚度 (m)	类 别
>0 · AND · <0.2	薄层
≥0.2 · AND · <0.4	较薄
≥0.4 · AND · <0.6	中厚
≥0.6 · AND · <0.8	较厚
≥0.8	厚层

这种分析方法，可进行属性数据的合并式转换，把复杂的属性类别合并成简单的类别，以实现空间聚合。

（四）多变量统计分析

多变量统计分析主要用于数据分类。在 GIS 中存储的数据具有原始的性质，以便用户可以根据不同的使用目的，进行任意提取和分析，特别是对于观测和取样数据，随着采用的分类和内插方法的不同，得到的结果有很大的差异。因此，在大多数情况下，首先是将大量未经分类的属性数据输入信息系统的数据库，然后要求用户建立具体的分类算法，以获得所需要的信息。

1. 变量筛选分析

随着现代数据收集系统的不断改进，在一个取样点上常可以收集到几十种原始变量。在这些变量中有许多是相互关联的，可以通过寻找一组相互独立的变量，使多变量数据得到简化，这就是变量筛选分析。常用的变量筛选方法有主成分分析法、主因子分析法和关键变量分析法等。

主成分分析是以取样点作为坐标轴，以属性变量作为矢量，通过以相似系数建立相关矩阵，研究属性变量之间的亲疏关系。

主因子分析是以属性变量作为坐标轴，以取样点作为矢量，通过以相关系数建立相关矩阵来研究取样点之间的亲疏关系。

关键变量分析则是利用属性变量之间的相关矩阵，通过由用户确定的阈值，从数据库变量全集合中选择一定数量的关键独立变量，以消除其它冗余的变量。

现以关键变量分析法为例，讨论其计算方法。

设有几个数据点，每点有 m 个属性变量，则多变量数据可表示为矩阵

$$x=\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & & & \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} \qquad (7-1)$$

这些变量之间的关系可用相关系数矩阵表示

$$R=\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \vdots & & & \\ r_{m1} & r_{m2} & \cdots & r_{mm} \end{bmatrix} \tag{7—2}$$

r_{ij}是数据标准化后的夹角余弦，其计算公式为

$$r_{ij}=\frac{\sum_{k=1}^{n}(x_{ki}-\bar{x}_i)(x_{kj}-\bar{x}_j)}{\sqrt{\sum_{k=1}^{n}(x_{ki}-\bar{x}_i)^2\sum_{k=1}^{n}(x_{kj}-\bar{x}_j)^2}} \tag{7—3}$$

式中 $0\leqslant r_{ij}\leqslant 1$；$j=1, 2, \cdots, m$；$k=1, 2, \cdots, n$。显然，$r_{ij}$越近于 1，变量 i 与变量 j 的关系越密切；r_{ij}越近于零，i 与 j 的关系越疏远。

2. 变量聚类分析

所谓变量聚类分析就是一系列数据观测点的属性变量，按其性质上的亲疏远近程度进行分类。m 个属性数据组成了 m 维空间，两个数据点在 m 维空间的相似性可用变量空间欧几里德距离来表示

$$d_{ij}=\sqrt{\sum_{k=1}^{m}(x_{ik}-x_{jk})^2} \tag{7—4}$$

式中各变量及参数的含义与式（7—3）相同。

数据点的相似性取决于 d_{ij}，d_{ij}小则相似性大，d_{ij}大则相似性小。

设有根据欧几里德距离建立的距离系数矩阵（表 7—3），可以用最短距离法获得聚类图（图 7—3）。

计算步骤如下：

（1）用式（7—4）计算数据点之间的距离，形成距离系数矩阵 D（0）（表 7—3）。此时，各个数据点自成一类，显然有 $D_{pq}=D_{qp}$（对角矩阵）。

（2）选择 D（0）中最小的非对角元素，设为 D_{pq}，则 p、q 两个数据点合并为一类，记为 r，在表 7—3 中，首先将第 6 和第 12 数据点并为一类。

（3）计算新类与其它类的距离。由于第 p、第 q 数据点已合并成一类 r，须将 D（0）中的第 p、q 行及第 p、q 列删去，并计算新类 r 到其它数据点的距离，插入到原 p 行、p 列中去，组成新的距离矩阵 D（1）。

$$D_{rk}=\min\{D_{pk}, D_{qk}\} \qquad k=1, 2\cdots m, \quad k\neq p, q \tag{7—5}$$

（4）重复上述（2）、（3）两个步骤，得 D（2）、D（3）…，直到所有的数据点都得到归类为止。若某一步中得到的最小非对角元素不止一个，则对应的这些类可以同时合并。对表 7—3 数据的聚类结果见图 7—3。

二、缓冲区分析

缓冲区（buffer）分析即根据数据库中的点、线、面实体，在其周围建立一定宽度范围的缓冲区多边形（图 7—4）。这是 GIS 空间分析的基本功能之一。例如，某项工程建设的地

表7—3 距 离 系 数 矩 阵 表

序号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1																			
2	0.4197	1																		
3	−0.0555	0.0215	1																	
4	0.0146	0.2443	0.5520	1																
5	−0.1286	−0.1258	0.6521	0.0017	1															
6	0.9148	0.4794	−0.0013	0.0897	−0.1336	1														
7	−0.1272	−0.2901	0.0923	−0.1626	0.0912	−0.1295	1													
8	−0.0216	−0.2681	−0.4527	−0.4438	−0.1182	−0.1131	−0.3763	1												
9	−0.0955	−0.5102	−0.0846	−0.3741	−0.1469	−0.1945	0.7495	0.0536	1											
10	−0.1983	−0.4163	0.3712	0.3499	0.0886	−0.1152	0.0924	−0.3346	0.1729	1										
11	−0.1419	−0.1401	0.2758	0.1303	0.1235	−0.1341	0.7621	−0.6021	0.3948	0.2742	1									
12	−0.0505	−0.0278	0.2721	0.4630	0.0872	−0.0047	−0.4045	−0.2361	−0.4099	0.3686	0.0673	1								
13	0.4632	0.5854	0.1813	0.0119	−0.0425	0.4932	−0.3295	0.1443	−0.3663	−0.2584	−0.1973	−0.1661	1							
14	−0.1989	0.0288	−0.5728	−0.1747	−0.3862	−0.2744	0.0075	−0.0004	−0.0362	−0.2117	0.0247	−0.1727	−0.4110	1						
15	−0.0900	0.2699	−0.2647	0.1417	0.2550	−0.0119	0.0613	−0.2873	−0.1113	0.0622	0.1865	0.2015	0.2092	−0.2986	1					
16	−0.1578	−0.4601	−0.4146	−0.1624	−0.3781	−0.1538	0.1325	0.2865	0.1152	0.1069	0.1982	0.0814	−0.1069	0.1082	−0.1984	1				
17	0.1612	0.4005	0.3855	0.2108	0.3742	0.1481	−0.1162	−0.3126	−0.0262	−0.0164	−0.1602	0.0853	−0.0309	−0.127	0.1187	−0.9357	1			
18	−0.0201	−0.3388	−0.4792	−0.3679	−0.2352	−0.0902	0.1987	0.2145	0.3245	−0.0492	0.0280	0.2987	−0.1219	0.1823	−0.2686	0.5614	−0.5396	1		
19	0.2291	0.6400	0.5935	0.4100	0.2019	0.3345	0.0702	−0.5273	−0.3254	−0.0692	0.2189	0.0437	0.4329	−0.3393	0.4440	−0.5603	0.4703	−0.5749	1	
20	0.2724	0.5739	0.6306	0.4543	0.1860	0.3827	0.0435	−0.4884	−0.3265	0.0099	0.2436	0.1031	0.4599	−0.4296	0.4151	−0.4744	0.3926	−0.5286	0.9535	1

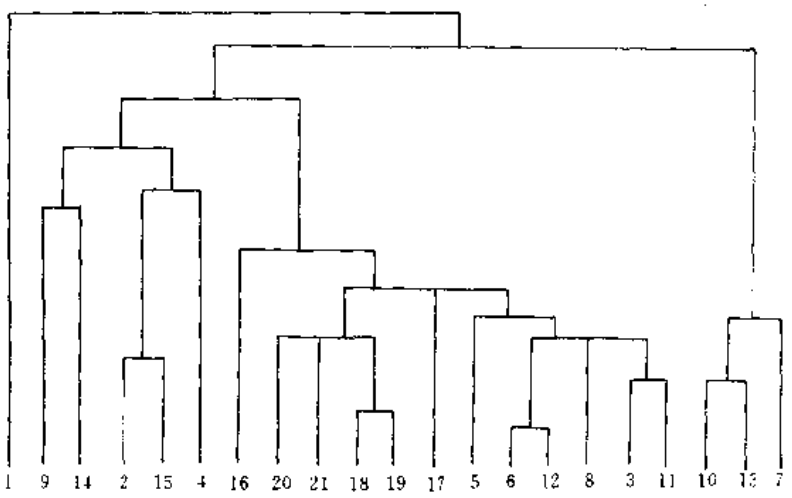


图 7—3 变量聚类结果（据黄杏元，1990）

址已选定，应根据工程建设的影响范围（如 500m）通知周围居民作搬迁或采取其它措施；在森林规划中，需要按照距河流一定纵深的范围来规划森林的砍伐区，以防止水土流失；在野生动物保护评价中，许多动物的活动范围总是局限于栖息地（河流、洞穴、巢）的一定范围内。在这些例子中，都需要沿这些点、线、面建立一定范围的缓冲区

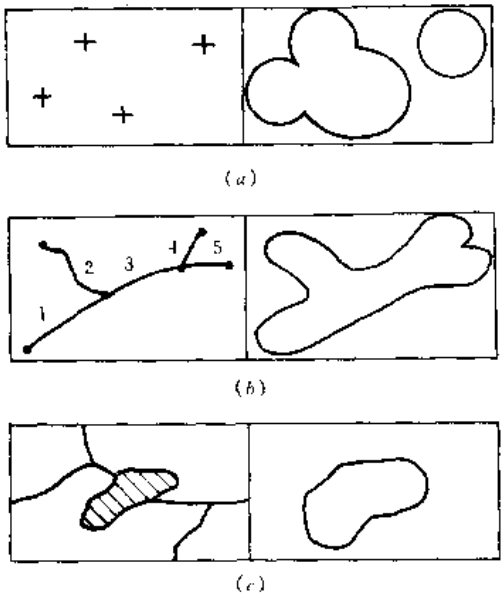


图 7—4 不同类型实体缓冲区的建立（据 ESRI，1988）
a—点缓冲区；*b*—线缓冲区；*c*—面缓冲区

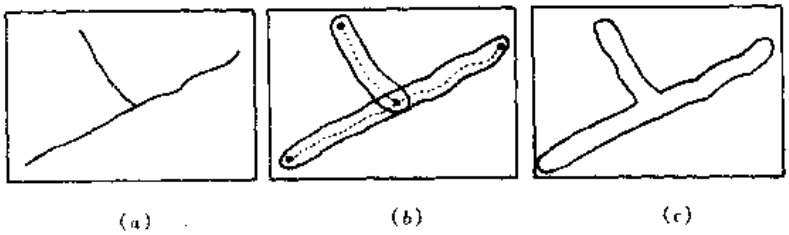
缓冲区是一些新的多边形，不包含原点、线、面要素。缓冲区的大小由缓冲宽度确定。在建立缓冲区时，应注意几个特殊问题。

1. 缓冲区发生重叠时的处理

缓冲区的重叠包括多个特征缓冲区之间的重叠（图 7—5）以及同一特征缓冲区图形的重叠（图 7—6）。对于前者，首先通过拓扑分析的方法，自动地识别出落在某个缓冲区内部的那些线段或弧段，然后删除这些线段或弧段，得到经处理后的连通缓冲区；对于后者，可通过缓冲区边界曲线逐条线段求交。如果有交点并且在该两条线段上，则记录该交点，从此点截断曲线，而线段的其余部分是否保留则应判断它位于重叠区内还是位于重叠区外。若位于区内则删除，区外则记录之，便可得到包含岛的缓冲区。

2. 特征要求缓冲区宽度不同时的处理

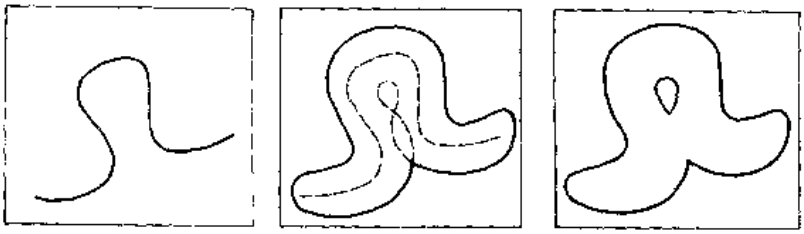
在进行缓冲区分析时，经常发生不同级别的同一类要素具有不同的缓冲区大小。例如，在城市土地地价评估时，沿主要街道两侧的通达度、繁华度的辐射范围大，而小街道则较



(a) (b) (c)

图 7－5 多个特征缓冲区图形的处理（据 ESRI，1988）

a－原图；b—缓冲区操作；c—缓冲区结果

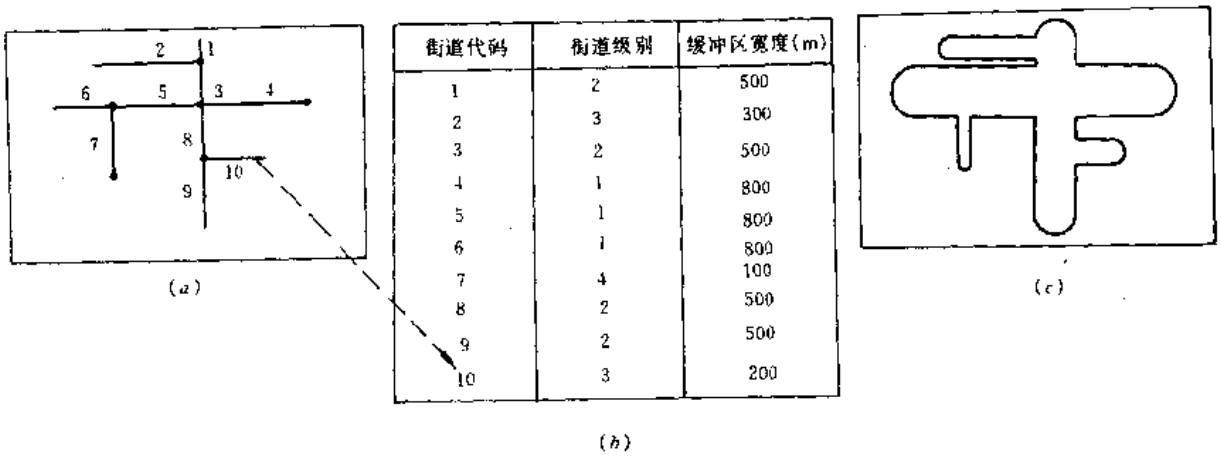


(a) (b) (c)

图 7－6 同一特征缓冲区图形的处理

a—输入数据；b—缓冲区操作；c—重叠处理后的缓冲区

小，这与要素的类型和特点有关。在建立这种缓冲区时，首先应建立要素属性表，根据不同属性确定不同的缓冲区宽度，然后再产生缓冲区（图 7—7）。



(a)

街道代码	街道级别	缓冲区宽度(m)
1	2	500
2	3	300
3	2	500
4	1	800
5	1	800
6	1	800
7	4	100
8	2	500
9	2	500
10	3	200

(b)

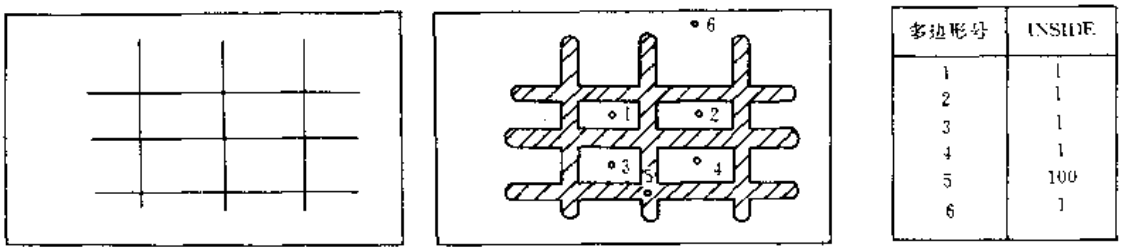
(c)

图 7—7 不同宽度的缓冲区处理

a－原图形；b—街道属性表；c—缓冲区结果

3. 复杂图形缓冲区的内外标识处理

复杂图件经缓冲区分析后产生许多多边形。图 7—8 表示在缓冲带内、外的多边形区域中，为了标识哪些区域是缓冲带内还是缓冲带外，应在这些多边形中加入特征属性。如在多边形属性表中加入 INSIDE 栏，INSIDE 值为 1 表示该多边形在缓冲区外，INSIDE 为 100，则该多边形位于缓冲带内。



多边形号	INSIDE
1	1
2	1
3	1
4	1
5	100
6	1

(c)

图 7—8　缓冲区内外多边形的标识

a—输入图形；*b*—缓冲区结果；*c*—缓冲区属性表

三、空间合成叠置分析

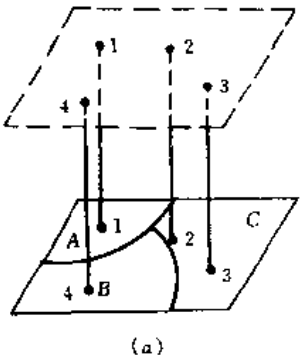
空间信息（多边形网络叠置层）的合成叠置，就是把同一地区，同一比例尺的两幅或两幅以上的图层重叠在一起，产生新的空间图形或空间位置上新的属性。因此，叠置后产生的新的图形属性就是原叠置相应位置处的图形对应属性的函数，可用下述关系式表达：

$$U=f\ (A、B、C\cdots) \qquad (7—6)$$

式中 A、B、$C\cdots$表示原叠置层的图形的属性，f 函数取决于各层上的属性与用户需要之间的关系。

（一）包含分析

包含分析主要用于确定点、线、面之间的相互联系，如确定某区域内矿井的个数，这是点与面之间的包含分析（*Point-in polygon*）（图 7—9）。再如确定某一县境内公路的类型以及不同级别道路的里程，是线与面之间的包含分析（*line-in polygon*）（图 7—10）。分析方法是：首先把这些矿井、公路等点、线要素数字化，经处理后形成具有拓扑结构的相应图层，然后和已经存放在系统中的多边形进行点与面、线与面的叠加；最后对各个多边形或区域进行这些点或线段的自动计数或归属判别。其中，点的归属判别，采用铅垂线算法（见第 6—3 节中开窗处理）；线段是由点组成的，因此可用相同的算法处理，只是此时需进行与多边形边界交点的计算，在交点处截断线段，并对线段重新编号。



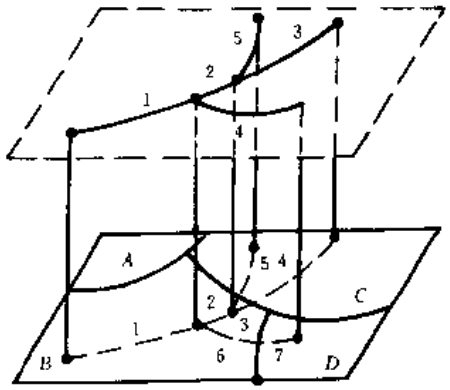
点号	多边形号
1	A
2	C
3	C
4	B

(b)

图 7—9　点包含分析

（二）多边形叠置（*overlay*）

多边形叠置方法亦称为 *polygon-on-polygon* 叠置。参加叠置分析的两个图层应都是矢量多边形结构。若需进行多层叠置，也是两两叠置后再与第三层叠置，依次类推。其中被叠置的多边形为本底多边形，用来叠置的多边形为上覆多边形，叠置后产生具有多重属性的新多边形（图 7—11）。



(a)

叠置后线号	原线号	多边形号
1	1	B
2	2	B
3	3	B
4	3	C
5	5	C
6	4	B
7	4	D

(b)

图 7—10　线包含分析

多边形叠置的目的是通过区域多重属性的模拟，寻找和确定同时具有几种地理属性的分布区域（图 7—12）；或是按照确定的地理指标，对叠置后产生的具有不同属性级的多边形进行重新分类或分级（图 7—13）。

应用实例：某矿区由于岩溶陷落柱的存在，严重影响了生产安全和机械化采煤效率。对陷落柱这种复杂地质现象的探测，宜运用综合探测技术。在综合研究遥感图像、地质资料

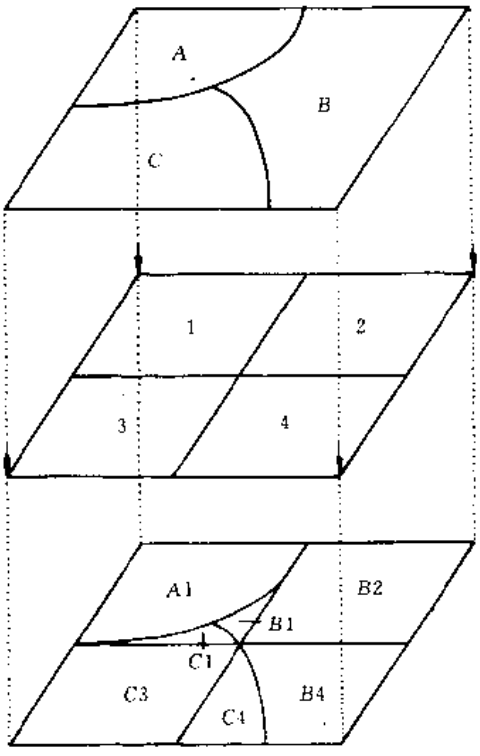


图 7—11　多边形叠置示意图

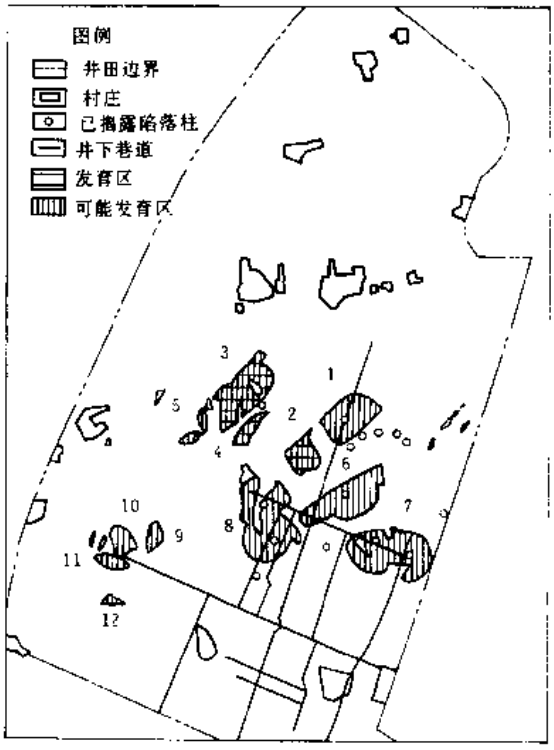


图 7—12　多边形叠置后输出的煤矿区岩溶陷落柱预测分布图

的基础上，可得到陷落柱分布有如下规律：

（1）位于断层两侧约 500m 范围内；

（2）褶皱轴线两侧约 300m 范围内；

（3）已知陷落柱分布带两侧约 300m 范围内；

（4）地表沟谷线两侧 200m 范围内；

（5）遥感图像环形或线性构造两侧 200m 范围内；

（6）热红外扫描图像热屏带区周围 100m 内。

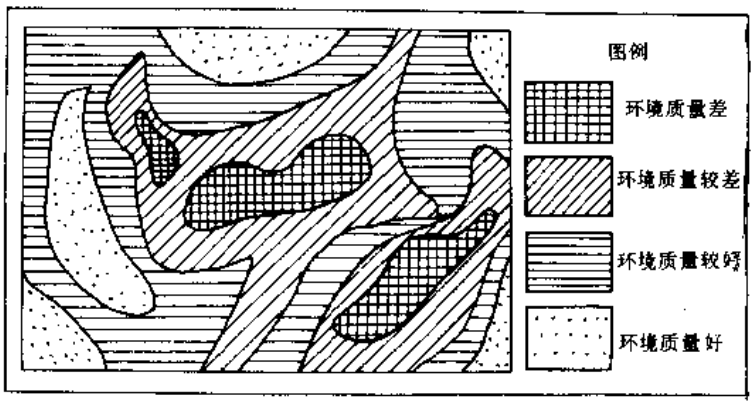


图 7—13　某市环境质量综合评价图

把以上获取的图件经数字化输入 GIS 中，经编辑及其它处理后，形成空间图层，并用缓冲区分析法建立各图层要素的缓冲区，最后进行多边形叠置分析，输出矿区陷落柱预测分布图（图 7—12）。

此外，还可进行多边形包含统计叠置（图 7—14），以精确地统计一种要素（如土地利用）在另一种要素（如行政区域）的某个区域多边形范围内的分布状况和数量特征（包括拥有的类型数、各类型的面积及其所占总面积的百分比等），或提取某个区域范围内某种专题内容的数据。这种叠置所产生的结果为统计报表。

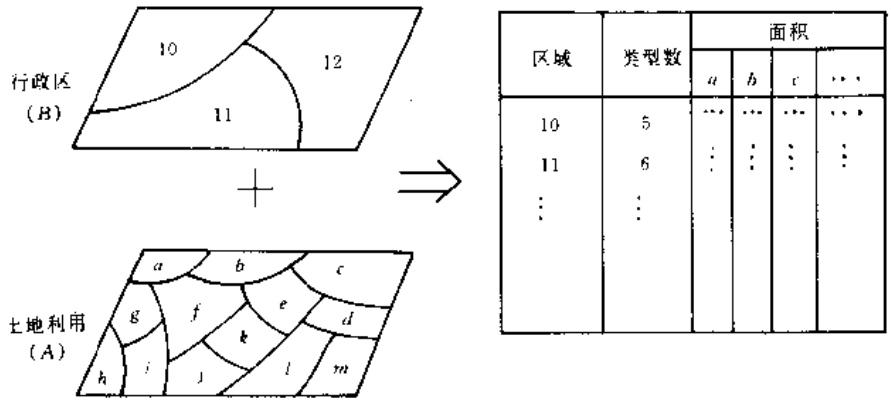


图 7—14　统计叠置示意图

（三）栅格数据的叠置分析

栅格数据的叠置（图 7—15），有三种主要变换类型可控制式（7—6）所表达的函数关系，即点变换、区域变换和邻域变换。

1. 点变换函数

点变换只需把对应的栅格单元进行属性值的运算。点变换函数不受邻域点上属性值的影响，也不受区域内一般特征的影响。其运算函数可以是加、减、乘、除、指数、对数、三角函数等。运算得到的新属性值可能与原叠置层的意义完全不同，可以是分类值、名称等，也可以是基本元素分析或集群分析的结果。栅格数据叠置所得到的新的属性值是进一步空间聚类或聚合的依据（空间聚类或聚合的原理已在综合属性分析部分介绍过）。

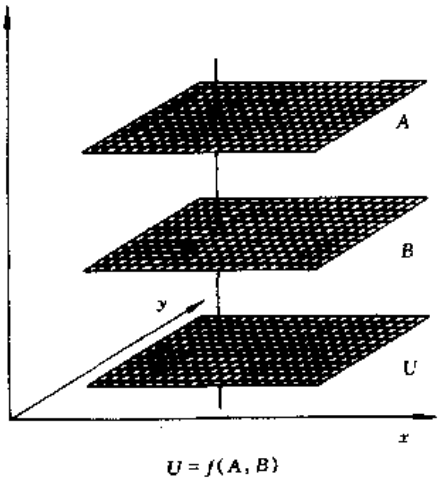


图 7—15　栅格数据的叠置

例如，土地评价涉及的多因素分析中可能包括土壤类型、土壤深度、排水性能、土壤结构、地形地貌等各层。这些层中对应栅格属性值的加权数学和逻辑运算结果，经分类后可得到土地适应性分析结果。

点变换函数除上述基本函数外，还可由基本函数组成复杂的函数式，对各分层的每一栅格单元进行叠置后再做新层属性值的计算。GIS 中一些普通的点运算函数如表 7—4 所示。

表 7—4　栅格叠置点运算函数

运算方法	关键字	修饰语	简　　述
加法	ADD	加权	各层上的属性相加得新层上相应点的值
减法	SUBTACT	加权	两层上对应点的属性相减
乘法	MULTIPLY	加权	相乘
除法	DIVIDE	加权	相除
乘方	EXPONENTIATE	加权	乘方
三角函数			变换成三角函数
逻辑函数			变换成逻辑函数
平方根			取平方根值
平方值	AVERAGE		取加权或不加权的平均值
覆盖	COVER		用第二幅地图上的非零象元覆盖第一幅地图产生第三幅
逻辑组合	CROSS		把多幅图的属性按布尔逻辑关系组合成新图上的属性
查找范围	MAXIM12E MINIM12E		找出多幅图上属性的最大或最小值作为第三图幅上的新值
提取或分离	EXTRACT		从一层上选取特定属性值或值的区间产生新层
确定频率分布	HISTOGRAM		计算直方图
频率分布与理论分布匹配	DISTRIBUTION STRETCH		重新计算象元值使其分布与理论分布一致
赋常数	CONSTANT		将一层上的所有象元赋与同一个常数
分　类	SLICE		将连续值（或离散值）按数值区间分成类别

续表

运算方法	关键字	修饰语	简　述
两层的双向比较	CROSS	平均值 总和 极大值 极小值 中值 多（少）数 差异	建立属性的交叉表格，主要是对两层比较后，以点对点为基础的统计结果产生新地图
多变量运算和类别定位	PCOMP		n层象元值的基本元素变换，并写到输出层上去
	CLUSTER		完成聚群分析
	ALLOCATE		按分类标准把象元分配到输出地图上的某类别
傅立叶变换＋滤波	FOURIER		计算二维傅立叶变换并用滤波器滤除高或低频成分

举例，若 A、B 两个栅格数据层分别为不同时期城市建成区，则两个图层的相减运算可得到这期间内城市建成区的扩展变化。

2. 区域变换函数

区域变换是指在计算新层属性值时，不仅与原叠置层对应的栅格属性值有关，也与栅格所在的区域长度、范围、周长、形状等特性有关，与某一层上等值栅格的个数有关。可能用到的区域变换函数见表 7—5。

表 7—5　栅格叠置区域变换函数

运算方法	关键字	简　述
重分类或重编码	RENUMBER	对指定的单元或图例中的值或地图上的数值范围赋予新值。此法也可用于图象反差增强
确定数量	REPORT SIZE	计算或从数据库中读取区域的相应数据并显示计算结果或写进文件
确定区域的形状	EULER	确定一定区域的形状和某一层上分离开的区域数量
按面积分类	SIZE	按计算出的每个区域的面积大小和输出地图的要求重新给予每个区域一个名称
两多边形网叠置	OVERLAY	两层多边形网络叠置产生新多边形网络（矢量或栅格均可）
按区域进行逻辑选择		与点运算相似
按区域的内容分类		与点运算相似

举例，A 表示土地利用现状图，B 表示自然区域或行政区域图，叠置后，用区域变换函数，可得到各区域内不同土地利用单元的类型和面积统计结果，如图 7—14 所示。

3. 邻域变换函数

邻域变换是指在计算新层的属性值时，不仅考虑原叠置层对应栅格本身和其它属性值

相同的区域，而且还要考虑与它属性值不同的邻域栅格在计算新层属性时的影响。与区域函数的主要区别是涉及的区域范围不是按区域边界确定，而是由指定的距离确定（图 7—16）。

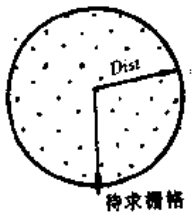


图 7—16 邻域变换函数示意图

邻域函数中较为简单的运算是加权平均值和最大最小值的计算，得到的是叠置栅格周围一定范围内的属性值统计特征量，如邻域范围内栅格属性值的变化和比率等。表 7—6 列出了一些邻域变换函数及其作用。

栅格数据叠置在方法上比矢量多边形叠置简单，但由于其数据量大，因此，关键在于如何提高其运算速度。为了压缩栅格数据量，一般采用游程编码或四叉树编码法存储栅格数据。因而需要研究相应编码法栅格数据叠置的有效算法。

表 7—6 邻域变换函数及其作用

运算方法	关键字	修饰词	简　介
邻接单元的延续	CLUMP	距离限制	建立新地图时按唯一性识别标准确定延续的单元组，一般以距离作为识别标准
	SCAN	距离，均值总和，极值中间值、偏差、方差等距离	按地图上一定距离范围内邻域单元的统计特征计算每个单元的累积值，从而建立一幅新的地图
	SPAN		按某象元到其他类别的象元之间的距离计算离散地图上每个象元的新值
找出给定目标的取值区间或距离内的位置	DSEARCH	距　离 区　间 方　向	在指定的距离和数值区间内，从连续或离散地图数据中选择出需要的点，建立新地图
连续表面函数	SLOPE	均　值 极大值	对表示连续表面的数据微分产生坡度图作为输出地图
	ASPECT		同上，但主要是计算坡度方向
	PROFILE		分析斜坡断面类型
由离散数据建立连续表面的函数	SPREAD	距　离 越　过 向　上 向下穿过	按到起始点或起始线的距离值重新编码产生连续地图 控制延伸方向，即用连续或者离散地图来建立延伸通道 基底地图也可能包括某些障碍物等
等值化	CONTOUR		从栅格数据建立等值线图
内　插	TREND		计算趋势面和地图（这是连续的叠置层）
	INVERSE		计算每个象元的加权平均值，作为连续层，用协方差函数进行内插
	OPTIMAL		
搜寻最短通　道	VIEW		根据某层上的目标点向四周的连续三维表面观察能否看到的条件把其他象元（点）分成两类（能看到和不能看到）
	STREAM		搜寻连续表面上两点间的最短路线
	PATHWAY		按时间、费用、距离等在网络系统（如道路网络系统）中找出最短通道
点在多边形中	PPOINT		找出包含指定点的多边形

四、网络分析

（一）基本概念

网络是一系列相互联结的弧段，形成物质、信息流通的通道。例如：水从水库流向各种水渠；从发电厂经电网向用户供电；城市的道路网均构成网络，如图 7—17 所示。网络是现代生产、生活必不可少的条件。

网络分析的主要用途是：选择最佳路径；选择最佳布局中心的位置。所谓最佳路径是指从始点到终点的最短距离或花费最少的路线（图 7—18）；最佳布局中心位置是指各中心所覆盖范围内任一点到中心的距离最近或花费最小（图 7—19）；网流量是指网络上从起点到终点的某个函数，如运输价格，运输时间等。网络上任意点都可以是起点或终点。

一个网络由如下一些基本要素组成（图 7—20）：

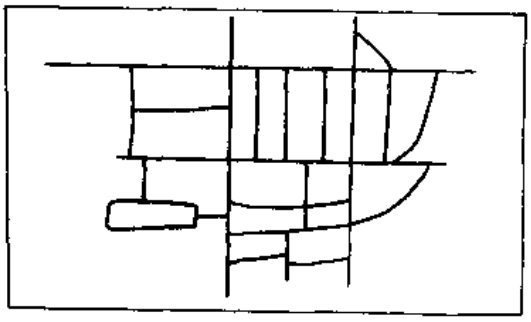


图 7—17　城市道路网络结构示意图

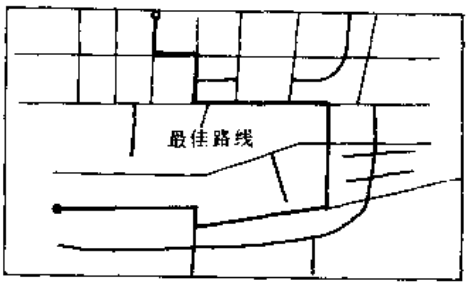


图 7—18　选择从始点到终点的最佳路线

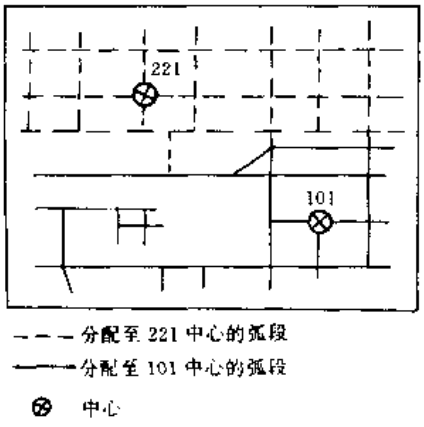


图 7—19　选择最佳布局中心

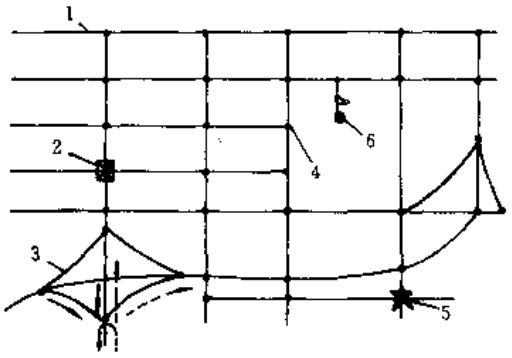


图 7—20　网络的基本要素

1—连通路线；2—障碍；3—转弯；4—结点；5—停靠点；6—中心

1）结点　网络中任意两条线段的交点。

2）连通路线或链　连结两个结点的弧段要素，是网络中资源运移的通道，与结点一起，构成了网络中的最基本要素。链间的相互联系在 GIS 中应具有拓扑结构。

3）转弯　在连通路线相连的结点处，资源运移方向可能转变，运移方向从一个链上

经结点转向另一个链。特定方向的转弯通常限制了资源在网络中的运移。例如在道路网中的高架桥使得车辆不能向左或向右拐弯。

4）停靠点　网络路线中资源装、卸的结点点位，如邮件投放点、公共汽车站等。

5）中心　网络线路中具有接收或发放资源能力，且位于结点处的设施。如水库具有调节各支流的水量并能向各渠道开闸放水的能力。

6）障碍　资源不能通过的结点。

上述要素除障碍和结点之外，都用图层要素形式表示，并用一系列相关属性来描述，一般以表格的方式存储在GIS系统数据库中，以便构造网络模型和网络分析。这些属性是网络中的重要部分。例如，在城市交通网络中，每一段道路（链）都有名字、速度上限、宽度等；停靠点处有大量的物资等待装载或下卸等属性。

在这些属性中，有三个重要的概念：

1. 阻强

阻强是指资源在网络中运移阻力的大小。它是描述链与拐弯所具有的属性。链弧的阻强是指从链的一个端点至另一个端点所需克服的阻力，如链弧段的长度可作为阻强的描述参数，因为物资在长链弧上运移花费的时间比短链弧上要多。阻强的大小应根据多种因素来确定，如弧段的特性，网络中运移资源的种类、运移的方向，弧段中的特殊情况等。

转弯的阻强描述了从一条链弧经结点到另一条链弧的阻力大小，它随着两相连链弧的条件状况而变化。

为了便于分析计算，不同类型的阻强都应使用同一种量纲。

运用阻强概念的目的在于模拟真实网络中各路线及转弯的变化条件。网络分析中选取的最佳布局中心和最佳路线随要素阻强的大小而变化。最佳路线是最小阻力的路线。对不构成通道的弧段或转弯往往赋以负的阻强，这样在选取最佳路线时可自动跳过这些弧段或转弯。

2. 资源需求量

资源需求量系指网络中与弧段和停靠点相联系资源的数量。如在供水网络中每条沟渠所载的水量；在城市网络中沿每条街道所住的学生数；在停靠点装卸货物的件数等。

3. 资源容量

资源容量系指网络中心为了满足各弧段的需求，能够容纳或提供的资源总数量。如学校的容量指学校能注册的学生总数；停车场能停放机动车辆的空间；水库的总容量等等。

（二）网络要素及其属性表示

1. 链弧的属性表示

链弧是有向线段，除用拓扑关系描述外，还有相应的属性如阻强、需求量等（图7—21）。其中需求量对于选择最佳布局中心及网流量计算是必不可少的属性值。

2. 转弯及其属性表示

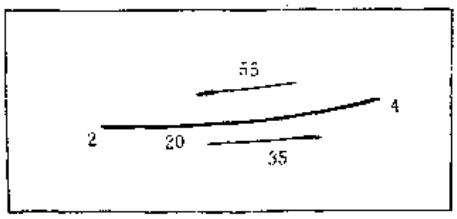
在网络结点处，可能产生的转弯个数为

$$N=m^2 \qquad (7—7)$$

式中 m 为在结点处相连的弧段条数。当3条弧段相连时，转弯的个数为9（图7—22）。

在转弯处往往有一些限制，对不同限制类别及其属性的描述见表7—7。

3. 停靠点、中心的属性



(a)

属性表

从结点	至结点	长度(km)	链弧号	正方向阻强(km/h)	反方向阻强(km/h)	资源需求量
2	4	175.50	20	35	55	5

(b)

图 7—21　链弧的属性表示



图 7—22　转弯的类型

这两种网络要素的属性表示法很简单。停靠点仅在选择最佳路线时使用，其属性为资源需求量，正值表示装载，负值表示下卸（表 7—8）

中心仅在寻求最佳状况时使用，其属性是资源最大容量、服务范围（从中心至各可能路径的最大可能距离）和服务延迟数（在其它中心的服务量达到该数量时开始服务）。

（三）选取最佳路径

在网络数据中分析从起点经一系列特定的结点至终点的资源运移最佳路线，即受阻力最小的路径。通俗的表达方式是，怎么在最短的时间内，交通工具通过最近的路段，绕过障碍而到达目的地？目的地如一些服务中心、消防队、急救中心等。GIS 通过共享数据库属性数据模型进行模拟、分析和判断，迅速显示出最佳路线。其算法一般采用比较法。

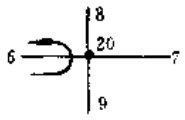
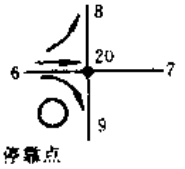
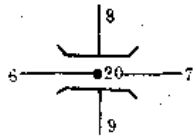
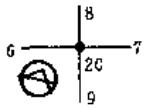
设从起点 P_1 至终点 P_2 的路径有几条，则应有

$$IMPED=\min\ [IMPED_i\ (P_1,\ P_2)] \quad i\leqslant n \qquad (7—8)$$

式中 $IMPED$ 为最小阻强，$IMPED_i\ (P_1,\ P_2)$ 为第 i 种路径所受的阻强，则 $IMPED$ 所对应的路径为最优路径。

举例：在图 7—23 所示的网络中，从起点 P_1 至终点 P_2 的路径有 3 条：第 1 条是经 10 和 11 号链弧，其阻强 $IMPED_1\ (P_1,\ P_2)=3+1+2=6$；第 2 条是经 20 与 21 号链弧，其阻强 $IMPED_2\ (P_1,\ P_2)=2+2-4=8$；第 3 条是直接从 P_1 点至 P_2 点的 15 号链弧，其阻强 $IMPED_3\ (P_1,\ P_2)=7$；因此，取 $IMPED=6$。第 1 条路线是最优路径，在显示网络时，予以特别显示。

表 7—7　转弯的类型及其属性描述

转弯类型	描　述	属性表	0=无阻强		−1=不允许拐弯	
		结点号	从弧段	至弧段	角度	时间阻强(s)
U 型拐弯 	U 型拐弯指从 6 号弧至 20 号结点并从 20 号结点转回 6 号弧，这是一个 180 度转弯，花费 20 秒时间	20	6	6	180	20
	停靠点使得从 6 号弧至其它弧段——直通 7 号弧，向左转至 8 号弧，向右转至 9 号弧的运移减慢	20	6	7	0	15
		20	6	8	90	20
		20	6	9	−90	10
高架道或地道 	高架道或地道允许直通而无延迟，如从 6 号弧至 7 号弧；但不允许转弯，此时以负的阻强表示(如从 6 号弧至 8、9 号弧)	20	6	7	0	0
		20	6	8	90	−1
		20	6	9	−90	−1
		20	9	8	0	0
		20	9	7	−90	−1
		20	9	6	90	−1
不准右转弯 	不允许从 6 号弧转向 9 号弧，并赋于负值阻强；允许其它方向的转弯，其阻强为正	20	6	9	−90	−1
		20	6	7	0	5
		20	6	8	90	10

（四）选择最佳布局中心

在网络中，中心点与网络各路径的关系是固定的。选择最佳布局中心的目的是把所有链弧都分配到某一中心，并把中心的资源分配给这些链弧以满足其需求，也即既要满

表 7—8　停靠点与其属性

结点号	需求量
546	3.5
547	−2.0

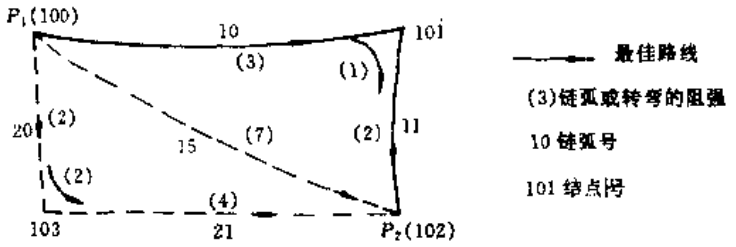


图 7—23　最佳路径选取分析实例

足需要，又不能浪费中心的资源。GIS 中网络分析功能通过对比的方法，模拟分析中心覆盖范围和服务对象数量，筛选出最佳布局和布局中心的位置。对比条件包括中心的性质、网络覆盖范围及网络状况描述等属性数据。在模拟和分析过程中，也可以采取人机对话的方式，由用户设置模拟条件，以使中心保证最大覆盖范围，并且为用户提供最佳服务。

图 7—24 表示有阻强链弧的最佳布局算法。有 3 条弧段已分配到相应的中心，而弧段 15 尚待分配。分配的原则是已有路径的阻强总数与该弧阻强之和作比较，选取阻强最小的路径，该路径相连的中心为最佳中心。15 号弧的分配方案与阻强如下：

至 101 点：$IMPED=1+2=3$

至 120 点：$IMPED=1+2+2=5$

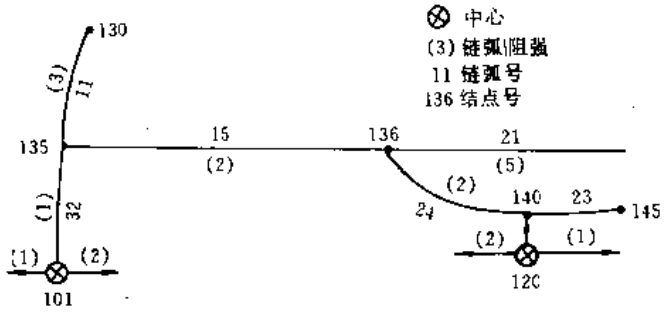


图 7—24　弧段分配到最佳中心

因此，应把该弧分配到 101 中心去。假如两个中心的阻强总和相等，则可任意取一中心，或参考其它条件取一。

把所有链弧分配到所有对应中心的操作应同时有序地进行，但发生下列情况之一时则应停止分配：

（1）沿路径的累积阻强已达到中心的服务范围；

（2）沿路径的积累资源需求量已达到中心的总资源量；

（3）网络中所有链弧均已分配到最近中心。

下面举一网络分析应用的实例。

1986 年美国加州环境研究所承担了 *Redland* 市一所私立小学改址新建的项目。校方提出了两处新校址方案，要求使用网络分析以帮助决策。这是一个网络分析的最佳布局中心选择的问题。

首先，系统地统计和编录每个学生从住址到旧校址所需的时间，并将统计结果输入系统数据库。在数据库中已经建立了该市街道的网络模型。在相应的路段注明学生上学所需使用的平均时间。街道口标以学生过路口使用的平均时间（时间长短取决于街道宽度和街口类型，是右转弯或左转弯，一般右转弯比左转弯使用的时间长）。按规定小学生不准过高速公路，在这些路上设置有障碍。

然后，将学生的家庭住址、旧校和新校址的位置与网络图形坐标配准，按路段把学生的统计数据作为数据库属性数据。属性数据与网络图形数据自动连接。在此基础上，用系统中最佳布局中心选择程序模块，对所有学生从家到旧校址使用的平均时间和最长时间进行分析和判断。最终结果表明，旧校址的位置是最佳位置。这个结果也证明系统的分析和判断模型是可靠的。

最后，还使用该网络模型对新校址进行了模拟分析。首先对几个学生从家到新校址使用的时间进行测定，测定的时间数作为全部学生的平均时间。对两个新校址分别进行分析后，对所有学生上学时间作了评价（表 7—9）。

项目对保留旧校址，再建立一所新校的方案进行了模拟分析和实验。条件是旧校址不

表 7- 9 学生从家到旧址和新址所需时间

校　址	平均时间（分钟）	最多时间（分钟）
旧 校 址	4.8	7.6
东区新校址	5.1	9.2
西区新校址	5.5	9.9

变，模型中只改变新校的位置。有两种组合方案，第一组合：旧校与东区新校；第二组合：旧校和西区新校。模型的结果是上学的最短时间和到各校注册的学生数（表 7—10）。其中第一组合方案的优点是可以减少多数学生上学途中所耗费的时间，比第二种组合好。

表 7—10 两种组合方案（学生上学时间和可能注册生数）

组合方案	平均时间（分钟）	注册学生数量	
		旧　校	新　校
旧校和东区新校	4.1	84	54
旧校和西区新校	4.4	90	48

综合考虑各种其它因素（如路面、上下班高峰和学校最大容量）并加入属性数据库中，再进行分析，最后选择既能减少多数学生上学路上花费的时间，又满足新、旧校均有足够生源的最佳方案（表 7—11）。

表 7—11 最佳组合方案（学生上学时间与注册学生数）

最佳组合方案	平均路上花费时间（分钟）	注册学生数量	
		旧　校	新　校
旧校和东区新校	3.9	113	25

第 7—2 节　数字地面模型（DTM）与地形分析

一、DTM 的概念与作用

数字地面模型（Digital Terrain Model，DTM）是描述地面诸特性空间分布的有序数值阵列，它可用二维区域上的一个有限项的向量序列来表示，即

$$K_P = f_k\ (u_p,\ v_p) \qquad (k=1,\ 2,\ 3\cdots m,\ p=1,\ 2,\ 3\cdots n) \qquad (7—9)$$

式中，K_P 为第 p 号地面点（可以是单一的点，但一般是某点及其微小邻域所划定的一个地面单元）上的第 k 类地面特性信息的取值；u_p，v_p 为第 p 号地面点的二维坐标，可以采用任一地图投影坐标，如经纬度，矩阵的行列号等；m（$m \geqslant 1$）为地面特征信息类型的数目；n 为地面点的个数。

在许多情况下，所记的地面特性是高程 z，它的空间分布由 x、y 平面坐标系统来描述，也可用径度 λ 和纬度 φ 来描述海拔 h 的分布，对于这种 DTM 也称作数字高程模型（Digital Elevation Model，DEM）。在本书中，如不作特别强调，DTM 所指的即为 DEM。

数字地面模型的用途广泛，主要包括：

（1）在国家地形数据库中存储数字地形图的高程数据；

(2) 在道路等工程设计中计算挖填土石方量；

(3) 显示三维地面景观；

(4) 线路、坝址的自动选定，以及库容、淹没损失的计算；

(5) 与遥感图像复合，以提高分类精度和进行图像几何畸变的校正；

(6) 地形因子（坡度、坡向、粗糙度等）的自动提取，进行地貌分析；

(7) 地表形态的自动分类；

(8) 不同地面的比较和统计分析；

(9) 越野通视情况分析（为军事和土地景观规划等目的服务）。

二、DTM 的数据采集与表示

(一) DTM 的数据源与采集方法

1. 以航空或航天遥感图像为数据源

这种方法是由航空或航天遥感立体像对，用摄影测量的方法建立空间地形立体模型，量取密集数字高程数据，建立 DTM。采集数据的摄影测量仪器包括各种解析的和数字的摄影测量与遥感仪器。

摄影测量采样法还可以进一步分成：

1) 选择采样　在采样之前或采样过程中选择所需采集高程数据的样点（地形特征点，如断崖、沟谷、脊等）。

2) 适应性采样　采样过程中发现某些地面没有包含必要信息时，取消某些样点，以减少冗余数据（如平坦地面）。

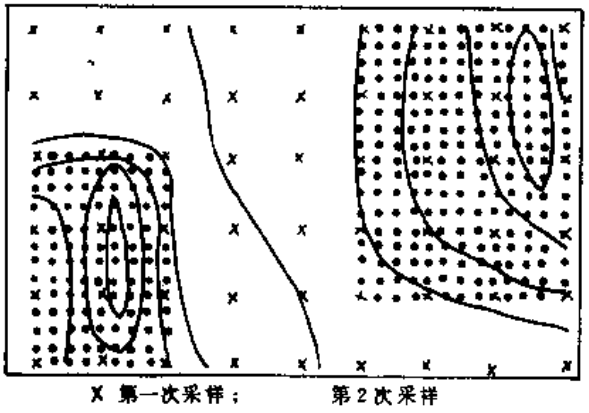


图 7—25　先进采样法

3) 先进采样法　采样和分析同时进行，数据分析支配采样过程。先进采样在产生高程矩阵时能按地表起伏变化的复杂性进行客观、自动地采样。实际上它是连续的不同密度的采样过程，首先按粗略格网采样，然后在变化较复杂的地区进行细格网（采样密度增加一倍）采样（图 7—25）。由计算机对前一次采样获得的数据点进行分析后，再决定是否需要继续作高一级密度的采样。

计算机的分析过程是，在前一次采样数据中选择相邻的 9 个点作窗口，计算沿行或列方向邻接点之间的一阶和二阶差分。由于差分中包含了地面曲率信息，因此可按曲率信息选取阈值。如果曲率超过阈值时，则必需进行另一级格网密度的采样。

2. 以地形图为数据源

主要以比例尺不大于 1∶1 万的国家近期地形图为数据源，从中量取中等密度地面点集的数字高程，建立 DTM。其方法有下列几种：

1) 手工方法　采用方格膜片、网点板或带刻划的平移角尺叠置在地形图上，并使地形图的格网与网点板或膜片的格网线逐格匹配定位，自上而下，逐行从左到右量取高程。当格网交点落在相邻等高线之间时，用目视线性内插方法估计高程值。这种方法获取的是规

则的格网DTM，它的优点是几乎不需要购置仪器设备，而且操作简便。

2）手扶跟踪数字化仪采集　采集方式有：沿主要等高线采集平面曲率极值点，并选采高程注记点和线性加密点作补充；逐条等高线的线方式连续采集样点，并采集所有高程注记点作补充，这种方式适用于等高线较稀疏的平坦地区；沿计曲线和坡折线采集曲率极值点，并补采峰-鞍线和水边线的支撑点，分别以等高线，峰-鞍链和边界链格式存储。

3）扫描数字化仪采集　这种方式采集速度最快，但目前仅能以扫描分版等高线图方式采集高程。随着研究的不断深入，一些难点和瓶颈问题被解决，从地图扫描数据中自动地建立DTM技术必将达到实用水平。

3. 以地面实测记录为数据源

用电子速测仪（全站仪）和电子手簿或测距经纬仪配合PC-E500等袖珍计算机，在已知点位的测站上，观测到目标点的方向、距离和高差三个要素。计算出目标点的x、y、z三维坐标，存储于电子手簿或袖珍计算机中，成为建立DEM的原始数据。这种方法一般用于建立小范围大比例尺（比例尺大于1∶5000）区域的DTM，对高程的精度要求较高。另外气压测高法获取地面稀疏点集的高程数据，也可用来建立对高程精度要求不高的DTM。

4. 其它数据源

采用近景摄影测量在地面摄取立体像对，构造解析模型，可获得小区域的DTM。此时，数据的采集方法与航空摄影测量基本相同。这种方法在山区峡谷、线路工程和露天矿山中有较大的应用价值。

另外，航空测高仪可获得精度要求不太高的高程数据，也可以依此来构造DTM。

（二）DTM的表示方法

1. 数学分块曲面表示法

这种方法把地面分成若干个块，每块用一种数学函数，如傅立叶级数多次多项式、随机布朗运动函数等，以连续的三维函数高平滑度地表示复杂曲面，并使函数曲面通过离散采样点。这种近似数学函数表示的DTM不太适合于制图，但广泛用于复杂表面模拟的机助设计系统。

2. 规则格网表示法

规则格网表示方法是把DTM表示成高程矩阵

$$\mathrm{DTM}=\{H_{ij}\},\ i=1,\ 2,\ 3\cdots m-1,\ m,\ j=1,\ 2,\ \cdots n-1,\ n \qquad (7—10)$$

此时，DTM来源于直接规则矩形格网采样点或由规则或不规则离散数据点内插产生。由于计算机对矩阵的处理比较方便，特别是以栅格为基础的GIS系统中高程矩阵已成为DTM最通用的形式。高程矩阵特别有利于各种应用，但规则的格网系统也有下列缺点：

（1）地形简单的地区存在大量冗余数据；

（2）如不改变格网大小，则无法适用于起伏程度不同的地区；

（3）对于某些特殊计算如视线计算时，格网的轴线方向被夸大；

（4）由于栅格过于粗略，不能精确表示地形的关键特征，如山峰、洼坑、山脊、山谷等。

为了压缩栅格DTM的冗余数据，可采用游程编码或四叉树编码方法。

3. 不规则三角网（TIN）表示法

不规则三角网（Triangulated Irregular Network，TIN）是专为产生DTM数据而设计

的一种采样表示系统。它克服了高程矩阵中冗余数据的问题，而且能更加有效地用于各类以 DTM 为基础的计算。因为 TIN 可根据地形的复杂程度来确定采样点的密度和位置，能充分表示地形特征点和线，从而减少了地形较平坦地区的数据冗余。TIN 表示法利用所有采样点取得的离散数据，按照优化组合的原则，把这些离散点（各三角形的顶点）连接成相互连续的三角面（在连接时，尽可能地确保每个三角形都是锐角三角形或是三边的长度近似相等），如图 7—26 所示。

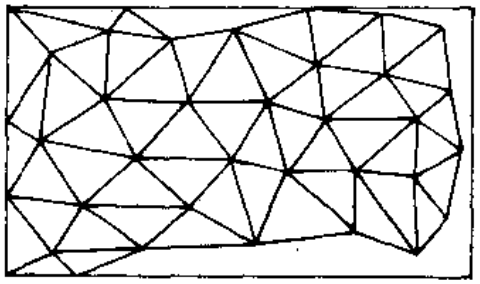


图 7—26　不规则三角网法构成的 DTM

在概念上，TIN 模型类似于多边形网络中的矢量拓扑结构，只是 TIN 中不必要规定“岛屿”或“洞”的拓扑关系。TIN 把结点看作数据库中的基本实体。拓扑关系的描述，是在数据库中建立指针系统来表示每个结点到邻近结点的关系、结点和三角形的邻里关系，列表是从每个结点的北方向开始按顺时针方向分类排列的（图 7—27）。TIN 模型区域之外的部分由拓扑反向的虚结点表示，虚结点说明该结点为 TIN 的边界结点，使边界结点的处理更为简单。

在图 7—27 中，给出了 TIN 网络模型数据结构中的一部分，包括 4 个结点和 2 个三角形，数据则由结点列表、指针列表和三角形列表组成。区域中包括了边界结点，故设置虚指针。

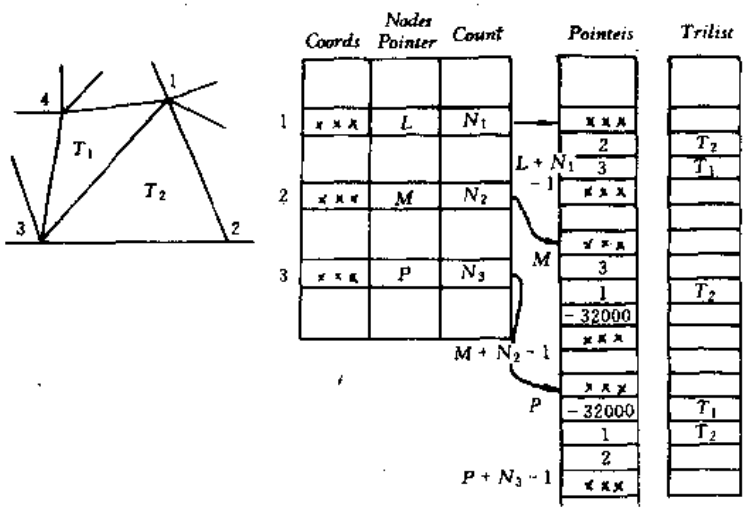


图 7—27　TIN 的数据结构模型

由于结点列表和指针列表包含了各种必要的信息和连接关系，因而可以完成 DTM 数值的有效提取和计算。

三、DTM 的空间内插方法

DTM 空间内插的概念十分简单，即在一个由 x、y 坐标平面构成的二维空间中，由已知若干离散点 P_i 的高程，估算待内插点的高程值。由于 DTM 采样的数据点呈离散分布形

式，或是数据点虽按格网排列，但格网的密度不能满足使用的要求，这就需要以数据点为基础进行插值运算。

DTM 内插按插点分布范围，可分为分块内插、剖分内插和单点移面内插三类，见图 7—28。

分块内插，是把需要建立 DTM 的地区，切割成一定大小的规则方块，形状通常为正方形。它的尺寸应根据地形复杂程度和数据源的比例尺确定。在每一个分块上展铺一张数学面，相邻分块之间有适当宽度的重叠带，以使重叠带内全部数据点成为相邻块展铺数学面时的共用数据，保证一张数学面能够较平滑地与相邻分块的数学面拼接（图 7—29）。这种内插方法的优点是可以得到光滑连续的空间曲面。

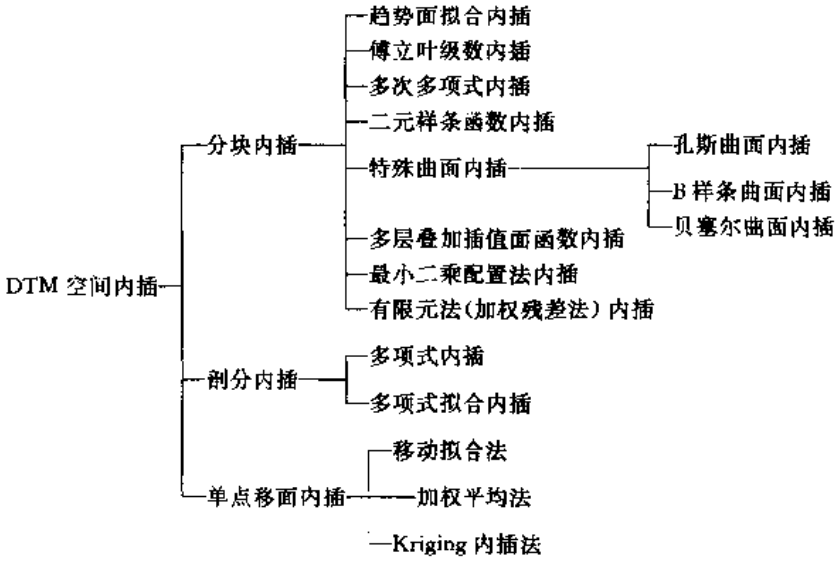


图 7—28　DTM 空间内插方法分类

剖分内插是把需要建立 DTM 的地区切割成大小和形状不同的子区（剖分），子区间拥有公共边但不重叠，在该子区内展铺一个数学面，内插剖分区内任意点的高程。该法只在剖分间边界端点处重合，通常没有严格重合的边界，所以既不连续，也不光滑。剖分多边形的顶点都是数据点，最常见的数据点个数为 3，与 TIN 结构相同（图 7—30）。

单点移面内插是以待插点为中心，以适当半径或边长的圆或正方形作为移动面去捕捉适当数目的数据点，并以此展铺一张数学面，内插该中心的高程（图 7—31）。

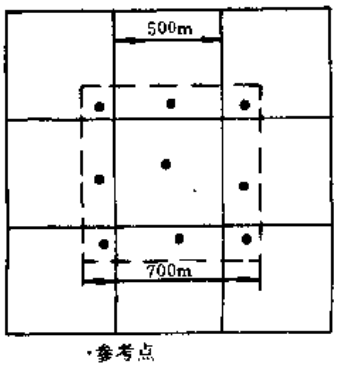


图 7—29　分块内插区域

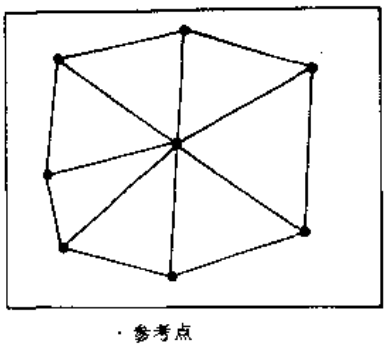


图 7—30　剖分内插区域

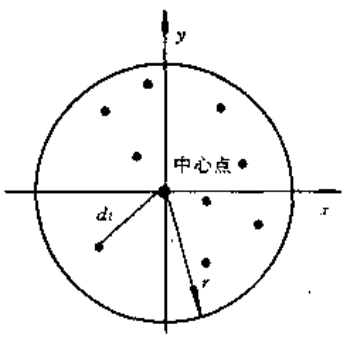


图 7—31　单点移面插值区域

下面介绍几种常用的内插方法。

1. 二元多项式拟合内插

用二元多项式来拟合地形表面，则有

$$z=f(x,\ y)=\sum_{i=0}^{m}\sum_{j=0}^{m}C_{ij}x^iy^j\quad(m>3)\tag{7—11}$$

在满项时，多项式的项数 $N=(m+1)^2$，缺项时 $N=(m+1)^2-N'$ $(N'\geqslant 1)$。如果在内插区域内，数据点个数 n 大于多项式的项数 k，这时可利用 n 个数据点的三维坐标 $(x_k,\ y_k,\ z_k)$ 以及逼近面和实际地面在数据点处的高程较差 V_k $(k=1,\ 2\cdots n)$，列出 n 个误差方程，按最小二乘法解算出多项式的 N 个系数，即

$$\underset{n\times 1}{V}=\underset{n\times m^2m^2\times 1}{A}\ \underset{}{C}-\underset{n\times 1}{Z}\tag{7—12}$$

按 $[VV]=\min$ 的最小二乘法原理，解得

$$C=[A^TA]^{-1}[A^TZ]\tag{7—13}$$

将所求参数代入二元高次多项式，可得二元高次多项式曲面拟合方程。若要求内插区任一点 P 的高程，只需把 $(x_p,\ y_p)$ 代入该方程，就可求出 Z_p。

这种内插的特点是拟合曲面不通过所有数据点，而是取得最靠近数据点的光滑曲面，以保证邻块间的光滑连续拼接。

2. 二元样条函数内插

二元样条函数是在分块范围内，按一定规则，用相邻数据点连线将块分割成若干个多边形分片（当数据点组呈正方形格网结点分布时，各分片是大小相等的正方形），通过每一分片上的全部数据点，展铺一张光滑的数学曲面，并使相邻分片间保持连续光滑的拼接。

对于 DTM 的内插，一般采用二元三次样条函数

$$z=f(x,\ y)=\sum_{i=0}^{3}\sum_{j=0}^{3}C_{ij}x^iy^j\tag{7—14}$$

写成矩阵形式为

$$z=(1,\ x\ x^2\ x^3)\begin{pmatrix}C_{00}&C_{01}&C_{02}&C_{03}\\C_{10}&C_{11}&C_{12}&C_{13}\\C_{20}&C_{21}&C_{22}&C_{23}\\C_{30}&C_{31}&C_{32}&C_{33}\end{pmatrix}\begin{pmatrix}1\\y\\y^2\\y^3\end{pmatrix}\tag{7—15}$$

设分块范围内的数据点按单位边长正方形格网结点排列，一个单位边长的正方形为一个分片（图 7—32）。取分片的左下角点为该分片平面直角坐标系统的原点，分片内任一点 P 的平面直角坐标为 $0\leqslant x_p\leqslant 1$，$0\leqslant y_p\leqslant 1$。为了保证展铺的曲面在相邻分片上连续且光滑，必须满足弹性材料的力学条件，即

(1) 相邻分片拼接处在 x 和 y 轴方向的斜率都应保持连续；

(2) 相邻分片拼接处的扭矩连续。

拼接后整个分块的逼近面，就是二元三次样条函数曲面。

由于每个分片仅有 4 个格网结点信息 $(x,\ y,\ z)$，只能列出 4 个方程式，而函数的待定系数为16个，因此其余12个方程只能根据上述力学条件建立。据此建立的12个线性方程中，要用到沿 x 轴方向的斜率 R，沿 y 轴方向的斜率 S 以及扭矩 T，它们可由下式求得：

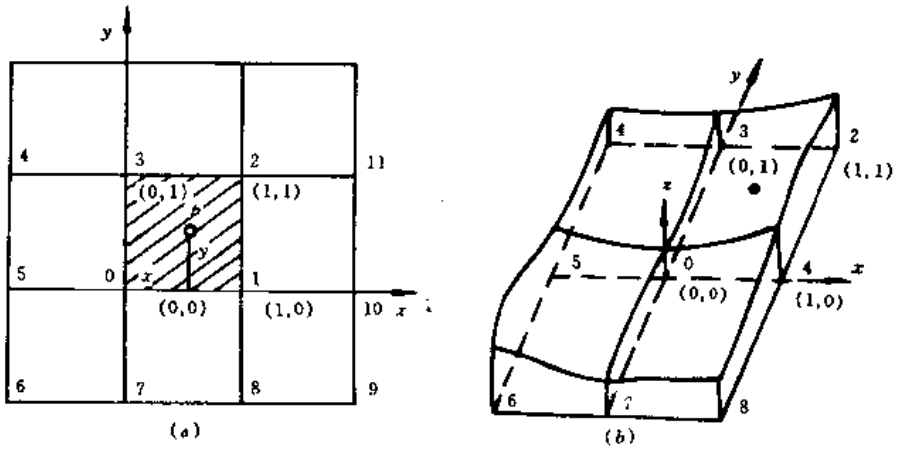


图 7—32　样条函数内插

$$R=\frac{\partial z}{\partial x}$$

$$S=\frac{\partial z}{\partial y} \tag{7—16}$$

$$T=\frac{\partial^2 z}{\partial x\partial y}=\frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right)$$

对于图 7—32 中画有阴影线的分片，其 4 个数据点的三维坐标分别是：0 (0，0，Z_0)，1 (1，0，z_1)，2 (1，1，z_2)，3 (0，1，z_3)。以 0 点为例，所建立的 4 个方程为

$$Z_0=C_{00}$$

$$R_0=\left.\frac{\partial z}{\partial x}\right|_0=\frac{z_1-z_5}{2}$$

$$S_0=\left.\frac{\partial z}{\partial y}\right|_0=\frac{z_3-z_7}{2}$$

$$T_0=\frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right)=\frac{1}{2}\left(\frac{z_2-z_4}{2}-\frac{z_3-z_6}{2}\right)=\frac{1}{4}\left[(z_6+z_2)-(z_5+z_4)\right]$$

按同样的方法可建立分别以 1、2、3 为顶点的共 12 个上述类型的方程。

用分片四个角点的高程，以及由各相关数据点高程计算得到的两个方向的斜率和扭矩数值组成一个 4×4 的常数矩阵 A

$$A=\begin{bmatrix} z_0 & S_0 & z_3 & S_3 \\ R_0 & T_0 & R_3 & T_3 \\ z_1 & S_1 & z_2 & S_2 \\ R_1 & T_1 & R_2 & T_2 \end{bmatrix}$$

按照对斜率 R、S 和扭矩 T 的定义，以及二元三次样条函数定义，得

$$R=\frac{\partial z}{\partial x}=(0\ 1\ 2x\ 3x^2)\ C\ (1\ y\ y^2\ y^3)^T$$

$$S=\frac{\partial z}{\partial y}=(1x\ x^2\ x^3)\ C\ (0\ 1\ 2y\ 3y^2)^T$$

$$T=\frac{\partial^2 z}{\partial x\partial y}=(0\ ,\ 1\ 2x\ 3x^2)\ C\ (0\ 1\ 2y\ 3y^2)^T$$

把分片的 4 个角点的平面直角坐标代入该点，得

$$\begin{bmatrix} z_0 & S_0 & z_3 & S_3 \\ R_0 & T_0 & R_3 & T_3 \\ z_1 & S_1 & z_2 & S_2 \\ R_1 & T_1 & R_2 & T_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix} C \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix}^T$$

写成紧凑矩阵形式

$$A=XCY^T$$

解此方程，有

$$C=X^{-1}A\ (Y^{-1})^T$$

把解得的系数阵代入式（7—14），则建立了二元三次样条函数式。对于分片中任意点 P，把它的平面直角坐标（x_P，y_P）代入，就可求出其高程 Z_{P0}

3. 移动拟合法内插

移动拟合法是典型的单点移面内插方法。对每一个待定点取一个多项式曲面来拟合该点附近的地形表面。此时，取待定点作为平面坐标的原点，并用待定点为圆心，以 r 为半径的圆内诸数据点来定义函数的待定系数（图 7—31）。

设取二次多项式来拟合数据，则待求点的高程可写成下列的一般式：

$$z_P=Ax^2+Bxy+Cy^2+Dx+Ey+F \qquad (7—17)$$

把坐标原点平移到待定点处，则 $\bar{x}=x-x_P$，$\bar{y}=y-y_P$，并代入上式，得移动拟合法二次多项式公式为

$$z_P=a\bar{x}^2+b\bar{x}\bar{y}+c\bar{y}^2+d\bar{x}+e\bar{y}+f \qquad (7—18)$$

为了求取式（7—18）中的待定系数，应以 x_P 为圆心，以 r 为半径作圆（r 的大小根据落入圆内的数据点个数 n 来确定，$6<n<20$），圆内的数据点均被采用，则可建立误差方程式

$$V_i=a\bar{x}_i^2+b\bar{x}_i\bar{y}_i+c\bar{y}_i^2+d\bar{x}_i+e\bar{y}_i+f-\bar{z}_i\ \ (i=1,\ 2\cdots n) \qquad (7—19)$$

式中（$\bar{x}_i$，$\bar{y}_i$，$\bar{z}_i$）为坐标原点平移至 P 点后，各数据点的三维坐标。

在求解系数时，可根据数据点至待定点的距离给予适当的权。权的给定按下式进行：

$$w_i=\frac{1}{d_i} \qquad (7—20)$$

或

$$w_i=\left(\frac{r-d_i}{d_i}\right)^2 \qquad (7—21)$$

式中 $d_i=\sqrt{\bar{x}_i^2+\bar{y}_i^2}$。

根据最小二乘法原理，按 $v\omega v^T=\min$ 的方法，建立法方程式，求解各待定系数。把系数代入插值公式，就能求得待定点高程 z_P。

4. 加权平均内插法

加权平均法是最简单的单点移面内插方法，它是搜索区域内的高程数据点，并求得加权平均值作为待定点的高程值。一般情况下，所考虑的权仅是距离的单调下降函数。为了提高插值精度，还应考虑数据点的分布方向，权衡搜索圆内所有点的方向和距离的分布情况赋权。搜索圆内数据点一般为 4 个，最多 10 个，这由调节搜索圆半径 r 的方法确定（图 7—31）。美国地质调查局推荐的考虑了方向和加权的内插公式为

$$
\begin{aligned}
&D_{xi}=x_P-x_i,\quad D_{yi}=y_P-y_i\\
&R_i^2=D_{xi}^2+D_{yi}^2\\
&w_{xi}=D_{xi}/R_i,\quad w_{yi}=D_{yi}/R_i\\
&S_x=\sum_{i=1}^{n}w_{xi}\quad S_y=\sum_{i=1}^{n}w_{yi}\quad S_k=\sum_{i=1}^{n}1/R_i\\
&w_{Ti}=\left[S_R-w_{xi}\left(S_x-w_{xi}\right)-w_{yi}\left(S_y-w_{yi}\right)\right]/R_i\\
&z_P=\sum_{i=1}^{n}w_{Ti}\cdot z_i/\sum_{i=1}^{n}w_{Ti}
\end{aligned}
\qquad (7—22)
$$

式中 x_P，y_P ——待定点的平面直角坐标；

z_P ——待定点高程；

$(x_i,\ y_i,\ z_i)$ ——搜索圆内数据点 i 的三维坐标，$i=1,\ 2\cdots n$，$4<n<10$；

w_{Ti}——数据点的权。

这种加权平均内插方法，适合于数据点分布不均匀、地形变化大的区域。

四、DTM 的地形分析

尽管DTM 的应用十分广泛，但地形分析是其基本应用，其它应用都可由此推演、扩展。地形分析的内容有地形因子提取、地表类型分类以及剖面图的绘制等。现以栅格结构的DTM 为例讨论地形分析。

（一）地形因子的自动提取

1. 坡度和坡向分析

坡度定义为水平面和地形表面之间夹角的正切值；坡向为坡面法线在水平面上的投影与正北方向的夹角（图 7—33）。

坡度和坡向的计算通常在 3×3 个 DTM 格网窗口中进行。窗口在 DTM 数据矩阵中连续移动后完成整幅图的计算工作。

坡度 $\text{slope}=\text{tg}P=\left[\left(\partial z/\partial x\right)^2+\left(\partial z/\partial y\right)^2\right]^{1/2}$ (7—23)

坡向 $\text{Dir}=\left(-\partial z/\partial y\right)/\left(\partial z/\partial x\right)$ (7—24)

式中的 $\partial z/\partial x$、$\partial z/\partial y$ 一般采用 2 阶差分方法计算。对图 7—34 所示的格网，有

对于 $(i,\ j)$ 点

$$
\frac{\partial z}{\partial x}=\frac{(z_{i,(j+1)}-z_{i,(j-1)})}{2\delta x}\qquad \frac{\partial z}{\partial y}=\frac{(z_{(i+1),j}-z_{(i-1),j})}{2\delta y}
$$

式中 δx、δy 为格网结点 x、y 方向的间隔。

2. 地表粗糙度的计算

地表粗糙度是反映地表的起伏变化和侵蚀程度的指标，一般定义为地表单元的曲面面积与其在水平面上的投影面积之比。但根据这种定义，对光滑而倾角不同的斜面所求出的粗糙度，显然不妥当。实际应用中，以格网顶点空间对角线 L_1 和 L_2 的中点距离 D 来表示地表粗糙度（图 7—35），D 值愈大，说明 4 个顶点的起伏变化也愈大。其计算公式为

$$
R_{i,j}=D=\left|\frac{z_{(i+1),(j+1)}+z_{i,j}}{2}-\frac{z_{i,(j+1)}+z_{(i+1),j}}{2}\right|=\frac{1}{2}\left|z_{(i+1),(j+1)}+z_{i,j}-z_{i,(j+1)}-z_{(i+1),j}\right|
\qquad (7—25)
$$

3. 高程变异分析

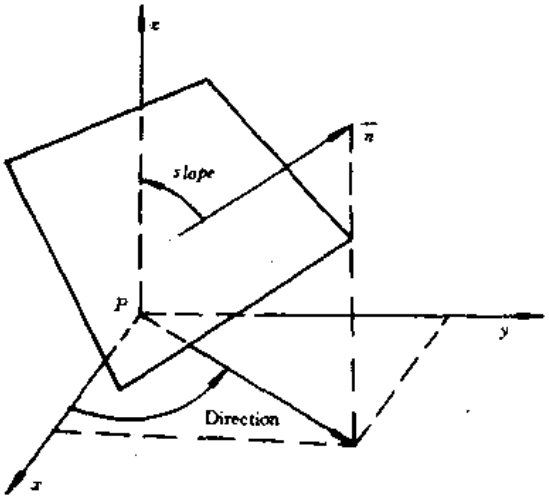


图 7—33　地表单元坡度、坡向示意图

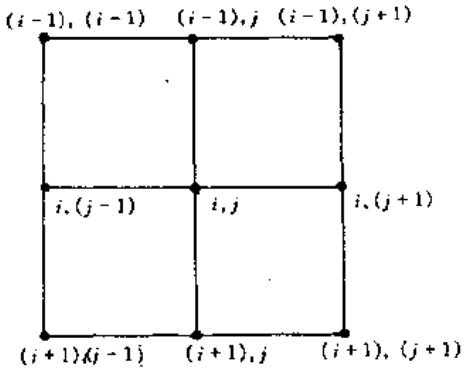


图 7—34　格网结点示意图

高程分析包括平均高程和相对高程的计算。通常以格网的 4 个顶点 P_k（$k=1$，2，3，4）的高程平均值定义为该格网单元的平均高程，即

$$\bar{z}=\frac{1}{4}\sum_{k=1}^{4} z\ (P_k) \qquad (7—26)$$

以格网的平均高程与研究区域某一最低点高程 z_{min} 之差定义为该单元的相对高程，即

$$D_z=\bar{z}-z_{min} \qquad (7—27)$$

高程变异是反映地表单元格网各顶点高程变化的指标，它以格网单元顶点的标准差与平均高程的比值来表示，即

$$V=S/\bar{z} \qquad (7—28)$$

其中，标准差 $S=\left[\frac{1}{4}\sum_{k=1}^{4}(z_{pk}-\bar{z})^2\right]^{1/2}$

(7—29)

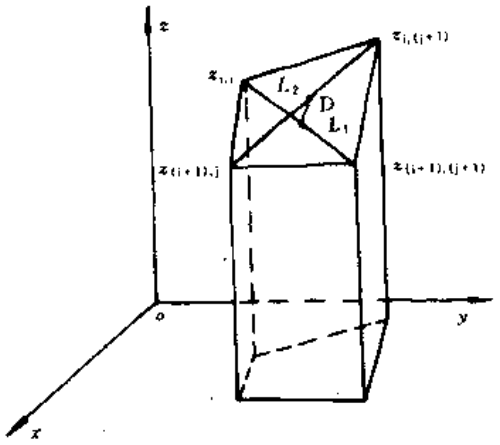


图 7—35　地表粗糙度计算

4. 谷脊特征分析

谷和脊是地表形态结构中的重要部分。谷是地势相对最低点的集合，脊是地势相对最高点的集合。在栅格 DTM 中，可按照下列判别式直接判定谷点和脊点：

(1) 当 $(z_{i,(j-1)}-z_{i,j})\cdot(z_{i,(j+1)}-z_{i,j})>0$ 时

若 $z_{i,(j+1)}>z_{i,j}$，则 $V_R\ (i,\ j)=-1$

若 $z_{i,(j+1)}<z_{i,j}$，则 $V_R\ (i,\ j)=1$

(2) 当 $(z_{(i-1),j}-z_{i,j})\cdot(z_{(i+1),j}-z_{i,j})>0$ 时

若 $z_{(i+1),j}>z_{i,j}$，则 $V_R\ (i,\ j)=-1$

若 $z_{(i+1),j}<z_{i,j}$，则 $V_R\ (i,\ j)=1$

在其它情况下，　$V_R(i, j)=0$

其中 $V_R(i, j)=\begin{cases}-1 \text{ 表示谷点}\\ 1 \text{ 表示脊点}\\ 0 \text{ 表示其它点}\end{cases}$

这种判定只能提供概略的结果。当需对谷脊特征作较精确分析时，应由曲面拟合方程建立地表单元的曲面方程，然后，通过确定曲面上各种插点的极小值和极大值，以及当插值点在两个相互垂直的方向上分别为极大值或极小值时，则可确定出谷点或脊点。

（二）地表地貌形态的自动分类

在DTM数据库的基础上进行地形的自动分类，首先是根据区域的地形地貌特点，拟定分类决策表，然后按照自动提取地形类型信息的流程，进行区域的地形分类（图7—36）。

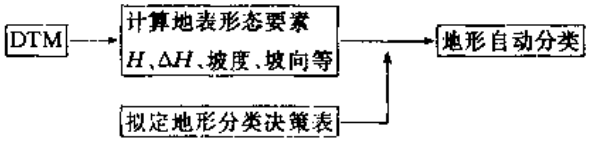


图7—36　地表形态自动提取过程

70年代末，我国曾运用DTM分析技术对云南腾冲地区地表形态进行自动分类试验。分类运用了平均高度、相对高度、单元高度、变异系数、坡度、粗糙度、水系密度等作为地形类型的特征参数，区分为7个地貌类型：山地、安山岩火山锥、安山岩熔岩台地、台地、山间盆地、坝地和谷地。采用统计模式识别的监督分类算法，设计最大似然分类器，对腾冲打直区地貌进行自动分类，经实地对比，自动分类精度达67.5%。

（三）地形剖面图的绘制

DTM表示了一个区域地形的整体状况。在应用中经常需要了解从地面上一点至另一点沿途的地形变化情况，如通视性、地貌轮廓形状、斜坡特征、地表切割强度等，这就需要绘制地形剖面图。

设$z_{i,j}$为格网结点(i, j)上的高度值，如图7—37所示。因此，只要知道待绘剖面线在DTM矩阵中的起点位置(i_1, j_1)和终点位置(i_2, j_2)，就可唯一地确定这条剖面线与DTM格网各个交点的平面位置及其高程。

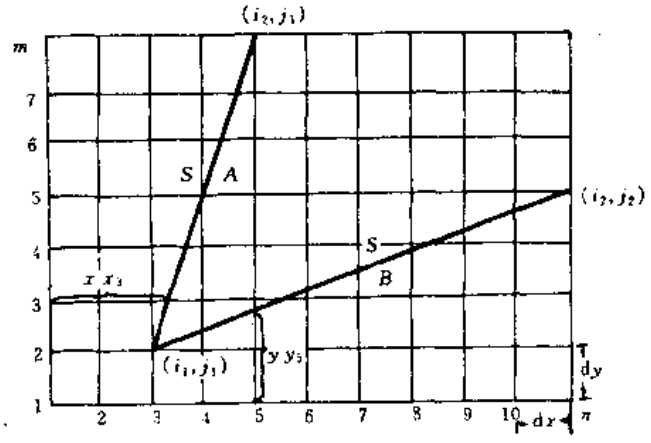


图7—37　剖面线交点的内插

令$\Delta x=j_2-j_1$，$\Delta y=i_2-i_1$，显然，当$\Delta x\neq 0$，且$|\Delta y/\Delta x|-1\geqslant 0$时，应求剖面线（图中$A$线）与DTM格网横轴的交点，它们在栅格DTM坐标系中的平面直角坐标和高程分别为

$$\begin{aligned} yy_k &= i_1+(k-1)\times ISG_2\\ xx_k &= j_1+|(yy_k-i_1)\times\Delta x/\Delta y|\times ISG_1\\ zz_k &= (xx_k-IA)(z_{ik,IB}-z_{ik,Ic})+z_{ik,Ic}\end{aligned} \qquad (7—30)$$

式中：IK=INT（yy_k）；IA=INT（xx_k）；IB=（IA+1）×ISG_1；IC=IB−ISG_1；INT 为取整函数；ISG_1、ISG_2 为“+”、“−”号的取值标志；$k=i_1$，(i_1+1)，$\cdots|i_2-i_1|$。而 ISG_1 和 ISG_2 的值由 Δx 和 Δy 的符号来确定（表 7—12）。

同理，当 $\Delta x\neq 0$，且$|\Delta y/\Delta x|-1<0$ 时，应求剖面线（图 7—37 中 B）与 DTM 格网纵轴的交点，此时：

$$\begin{aligned} xx_k &= j_1+(k-1)\times ISG_1 \\ yy_k &= i_1+|(xx_k-j_1)\cdot \Delta y/\Delta x|\times ISG_2 \\ zz_k &= (yy_k-IA)\cdot(z_{IB,Ik}-z_{IC,Ik})+z_{IC,Ik} \end{aligned} \tag{7—31}$$

其中 IK，IA 含义同式（7—28）；IB=（IA+1）×ISG_2；IC=IB−ISG_2；$k=j_1$，$j_{(1+1)}+\cdots|j_2-j_1|$。

表 7—12　ISG_1 和 ISG_2 值的确定

Δx	>0	>0	<0	<0
Δy	>0	<0	>0	<0
ISG_1	1	1	−1	−1
ISG_2	1	−1	1	−1

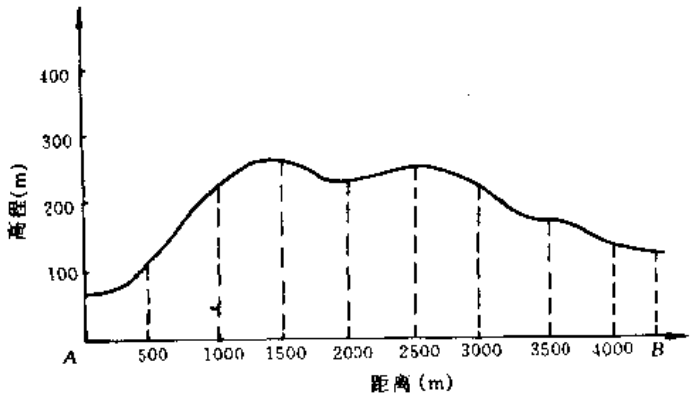


图 7—38　沿某一方向 AB 绘出的剖面曲线

当 $\Delta x=0$ 或 $\Delta y=0$ 时，说明剖面线方向与 DTM 格网的纵、横轴相同，剖面线上各点的高程可直接读出。

在计算出剖面线上各点的高程 zz_k 和剖面线相邻两点的实际距离 S，就可以根据选定的垂直比例尺 m_y 和水平比例尺 m_x，自动绘出所需要的地形剖面图，如图 7—38 所示。

剖面线相邻两点的实际距离 S，可由下式算出：

$$S=\begin{cases} \sqrt{\left(\dfrac{\Delta x}{\Delta y}\times dx\right)^2+dy^2}, & \text{当 } \Delta x\neq 0,\ \left|\dfrac{\Delta y}{\Delta x}\right|-1>0 \text{ 时} \\ \sqrt{\left(\dfrac{\Delta y}{\Delta x}\times dy\right)^2+dx^2}, & \text{当 } \Delta x\neq 0,\ \left|\dfrac{\Delta y}{\Delta x}\right|-1<0 \text{ 时} \\ \sqrt{dx^2+dy^2}, & \text{当 } \left|\dfrac{\Delta y}{\Delta x}\right|=1 \text{ 时} \\ dy, & \text{当 } \Delta x=0 \text{ 时} \\ dx, & \text{当 } \Delta y=0 \text{ 时} \end{cases} \tag{7—32}$$

（四）真实感地形表示

把地形表面以真实感强的三维图形表示出来，是DTM应用中的一个新兴课题。若把地形表示成色调连续变化的三维图像，则在土地景观设计、森林覆盖模拟等应用中更适用、美观和形象。其处理过程是：透视变换；色调计算；隐藏面消除；图形输出等。

1. 透视变换

透视变换是根据选定的观察点位置，适当旋转DTM模型坐标系，并把三维坐标投影到二维平面上，设任意点 P 的三维坐标为（x_P，y_P，z_P），经旋转后，P点在新坐标系中的坐标为

$$\left.\begin{aligned} x_P' &= l_1 x_P + l_2 y_P + l_3 z_P \\ y_P' &= m_1 x_P + m_2 y_P + m_3 z_P \\ z_P' &= n_1 x_P + n_2 y_P + n_3 z_P \end{aligned}\right\} \tag{7—33}$$

式中 l_1、l_2、l_3 —— x' 轴的方向余弦；

m_1、m_2、m_3 —— y' 轴的方向余弦；

n_1、n_2、n_3 —— z' 轴的方向余弦。

规定新坐标系的 x' 轴与过观察点垂直于投影平面的直线重合，从垂足（原坐标系与新坐标系的共同原点）指向观察点的方向为正，并规定 y' 轴为水平轴。设观察点的坐标为 L（x_L，y_L，z_L），可按下式确定 x' 轴的方向余弦：

$$\left.\begin{aligned} l_1 &= \frac{x_L}{r} \\ l_2 &= \frac{y_L}{r} \\ l_3 &= \frac{z_L}{r} \end{aligned}\right\} \tag{7—34}$$

式中 $r=\sqrt{(x_L^2+y_L^2+z_L^2)}$。

y' 轴与 z' 轴的方向余弦按下式求得：

$$\left.\begin{aligned} m_1 &= -l_1/S \\ m_2 &= l_1/S \\ m_3 &= 0 \\ n_1 &= -l_1 l_3/S \\ n_2 &= -l_2 l_3/S \\ n_3 &= S \end{aligned}\right\} \tag{7—35}$$

式中 $S=\sqrt{l_1^2+l_2^2+l_3^2}$。

由新坐标系的设置可知，y' 轴对应于显示屏的 x 轴，z' 轴对应于显示屏的 y 轴。因此，DTM中任意点 P 在显示屏上的二维坐标为

$$\left.\begin{aligned} \overline{x}_P &= m_1 x'_P + m_2 y'_P \\ \overline{y}_P &= n_1 x'_P + n_2 y'_P + n_3 z'_P \end{aligned}\right\} \tag{7—36}$$

2. 色调计算

色调计算的基本原理是，DTM曲面单元的明亮程度（色调）与光源和观察者的位置有关，也与各曲面单元的法线方向有关。

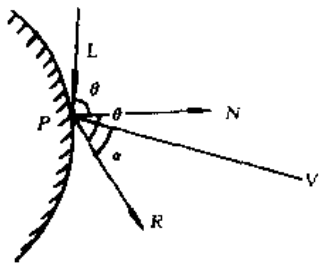


图 7—39　地形表面的基本矢量

在图 7—39 中，L 为光线，N 为曲面单元法线，V 为视线，R 为镜面反射和漫反射的合成反射光线，则明亮度的算法表达式为

$$\left.\begin{aligned} R&=R_1+w\ (\cos\alpha)^n\ (1-R_1)\\ G&=G_1+w\ (\cos\alpha)^n\ (1-G_1)\\ B&=B_1+w\ (\cos\alpha)^n\ (1-B_1)\\ R_1&=[d+\ (1-d)\ \cos\theta]\ \cdot X_0\\ G_1&=[d+\ (1-d)\ \cos\theta]\ \cdot Y_0\\ B_1&=[d+\ (1-d)\ \cos\theta]\ \cdot Z_0 \end{aligned}\right\}\qquad (7—37)$$

式中　R、G、B ——分别为观察者得到的屏幕上地形的三色（红、绿、蓝）刺激值；

R_1、G_1、B_1 ——分别为计算过渡值；

X_0、Y_0、Z_0 ——分别为数字地面模型的设计彩色三刺激值；

w ——反射强度，取 $0\leqslant w\leqslant 1$；

d ——背景光的比率，一般取 $d\leqslant 1$；

θ ——光源和法线间的夹角；

α ——视线和反射光间的夹角；

n ——地形表面的光泽度，一般取 $n=15$。

3. 隐藏面消除

地形立体显示时，靠近观察者的剖面有可能遮挡较远的剖面，应把受遮挡的局部剖面抹去，不予显示。彻底消隐技术难度很大，近似的解决方法是：仅对 $\bar{x}$ 坐标相同的前后剖面对应点的 $\bar{y}$ 坐标进行比较，如果某剖面上的一点 $\bar{y}$ 值大于前面所有剖面上 $\bar{x}$ 坐标相同点的 $\bar{y}$ 值，则认为该点是可见的，否则为隐藏点，应予消除。

4. 显示输出

经以上处理过程后，在显示屏上就可得到 DTM 的真实感显示三维图形，如图 7—40 所示。

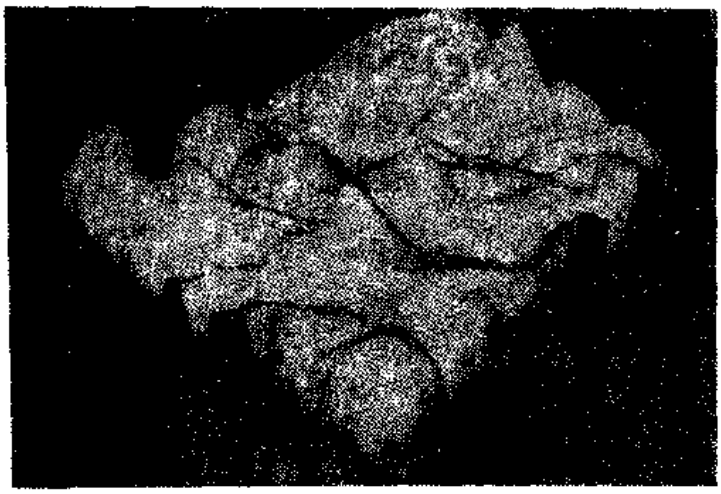


图 7—40　地形的三维真实感图形

第 7—3 节　数据输出与自动制图

数据输出是按用户能理解的形式或能够传送到其它计算机，并能以立即使用的形式表示数据处理结果的一种操作过程。人们能够理解的共同形式是地图图形、图像、表格、文字等；计算机兼容的形式，是指能读入其它计算机系统的磁盘、磁带记录形式，还包括某些通讯网、电话线、无线电设施以及多媒体电子传输形式。

大多数 GIS 的数据输出模块都有多功能数据输出的内容，比如：

(1) 选择输出范围的缩放/窗口功能；

(2) 比例尺放大、缩小、图形旋转功能；

(3) 图形自动综合功能；

(4) 颜色、线型变换及多边形充填功能；

(5) 文字、符号注记功能；

(6) 三维显示、动画显示功能；

(7) 用层式叠置层选择图形数据的功能；

(8) 输出设备驱动功能；

(9) 数据输出到磁介质上的转存功能；

(10) 文本文件输出的打印功能。

随着计算机多媒体技术的发展，GIS 数据输出模块还应具有多媒体播放功能。

一、数据输出设备

GIS 中数据输出设备可分为二大类：一类是电子屏幕上暂时显示分析结果或输出内容的显示装置，如图形显示器、电子投影屏幕等。另一类是在纸张、聚脂薄膜或其它材料上产生永久性图形图像的装置，如绘图机、打印机、硬拷贝机等。它们按输出产品的格式又可分为：栅格输出和矢量输出两类，并且通常与所输出空间信息的数据结构（矢量与栅格）相对应，但栅格输出设备可输出矢量图形，矢量输出设备也可输出栅格图形。GIS 中输出设备与系统的关系可用图 7—41 来描述。

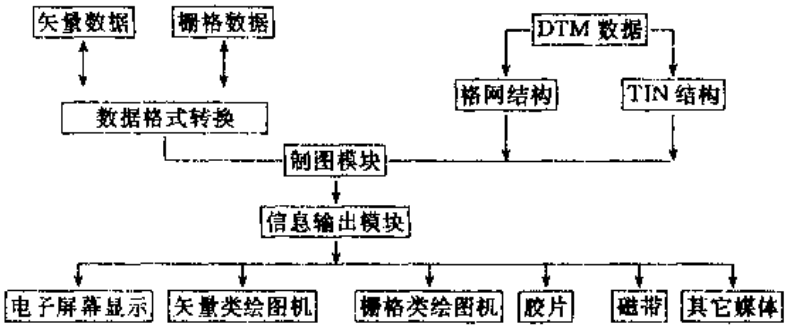


图 7—41　GIS 中数据输出设备与数据的联结

1. 图形显示器

图形显示器是 GIS 制图过程中不可缺少的外围设备，在数据获取、数据处理和数据输出中都起着十分重要的作用。图形显示器包括随机扫描显示器，存贮管显示器和光栅扫描显示器三种类型。前二种显示器只能和绘图仪一样用画线条的方式显示图形的轮廓线，图形真实感不强。而光栅扫描显示器不仅能显示一般的图形，而且能显示具有多级灰度和多

种色彩的图像，这就使得输出具有真实立体感的图形成为可能。

光栅扫描图形显示器采用和家用电视机相同的扫描方式，即把显示屏分成有限个离散点。如 1024×768，每个离散点为一个像元，屏幕上的像元阵列组成了光栅。每个像元可具有某种灰度或颜色。所显示的线段由相邻像元串接而成，字符由一定大小的像元矩阵构成。当电子束按顺序扫描整个屏幕时，只有电子束经过图形所在位置上的像元才具有灰度或颜色，从而在屏幕上可观察到所需的图形。为了保证屏幕上图形的稳定性和不使图形闪烁，电子束每秒钟对整个屏幕刷新 30 次以上。

电子束在扫描过程中，每个像元上的信息（亮度、灰度、色调）都要存放在帧缓冲存贮器中。帧缓冲存贮器是一个很大的阵列，以“位”来描述像元的信息。目前，大多数帧缓冲器都至少有 8 位位平面（即每个像元用 8 位二进制数表示），能产生 256 种灰度级或颜色。图 7—42 描述了光栅扫描图形显示器的工作流程。

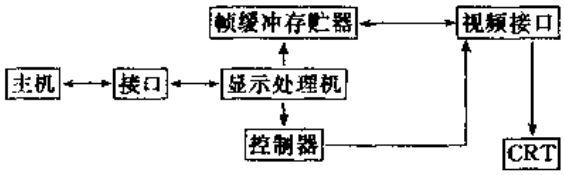


图 7—42　光栅图形显示器工作流程

2. 打印机

最简单的栅格输出设备是行式打印机。行式打印机制图就是根据计算机处理的数据，利用程序设置打印机上的符号，编排好打印的位置或重叠打印一些数字、符号和它们的组合，以制作分级统计图或类型图（图 7—43）。



图 7—43　用行式打印机打印的分级统计图

这种打印机输出的地图虽然不精致，但可显示制图现象的地理分布，以及在利用专门绘图机制图之前，提供概略的样图。若打印机采用 24 针打印头和多色色带，还能以肉眼难以分辨的针点输出精度要求不高的矢量图形。

通常，打印机的最重要作用是输出属性文本和制表。

3. 高分辨率阵列打印机

使用静电复印原理的阵列打印机可在每英寸上打印出 400 个点，即以 15.7 点/mm 的高分辨率输出图形或图像。这种打印机由两个基本部分组成：纸张输送机和安置细针的横架，每个细针彼此靠得很近又能单独操作。

这种打印机的优点是速度快，一幅 40cm×60cm 的地形图在笔式矢量绘图机上输出需要 30 分钟以上，而在阵列打印机上仅需 30 秒，而且与地图的复杂程度无关。

4. 扫描绘图机和喷墨绘图机

这两种绘图机都属于栅格类图形、图像输出设备。扫描绘图机用于制作胶片。它把光源聚焦于承片滚筒的面上，并逐点曝光，曝光量由像元的值来控制。扫描滚筒高速旋转，每转一周步进一定距离，如 25μm。曝光量分为 256 级，对应于像元的 256 个灰度级。这种扫描绘图机输出的栅格数据的彩色或黑白图像具有很高的分辨率，输出结果可与高质量相机媲美。

喷墨绘图机的构造与扫描绘图机相似，但它不是用光束在感光胶片上曝光，而是同时

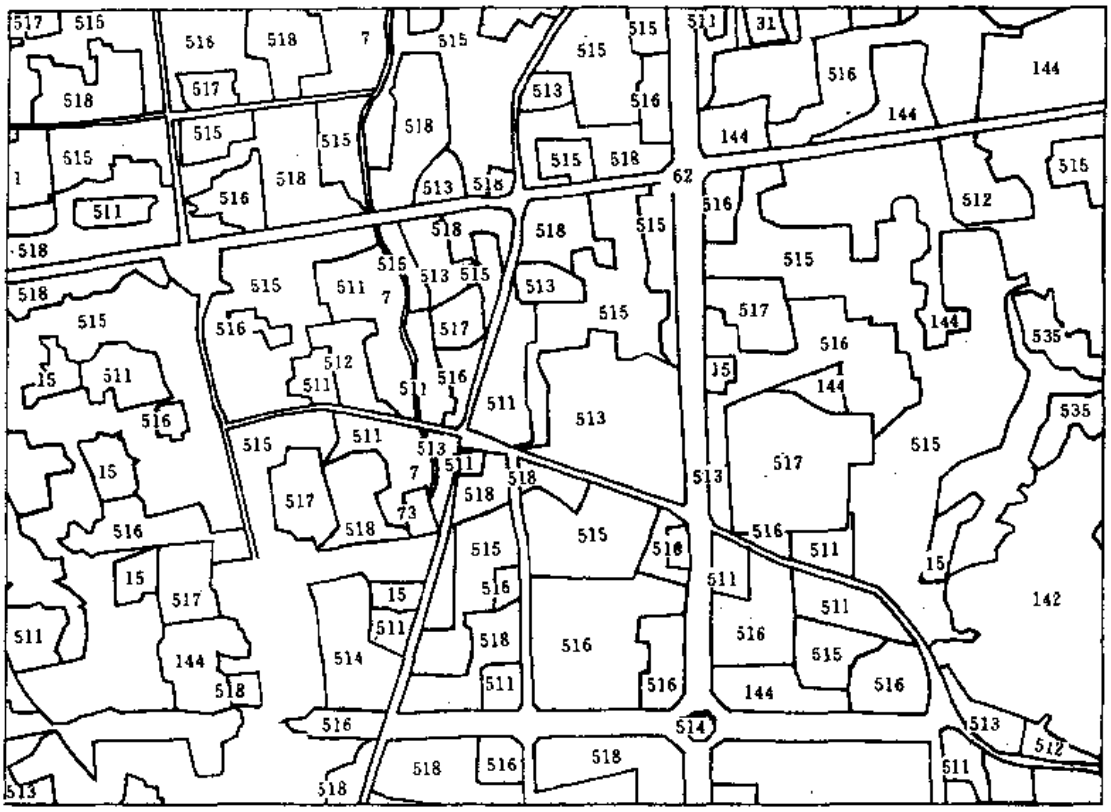


图 7—44　绘图机输出的土地利用图

用品红、黄、青三种色液从三个喷嘴喷出雾液，并按不同比例混合后在纸上得到所希望的图形或色调，其分解力为每毫米5点的多色绘图，且价格便宜。

5. 笔式绘图机

这是一种矢量类图形绘图机，有平台式和滚筒式两种。

平台式绘图机由放置图纸的平面，x，y方向独立移动的笔架等组成，而滚筒式绘图机用在y方向转动的滚筒代替了平面。

在笔式绘图机中，所有的信息都由一系列“绘出线段”这类命令绘出，即“把笔移到x，y点”、“落笔”、“换笔”、“抬笔”、“从x_1，y_1点绘到x_2，y_2点”等命令。笔的移动值可以是绝对坐标，也可以是坐标增量。

笔式绘图机的灵活性和绘图速度在很大程度上取决于绘图软件，即绘制字母和符号等复杂图形的程序功能。一般来说，绘图机内部存贮器中包括了预置的基本命令、印刷体字符和一些通用符号。制图程序只要调用这些命令就能完成复杂的图形绘制。

这类绘图机一般都具有较好的定位精度、重复精度和动态精度，基本能满足各种图形绘制的精度要求。笔式绘图机输出的图形见图7—44。

除以上的输出设备以外，还有许多种硬拷贝机可将栅格图形或图像复制成硬拷贝，如激光打印机、摄影硬拷贝机等，也是一种快速、经济的图形输出设备，但输出图幅较小，因而较少使用。

二、图形设备驱动软件

迄今，大多数图形输出软件都以输出设备为转移，造成了系统中每增加一个新的输出设备就不得不对绘图软件进行相应的修改。而重写或修改这些软件，以满足不同的输出设备，则是一种耗时、不经济的工作。解决的途径有两种：一是GIS输出模块中针对各种流行的输出设备预编写设备驱动程序，如ARC/INFO系统提供了数十种设备的驱动程序，程序中设置了设备参数，包括设备名称、型号、通讯参数、输出幅面大小、色彩种类、使用笔数等，执行这些程序就可把相应的设备与输出模块相联结，输出相应的产品；另一种途径是按国际通用标准生产各种设备，使所有的图形、图像输出软件不因设备的不同而改变，也即，这种标准与设备本身无关。GKS（图形核心系统）绘图软件则是这种标准软件。

GKS是德国标准化协会所开发的图形核心系统，并被全美国家标准协会所采纳，同时也作为国际标准而使用。

GKS为程序员提供了一整套二维图形基本单元子程序库，这些单元可通过FORTRAN-77等高级计算机语言编写的相容接口进行访问。

在GKS中，矢量数据结构的基本描述单元是线单元、标记单元和文本单元等。线单元用来绘制一系列连接在一起的直线段（链）；标记单元用来将符号放置在指定的点位上而不绘线；文本单元则用来把文字说明和注记绘在任意需要的位置和方向上。在栅格数据结构中，GKS支持充填单元和像元阵列单元，用于描绘特殊符号——面积填充符号和独立符号。GKS还能将线符号和不同颜色表示的属性附加在像元阵列单元上，同时还可以把基本单元当作简单实体加以定义和处理。

GKS的最重要特点是所有从程序模块输出到显示设备的程序都通过名为虚拟设备接口（VDI）程序来传送。这样VDI就成为用户程序与设备之间的通道，用户只需处理自己所需要的图形、图像，而不必去关心设备的种类和型号。GKS和VDI作为输出设备无关的

标准接口，在GIS中的地位如图7—45所示。

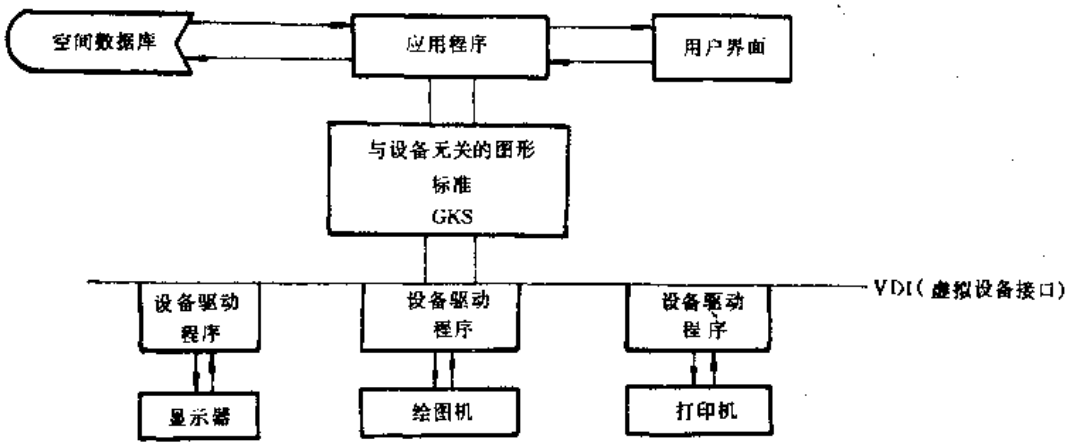


图7—45 GKS、VDI在GIS数据输出模块中的地位

在GIS中加入了GKS和VDI后，价格昂贵的图形处理软件可直接输出到许多不同的设备而不必修改输出程序。对程序员来说，GKS的效益十分明显。

三、制图符号与汉字注记

为了保证GIS输出产品的可视性，各种地图符号和汉字注记已成为GIS信息输出模块中必不可少的部分。

（一）制图符号库

制图符号库中存贮的主要是地图符号的颜色码和图形信息，每个符号组成一个信息块。地图符号库的设计遵循下列的原则：

（1）尽可能符合国家规定的各种专题地图图式；

（2）遵循图案化、精确性、逻辑性、对比性、统一性、色彩象征性、制图和印刷可能性等一般原则；

（3）符号库的可扩充性。

符号信息块的构成通常有下列两种方式：

1）直接信息法　信息块中存贮符号图形的矢量数据（图形特征点坐标）或栅格数据（足够分解力的点阵数据），直接表示符号图形的每个细节。这种信息块占用存贮空间大，但绘图程序的算法统一。

2）间接信息法　信息块中只存贮符号图形的几何参数（如图形的长、宽、间隔、半径、夹角等），其余数据都由计算机按相应绘图程序的算法解算出来。这种信息块占用存贮空间小，但需编制专门程序。

在GIS输出模块中，采用直接信息法较多。现以直接信息法为例，介绍矢量符号信息块的设计方法。矢量符号库用于矢量方式的图形输出，可按点、线、面三类符号分别建库。

1. 点状信息符号库

在点状符号信息块中记录图形的颜色码、图形特征点坐标及其联系（一般用抬落笔码来表示），信息块采用以符号定位点为原点的局部坐标系。图7—46是某种独立地物的放大表示。可按图中点号顺序在信息块中记录$\{P_i, x_i, y_i\}$ $(i=1, 2, \cdots 11)$，P_i为i点的抬落

笔码，（x_i，y_i）为该点在局部坐标系中的坐标。

点状信息块的结构描述见表 7—13。

绘图时，读出信息块，按图上指定位置和方向，并指定旋转和平移参数，调用基本绘图语句直接连线绘出。不难看出，各类点状符号均可用同一绘图程序输出。

2. 线状信息符号库

地图上各类线状符号往往是由某一图案（线状符号的基本单元），沿线状要素的中轴线串接而成，图案坐标系的 x 轴与线状要素中轴线重合，并在线状要素拐弯处随着弯曲而变形。绘图时由程序完成图案的定位、弯曲处理和串接等操作。图案坐标系的原点在图案的首端。如对图 7—47 所示的线状符号的图案，采用的描述结构如表 7—14。

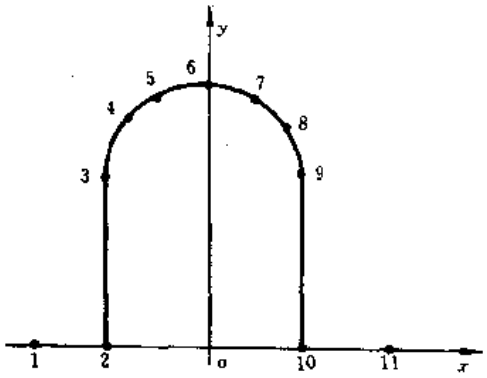


图 7—46　点状符号的特征点表示

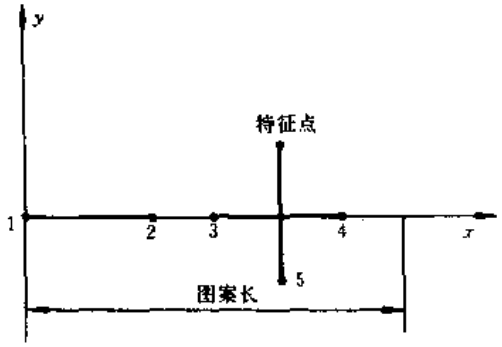


图 7—47　线状符号的特征点表示

表 7—13　点状符号信息描述

符号代码	颜色码	特征点数 n	抬落笔码 1	x_1	y_1	…	抬落笔码 i	…	x_n，	y_n

表 7—14　线状符号信息描述

符号代码	颜色码	特征点数 n	图案长	抬落笔码 1	x_1	y_1	…	抬落笔码 n	x_n，	y_n

另外，利用简单的线状符号可以构成复杂的符号，如双线铁路、陡崖线等。

3. 面状信息符号库

面状符号由填充符号在面域内按一定方式配置组合而成。多数情况下，填充符号在面域内按一定方向、一定间隔（行距）逐行配置。晕线是面状符号形式之一。其它各种面状符号也可按计算晕线端点的方法事先计算出各行与轮廓边的交点（图 7—48），然后在每对交点间配置相应的填充符号，这实质上相当于在两端点间的连线上绘制线状符号。

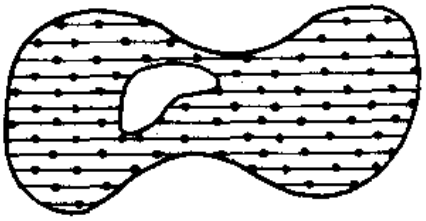


图 7—48　面状符号的配置

面状符号信息块中存储的是充填符号的图案信息。它的结构和线状符号信息块类似，但需增加三种信息：行距、行倾角、排列方式。其中行距为符号排列的行间距；行倾角是指行方向与 x 轴的夹角；排列方式一般有“#”型、“卄”形和散列式三种方式（图 7—49）。它们在信息块中以不同代码来表示。面状符号的信息描述如表 7—15。

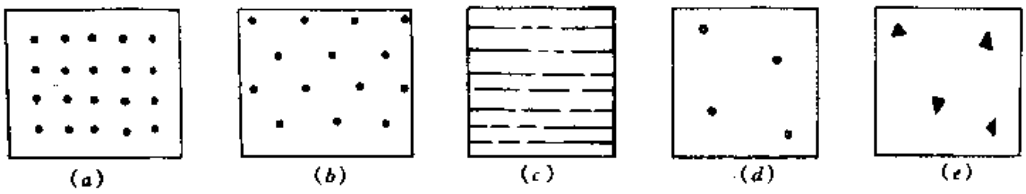


图 7—49　面状充填符号的排列方式

a—“#”形排列；*b*—“卄”形排列；*c*—攀线散列式；*d*—符号散列式；*e*—符号散列式

表 7—15　面状充填符号的信息描述

符号代码	特征点数 n	图案长	排列方式	行倾角	行距	抬落笔码 1	x_1	y_1	…	抬落笔码 n	x_n,	y_n

4. 各种符号库与 GIS 数据库的联结

为了在 GIS 产品中输出各种符号，应在空间要素（点、线、面）的属性表中加入相应的符号代码，这样，在输出时，对应的点、线用点线符号替代，而面状图斑以充填符号充填，其过程见图 7—50。

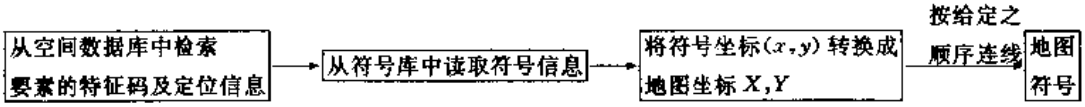
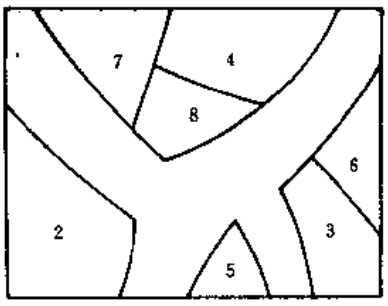


图 7—50　GIS 中符号库的调用过程

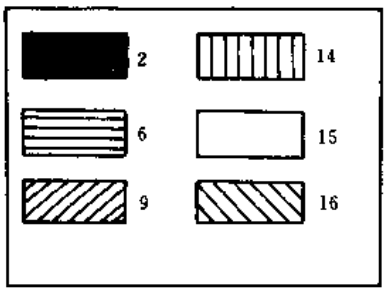
例如，在城市土地分类制图中，空间数据以完整的拓扑结构来描述，在其属性表中加入充填图案栏。制图输出时，调用面状符号库，绘出充填的土地分类图（图 7—51）。



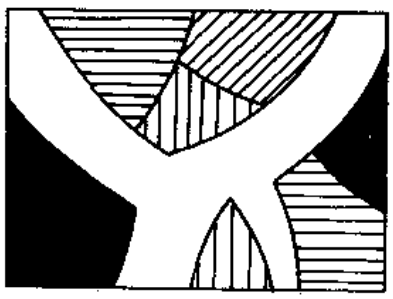
(a)

多边形号	土地使用类型	土地级别	充填符号代码
2	501	3	2
3	502	3	6
4	523	1	9
5	531	2	14
6	501	3	2
7	502	2	6
8	531	2	14

(b)



(c)



(d)

图 7—51　GIS 中面状要素图符库的充填制图

a—土地利用图斑；*b*—属性表；*c*—符号库；*d*—输出图形

（二）汉字注记

在地图上，各种要素除用线划、符号表示外，还需用文字和数字来说明各要素的名称、种类、性质和数量，即地图注记。在我国，为了使GIS输出的地图能被大众所接受，汉字注记是必不可少的。而从国外引进的许多GIS工具软件本身并未配置汉化工作环境，因此，必须为这些地理信息系统工具编写汉字注记软件，设计和建立专用汉字库。

汉字库分为栅格字库和矢量字库两类。通常，栅格汉字（或称点阵汉字）适用于屏幕显示和打印机输出，矢量字库适合于矢量绘图机输出。

从原理上看，汉字库与点状符号库没有本质区别。国家标准GB2312-80《信息交换用汉字编码字符集》中对7000多个汉字和特殊字符规定了统一的机内码。因此，不同的汉字系统其输入码和汉字信息块的存贮组织与地址码分配可能有多种方案，但每个字的机内码是统一的、规范的，为汉字的检索提供了方便。

矢量汉字库一般已商品化，可直接购买使用。但对于一般用户来说，读取现有各种汉字系统的点阵字库用以注记则更为便利。现以目前国内应用最广的CCDOS汉字系统字库CCLIB为例，探讨在GIS输出产品上注记汉字的方法。

1. 汉字点阵的提取

CCLIB字库是以不压缩字模存贮的16×16点阵字库，在字库中，每一个汉字或特殊符号用16个整数32个字节存放。CCLIB字库的组织严格遵循GB2312-80的规定。因此，每个汉字可通过其区位码去访问，由区位码计算汉字的机内码与汉字的地址。

点阵字库是以二进制码的形式存贮的，因此，要提取汉字的点阵信息，首先须将二进制代码转化成可见的ASCII码。在CCLIB中，其解决方法可采用把字库定义为记录长度为32字节的随机文件，则一个字模就是随机文件的一个记录。CCLIB汉字字模存贮第一个汉字的位置是第658个记录。如此类推，在区位码表中查找需注记的汉字，输入区位码，就可计算出其文件记录号。提取该记录，再利用位操作命令将其转换成ASCII码，即点阵的

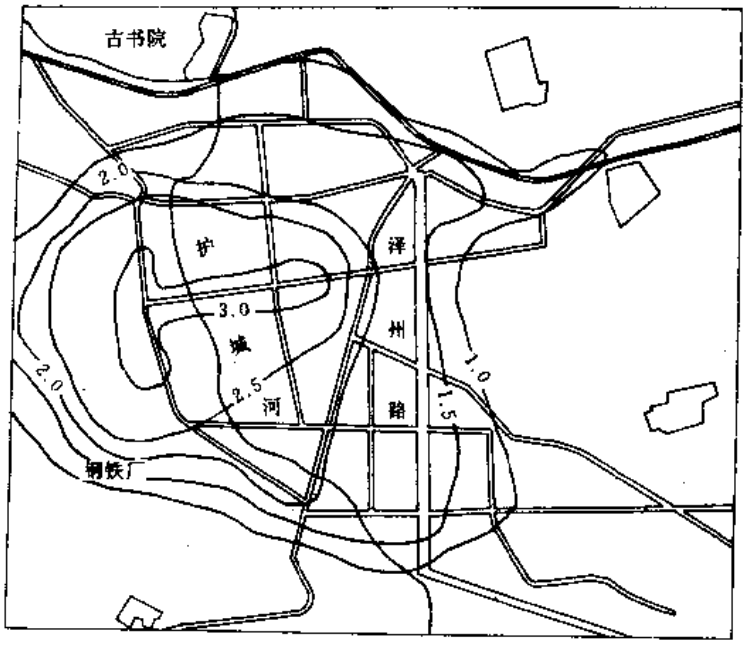


图7—52 晋城市污染等值线图

"0"、"1" 表示。

2. 点阵汉字矢量化

汉字矢量的方法与栅格数据矢量化采用同样的算法。矢量化后就可得到矢量汉字。为了使注记的汉字美观，还可进行一些特殊处理，如选用高分辨率（如 24×24 点阵、48×48 点阵）字库以及多体（黑体、宋体、仿宋、楷体等）字库，放大处理，字体旋转，字体边界平滑处理等。这样，输出的地图产品可与出版地图媲美。

图 7—52 显示了一幅在 PC-ARC/INFO 上开发的汉字库注记的环境地图。

四、制图数据的自动处理

一个完善的地理信息系统，必定是一个良好的制图系统，因为制图输出是 GIS 的一个基本功能。GIS 除了完成窗口剪折、比例尺缩放、图形旋转、文字符号注记等基本制图功能外，还应能提取某些制图要素，以及完成制图自动综合等功能。

（一）图形显示

图形显示是最常用的一种输出形式，它是把数据处理和分析结果，按需要以字符、数字或图形的方式显示出来。图形显示技术的最大优点是用户可以利用图形显示设备中的输入装置，如键盘、鼠标器等，对计算机进行实时控制，即通过把图形显示器的荧光屏当作图纸，对显示的内容和图纸进行修改、增删、旋转、移位、放大或缩小等实时处理。这样，既有效地提高了成图速度，又能灵活地进行输出产品的设计。现以微机配置的 VGA 图形适配卡（分辨率为 640×480，16 种色彩）为例，讨论图形显示的方法。

设显示的图形范围为（x_{min}，y_{min}，x_{max}，y_{max}），这样的图形只能在有效的屏幕范围内显示。计算机的屏幕坐标系是以左上角为坐标原点的设备坐标系（图 7—53），而图形坐标是笛卡尔坐标系坐标，因此，必须进行视见变换。

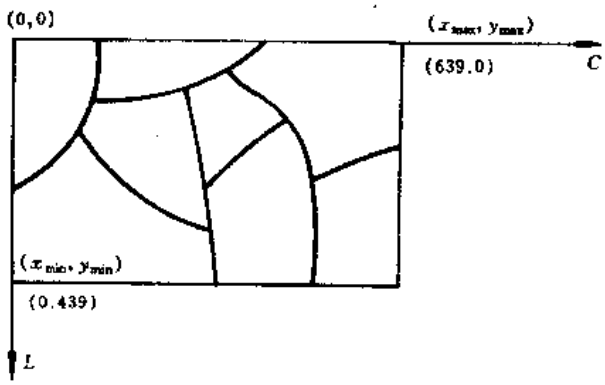


图 7—53 屏幕坐标与屏幕图形

屏幕坐标是以像素所在的行列位置表示的，则图形上任一点（x，y）在屏幕上的坐标为

$$\left.\begin{aligned} C&=\mathrm{INT}\left(\frac{x-x_{min}}{S}\right)\\ R&=\mathrm{INT}\left[479-\left(\frac{y-y_{min}}{S}\right)\right]\\ S&=\frac{\sqrt{(x_{max}-x_{min})^2+(y_{max}-y_{min})^2}}{\sqrt{640^2+480^2}}\end{aligned}\right\} \quad (7—38)$$

式中：C 为屏幕列坐标；R 为屏幕行坐标；INT 为取整数据；（x，y）为任一图形要素在笛卡尔坐标系中的坐标；S 为视见变换比例尺。

图形经变换后，图形要素对应于屏幕的像素。此时，在屏幕上用写点和绘线命令就可显示出图形。

（二）制图要素的提取

空间数据库中存储了大量的信息，在制作输出某种专题图时，仅需输出满足一定条件

的要素。这种提取的条件可以是定义的窗口，或是满足一定属性条件的图形，然后在空间数据库中查找满足这些条件的要素作相应的处理（如剪切），并输出到图形显示器或其它设备。

（三）制图综合

在地图制图学中，制图综合始终是地图制图的一大难题。离开了制图综合就无法制作地图，无法实现从制图资料到地图制作再到地图使用的信息传输过程。随着 GIS 规模的扩大、种类的增加和处理的问题日益复杂化，人们越来越认识到制图综合在 GIS 中的重要性。

所谓制图综合是指依据大比例尺地图数据库建立小比例尺地图以及在对多专题进行要素提取、空间分析、评价预测的基础上派生的各种专题图时，如何在地图上反映出本质的、主要的和整体规律性的内容，舍去非本质的、次要的和局部的细节。

近年来，国内外的学者和研究机构积极研究和推动制图自动化的实现，在 GIS 中建立了制图综合模块，推出了一些商品化的软件。如 Zeiss 开发的 PHOCUS 系统中的制图综合模块 CHANGE（1992），主要用于线状和面状要素的制图综合。美国 INTERGRAPH 公司研制的与其 MGE 系统相匹配的制图综合模块 MAP GENERALIZER，能够处理点、线、面各类数据，并具有选取/删除、化简、典型化/提炼、合并、退化、分级/符号化、夸大、移位、整饰等各种分项操作功能。其宗旨是以制图综合的科学原则和方法来指导和操纵 GIS 中的数据提取，专题要素分解与合成，小比例尺地图的派生等过程。集成化的制图综合模块可用图 7—54 来表示。

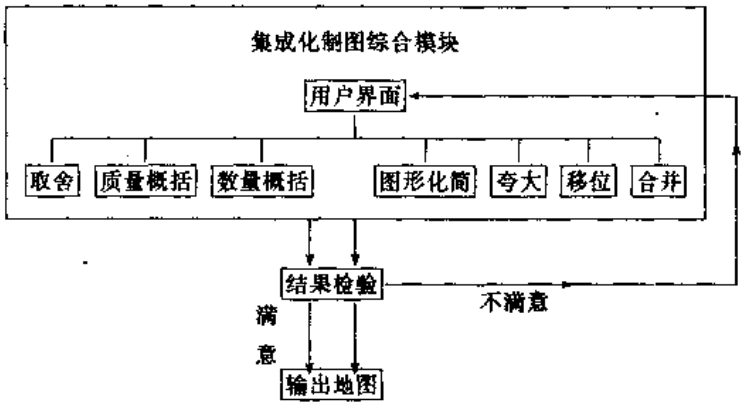


图 7—54　集成化制图综合模块的构成

集成化制图综合中各单项功能的实现方法如下：

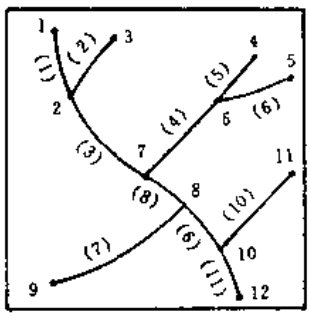
1. 要素取舍

在它的最简单形式中，要素的取舍依赖于其属性代码。该代码是按一定级别排列的具有优先权的编码。应用于消除许多次要类型要素，避免在缩小比例尺的地图图面上出现要素的堆积。如图 7—55 所示，对水系网中的主干河流、次干河流、支流等分级编码（1～4，4 为主干流），干流具有被保留的优先权，当图 a 被缩小比例尺时，编码为 1 的支流都被舍去。

对于面状要素，一般以其面积值的大小，采用临界值的方法，舍去小面积图斑。

2. 数量和质量概括

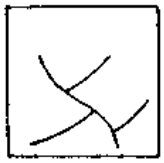
数量和质量特征的概括过程较简单，不需采用复杂的数学模型，仅需采取对数据库进行检索、判断和处理的方法即可完成。



(a)

弧号	从点	至点	长度	等级	优先权
1	1	2	3.4	1	1
2	3	2	3.8	1	1
3	2	7	5.4	2	2
4	7	6	5.0	2	2
5	6	4	2.8	1	1
6	6	5	3.1	1	1
7	9	8	5.8	2	2
8	7	8	2.1	3	3
9	8	10	2.7	4	4
10	10	11	5.6	2	2
11	10	12	2.6	4	4

(b)



(c)

图 7—55　水系网在比例尺缩小时的要素综合

a—原图；*b*—属性表；*c*—综合结果

制图对象数量特征的概括是指计算机按照需求扩大制图对象的数量分级间隔，减少分级数目，以达到减少其数量差别的目的。这种概括可以理解为自动修改同类要素的级别（或特征码），使其合并。例如土地利用图斑中的“梯田旱地”代码为 142，“坡旱地”代码为 143，均以 142 代替原 142、143 两级代码，则图斑数量随着同属性图斑的合并而减少。

质量特征的概括是指计算机按照要求用概括的分类代替详细的分类，以综合的质量概念代替各个物体的具体质量概念，达到减少制图对象之间质量差别的目的。根据不同的情况，质量概括有三种方法：

1）制图对象归类法　即按制图对象的分类体系，由低级类向高级类归并，直到符合新的比例尺能反映的分类级别为止。如“园地”分类的三级类“苹果园”、“葡萄园”等在制度综合时归并到二级类“果园”中去。

2）质量概念转换法　即把一种物体的质量概念转为另一种物体的质量概念。如在综合居民地时，把大面积建筑群中的小块空地转换为建筑区（图 7—56）。

3）图形等级转换法　即随着比例尺的缩小，把轮廓图形转换为符号图形，如矿井井筒以井筒符号表示。

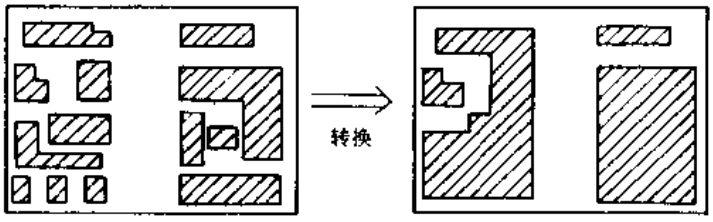


图 7—56　图形质量概念转换

3、形状化简

形状化简分线状地物化简和面状地物化简两种情况。

线状地物形状化简的基本思路是，通过某种算法，减少线状要素上点的数量，既达到简化图形的目的，同时又保持线划图形的基本特征。这种算法之一是 Douglus-Peucker 算法（详见第 6—3 节中的空间数据压缩部分）。

面状地物的图形化简分两个步骤来完成，首先进行轮廓线的化简，这可通过与线状地

物形状化简相同的方法来完成。然后进行内部结构化简，方法是把原图进行矢量→栅格变换，变成栅格形式，然后按褶积公式用一个单元权阵进行低通滤波，分离出应保留的图形信息，必要时，还可在此基础上再加上高通滤波，以便从剩余的信息量中分离出具有局部高频性质的地物。

4. 夸大和移位

移位是指图形要素整体向某一侧移动。将要素移出其几何精确位置，可以防止因比例尺缩小引起的拥挤。而且在许多应用中，综合的更合理的方法是，首先考虑一个稍有移动的重新定位，然后才决定舍去一个地图要素。

夸大是线划上某点向外移位或整体向四周移位，因此，夸大的算法也可归结到移位中去。在实施移位时，有三个问题必须明确：何处，哪个地物需要移位？向什么方向移位？移位量多大？第一个问题的解决方法是在空间数据库中检索并识别加宽后互相压盖的位置，然后按保持原位优先级的高低来确定移位的地物；第二个问题的解决方法是：两个地物沿它们质心连线相反的方向移动；第三个问题的解决方法是遵循人眼视觉可识别的地图上的最小间距，移位距离一般取图上 0.2～0.4mm。

位移算法的基本思路是：在确定带有夸大或移位特征的点后，将此点与其前后两点组成三角形，算出三角形顶角（需夸大或移位的角）平分线的长度，然后在此长度靠顶角一端加一个定值，得到一个新顶点。再计算出新点的坐标，并用它取代旧顶点，这就使图形的局部点得到夸大。若想使点向内移动，只要将定值取负值即可。不同点的连续夸大处理就可构成移位，即整条线向同一方向移位，此时其形状不变，但中心位置发生移动。图 7—57 表示了两条平行的铁路与公路在比例尺缩小时的移位表示。

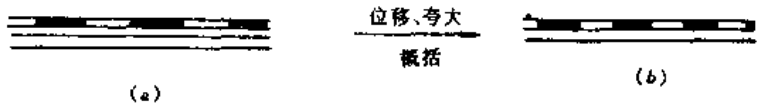


图 7—57　道路的移位、夸大表示

a—原图（1∶2.5 万）；*b*—处理结果（1∶10 万）

5. 合并

合并实质上是面状要素的形状化简。其方法是：利用数据库检索功能找出相距较近的图斑，判断哪几个合并较为合适，计算出可合并的数量。当能满足合并条件时，计算合并后的轮廓点位实现合并，其关键是确定合理的合并距离阈值。

五、GIS 数据制图与输出的动画和多媒体技术

1. 动画和多媒体技术简介

计算机动画（animation）技术，实际上是指用计算机来辅助制作动画片中每一序列，这些序列中每一幅画面又是完全静止的，并要求这些序列在快速显示时，能使人产生该片中诸角色逼真地不断运动的感觉。

动画片中的各幅画面，可以是对预先描述好的对象由计算机通过一些特定的算法自动生成的，也可以是事先安排好各画面，通过某种输入装置送入计算机，或者是两种方法结合生成的。计算机动画技术中，已生成好的动画序列向观众展示的方式，可以是在计算机

控制下实时地在显示设置上展示；也可以先用某种介质（如录像带或电影胶片）记录下来，然后再用相应设备放映出来。在此假定采用前一种展示方式。

计算机多媒体（multimedia）技术就是计算机综合处理多种媒体信息：文本、图形、图像、动画、声音和视频信号等，使多种信息建立逻辑连结，集成为一个系统，并具有交互性。多媒体计算机技术汇集了计算机体系结构、计算机系统软件、视频音频信号的获取、处理、特技以及显示输出等技术。应用多媒体技术是90年代计算机的时代特征。

目前，市场上流行的多媒体技术主要有五种，现以Intel和IBM公司推出的DVI系统为例，介绍多媒体计算机系统的结构。

DVI技术第二代产品ActionMedia750 I于1991年由美国Intel和IBM公司联合推出，其结构如图7—58所示。其核心是三块专用的DVI接口板，DVI视频板，DVI音频板以及DVI多功能板。

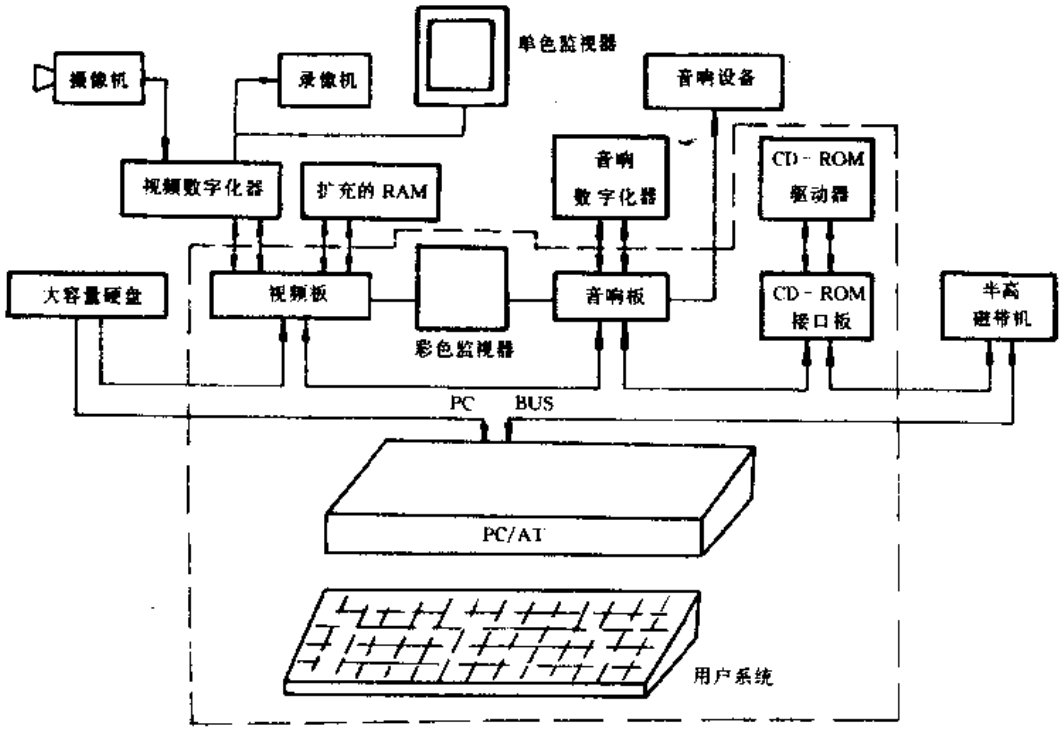


图7—58 DVI系统结构图

以IBM PC 386，486或其兼容型计算机作为工作平台，同时配有CD-ROM驱动器，带有放大器和音响效果的RGB彩色监视器，组成了DVI用户系统。在此基础上再配置与多媒体有关的外设，视频信息数字化器（连接到DVI视频板上），音响信号数字化器（连接到DVI音响板上），扩展的RAM（连接到视频板上），大容量的光盘或硬盘、磁带机、录像机、音响设备、监视器以及摄像机或扫描仪等，则组成DVI开发系统。

2. 动画技术和多媒体技术在GIS数据输出中的应用

应用GIS技术可以分析各种图形、影像的动态变化，便于进行数据的动态时空分析和显示。GIS制图输出方法有两种，一种是编制多时相的系列地图，另一种是使用特殊的描述、分类和索引方法在相同的地图上表示动态性，即应用计算机动画技术。例如：使地貌图或

DTM 按垂直投影方式完整地在三维空间表示，并使其能按规定的方向旋转，背面的特征从正面看似乎被隐藏，而旋转到正面却又出现，这样能得到真正的立体地形或地貌。动画技术能把地图中的水系以蓝色表示，并按规定的速度运动，从而产生碧波荡漾的效果。在网络分析中，使用动画技术，可以模拟资源在网络中的运移，如汽车奔驰在城市道路网中。这种动画技术使空间数据在电子屏幕上以逼真的效果显示，以向用户传递更多的信息。

多媒体技术包含了动画技术，它不仅能向人们传递视觉信息，还把声音也结合起来，向人们传递听觉信息。因此，多媒体技术应用于 GIS 空间数据制图，显著地扩大了单幅图面内的信息量，并提高传输效率。

一个典型的多媒体 GIS 制图输出模块的构成如图 7—59 所示。

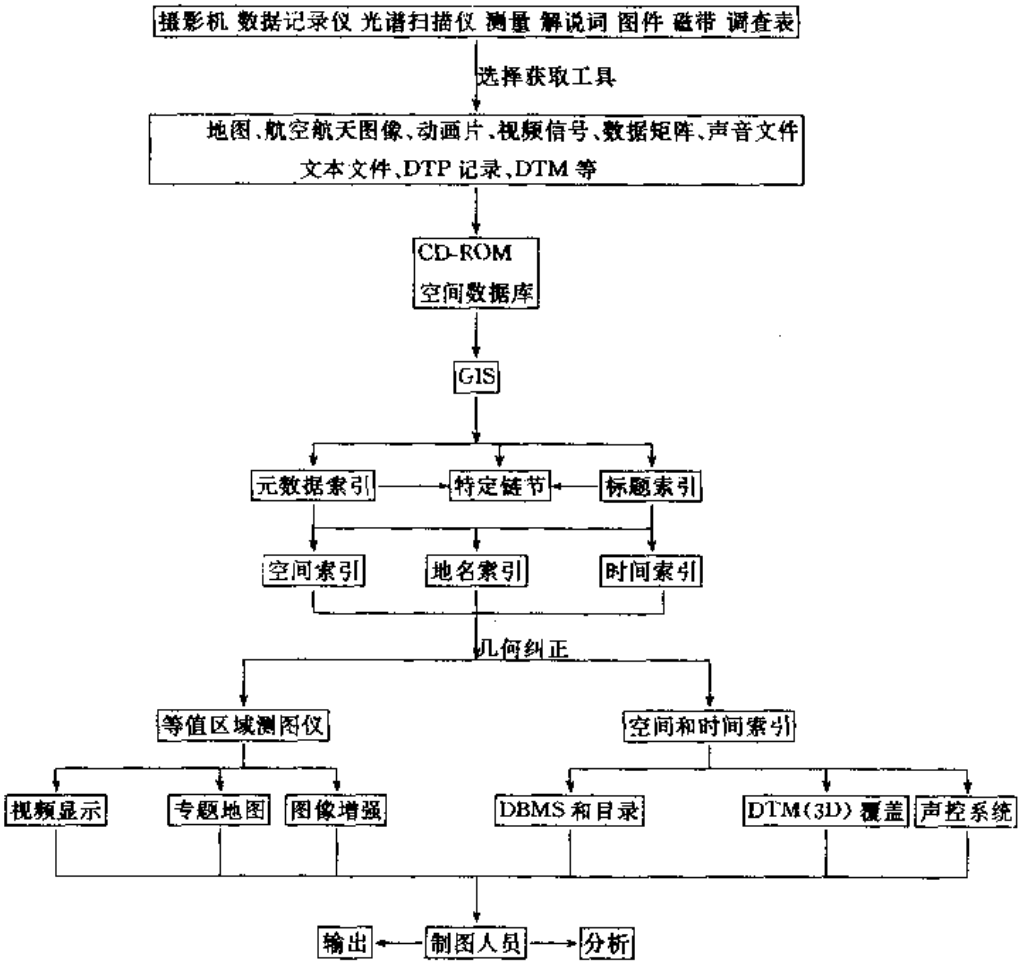


图 7—59　多媒体 GIS 输出模块的系统构成

现以城市地图制作为例，介绍多媒体技术在 GIS 数据输出中的应用。城市地图是近几年来发展较快的专题地图，随着城市经济的发展，城市地图的内容、主题、类型都得到了较大发展，出现了城市商业经济、城市环境、城市地价等一系列新品种的专题地图。目前城市地图主要印刷在纸质上，装订成册或以散页形式发行。随着科技的发展和居民素质的提高，城市电子地图的产生和发展是必然的，必将对城市生活和经济流通起重要的作用。

城市车站、码头是城市流通最重要的场所，通常大多数旅行者（尤其是外地人）到达

这里时，首先要买一份城市交通图，以便确定怎样到达目的地。现在，只需把城市地图与多媒体技术相互结合起来，开发城市导览系统，将城市概况，街道分布，交通走向、旅馆、娱乐场所，商业区等有关内容借助于图形、文字、声音、动画等技术对其组织和管理，然后在车站、码头或中心地带开设导览咨询服务，借助于大尺寸屏幕把城市有关内容显示出来，旅行者通过导览系统咨询，不仅可以知道怎样到达目的地，而且还知道目的地位置。例如，某旅客想找一旅馆，系统能立即列出城市不同等级的宾馆、饭店、招待所的名称，一旦选定某个旅馆，系列立即显示它在城市网络图中的位置、最佳行进路线、电话号码等，同时还伴有该旅馆情况介绍的内部设施图像和语言说明。显然，这种集图、分析、声形、动画于一体的导览系统比任何纸质交通图都优越、方便和准确。

同样，如果将多媒体技术应用于城市信息管理决策系统中，政府官员就可以通过屏幕，像看电视一样了解城市经济、环境、交通、贸易、生产等各方面的情况，使城市管理真正做到现代化、系统化。

由于在GIS数据输出模块中应用动画和多媒体技术的研究还刚刚起步，有许多实际问题尚待解决，但这种应用的真正实现将为期不远。

第7—4节　数据共享和标准化

信息时代的一个特征是实现信息或数据在社会中的通讯或共享。地理信息数据共享（也称作地理信息联合使用）是一种用数据字典完成的，用于不同地理信息系统之间数据交换的正规方法，其目的是有效地使用地理数据资源和合理地进行数据服务，使各种信息系统中不断增加的大量数据不仅被数据的拥有者所使用，而且被更多的用户在不同的系统中所使用。

数据共享的最关键一步是数据的标准化和规范化。数据标准化和规范化，不但是一个系统与另一个系统实现数据共享的需要，而且是在一个系统内保持数据的连贯性、持续有效性的需要。

一、数据共享

1. 数据接口

对于数据共享来说，数据资源本身就是基础。维护数据在各个系统是独立的，但是从用户的观点来看，他们对数据如何维护和组织并不感兴趣，仅对哪些数据可以得到，数据以什么形式表示等感兴趣，而标准的信息接口是用户最感兴趣的对象（图7—60）。

在数据共享中，交换分量必须具有共同的外观表示。有效的信息服务应该便于支持多种通讯载体和方法。

技术是数据提供者组织信息服务的手段，在某种程度上，它可以是专用的，但它必须能够支持标准的接口。

当用户应用的数据是从多个数据源获得时，质量信息就显得更加重要了，尤其是用来把坐标、时间、标识符等数据组合起来的信息质量更加重要。

2. 数据字典系统

数据字典不仅是空间数据管理的工具，也是数据共享中的重要工具（图7—61）。它不仅能告诉用户系统中存储了什么信息，而且还表示了数据管理的策略。

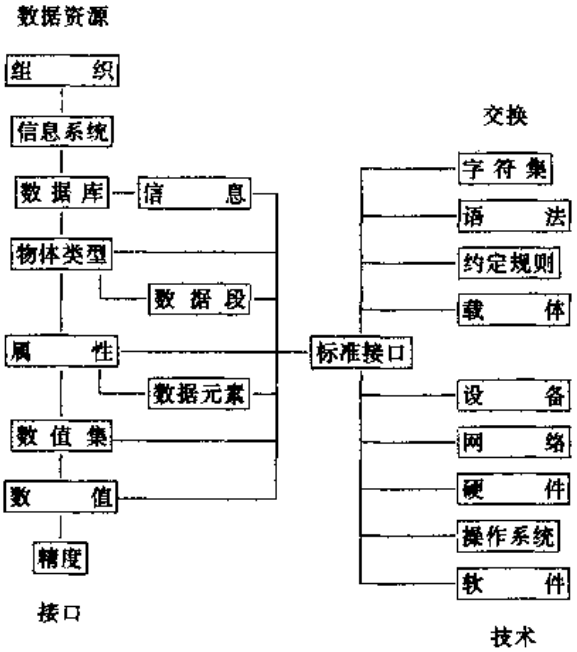


图 7—60　数据接口示意图

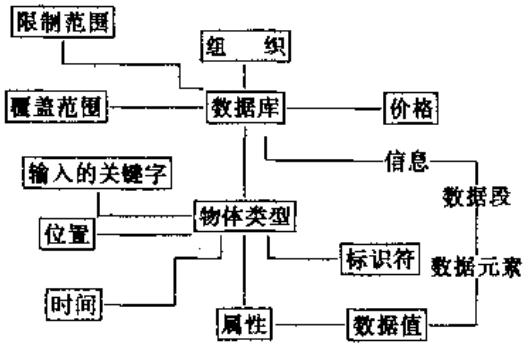


图 7—61　数据字典示意图

地理数据字典系统（GDDS）中记录了数据资源的信息描述：该系统为查询信息提供描述，为获取数据提供工具（图 7—62）。用户可以把该系统当作获取地理信息的目录。

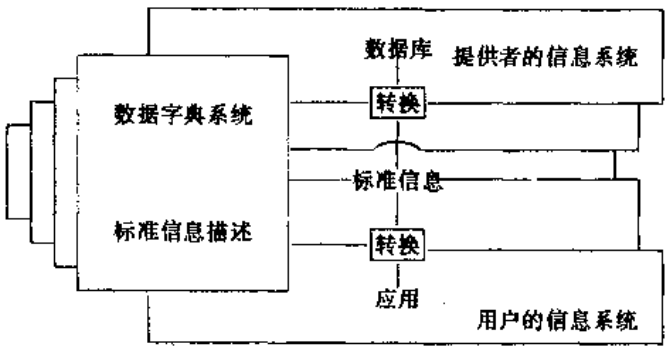


图 7—62　数据共享中的基本概念

3. 影响数据共享的因素

为了实现信息系统的数据交换、共享，要解决体制上和技术上的问题。就技术因素而

论，不同系统对空间数据采用的数据结构和数据格式不同，阻碍了数据的共享。对此，可通过数据的标准化、规范化、标准的数据转换接口等途径来解决。

另一个影响因素是人为的，如数据保密和有关奖励政策的问题。

因此，数据共享应该是建立在通用性电子数据交换的概念上，否则，数据保密会限制某些数据资源的使用，造成各部门对相同数据的重复输入而导致浪费。如果主要数据必需受到严格保密，则数据的拥有者应当安排一些代用品，即可传播的数据产品。

奖励政策是支持标准的数据交换良好的手段。从长远的角度来看，标准的服务会收到好的效益。因此，政府的专门机构应对执行标准化数据交换给予奖励，鼓励专题信息系统数据的共享。

二、数据标准化和规范化

地理数据标准化和规范化的基本内容有：统一的数据采集原则、统一的空间定位框架、统一的数据分类标准、统一的数据编码系统、统一的数据记录格式和统一的数据测试标准。

1. 统一的数据采集原则

GIS 数据库中涉及到众多自然资源和经济统计部门的数据。它具有数据量大、数据种类繁多，空间定位数据和统计数据并存，数据随时更新且有共享性、利于信息传输、交换等特点。根据系统的目标和功能，要求系统、全面而准确地占有尽可能多的有用数据，从应用急需出发采集数据。

在数据采集时应遵循以下原则：

（1）只采集、存储基本的原始数据，不储存派生数据，根据应用的频度、实现最小的冗余。如在采集、建立 DTM 时，只储存地形高程值，不储存派生出的信息，如坡度、坡向等地形因子。

（2）各级统计部门提供的数据为最基本数据，当其它部门提供的数据与它有矛盾时，以统计部门的为准。

（3）统计部门未进行统计的指标，以各地最直接的专业部门提供的数据为准。

（4）当上级部门和下级部门提供的数据有出入时，以上级部门核准的数据为准。

（5）对未进行统计的数据，需经多个专家研究，取得与实际情况相符的数据。

（6）地理底图应以最新出版的实测地图为准，专题图最好采用比例尺一致的区域系列图作为图形库的基本数据。

总之，要采集的是具有权威性、科学性和现势性的数据。

2. 统一的空间定位框架

统一的空间定位框架是各种地理空间数据的输入、输出和匹配处理的共同地理坐标基础，这种坐标基础可以归化成地理坐标、网络坐标和投影坐标这三种坐标系统。当数据信息的来源不同时，必须将它们转换到这三种坐标系统之一中来。

我国 GIS 国家规范研究组建议，我国 GIS 所采用的投影应与国家基本图系列所采用的投影相一致，即 1∶1 万～1∶50 万比例尺图幅采用高斯-克吕格投影为其输入、输出的基础；1∶100 万及更小比例尺的图幅采用等角圆锥投影。高斯-克吕格投影是一种保角投影，在每一幅图范围内无角度变形，其最大长度变形不超过 0.14%，面积变形不超过 0.28%，精度可以满足使用的要求。

对于格网系统的选择，研究组提出了基本格网、加密格网、合并格网和辅助格网四种

划分方法。其中基本格网分为3级；加密格网是在基本格网的基础上再细分成6级；合并格网也是以基本格网为基础，按需要进行格网整数倍的表示，分成5级；辅助格网是通过平面直角坐标建立的格网，覆盖于1∶1万～1∶5万比例尺地图上。

除辅助格网外，其余三种类型的格网均以经纬度的细分作为格网划分的依据，与我国基本比例尺地形图分幅系列一致，其目的是为了便于分幅检索和区域拼接。

3. 统一的数据分类标准

将成千上万个标准化数据组织成系统的数据库，必须要有一个较好的数据分类系统。分类系统划分是否合理，直接影响到系统的组织、系统间数据的联接、传输和共享，以及系统的质量。因此，它是系统设计和数据库建立的一项极为重要的基础工作。

国家规范研究组建议，信息分类体系采用宏观的全国分类系统与详细的专业系统之间相递归的分类方案（表7—16）。即低一级的分类系统必须能归并和综合到高一级的分类系统上。为此，首先按照社会环境、自然环境、资源和能源三大类，作为第一级；其次，根据环境因素和资源类别的主要特征和基本差异，再划分为14个2级类；第三，按每一个2级类包括的最重要的内容，作为第3级类别；最后；按照各个区域的地理特点和用户的需求，拟订区域的分类系统和每一个专业类型的具体分类标准。

4. 统一的数据编码系统

GIS存贮的空间数据具有时间、空间和属性的复杂特征，需要通过计算机能够识别的代码体系来提供地理分类和特征描述，同时需要制定统一的编码标准，以实现地理要素的计算机输入、存贮，以及系统间的数据共享。目前，我国正在进行地理信息及其属性的编码规范和标准的研究，尚未提供地理信息的统一编码方案。但是，在制定系统编码时，应遵循下列原则：

1）系统性　指信息系统中空间信息的编码应统一规划、统筹安排，不能各行其是，随意变动；

2）一致性　指任何专业名词、术语的定义必须严格保持概念的一致，对同一专业名词、术语必须是唯一的，且不能重复；

3）科学性　指编码能可靠地识别数据信息的分类，以较少的代码提供丰富的参考信息，以及根据代码结构能进行数据间关系的逻辑推理和判别；

4）标准化　指代码的内容和长度必须一致，码位的分配及格式必须一致，以及明确的代码含义，不能出现代码的多义性等；

5）扩展性　指系统编码的码位应留有充足的余地，当代码增加或删除时，不至破坏原有代码；

6）适用性　指代码不宜过长（一般为4～7位），必须以较少的码位提供丰富的参考信息，便于记录和查找，既可以减少出错可能性和操作量，又可减少存贮量及计算机处理时间。

例如，对于地形图中各种空间要素，可采用四位整数编码法。这种编码的方法主要是参考地形图图式符号的分类，用四位整数对地形要素进行分类编码。把全部编码分成1～9共9个类。最高位1代表测量控制点，2代表居民地；3代表独立地物；4代表管线和垣栅；5代表境界；6代表道路，7代表水系；8代表地貌与土质；9代表植被。其余三位代表线型、方式和种类。如居民地中的4位编码含义为

2 × × ×

代表线型　代表方式　代表种类

其中：第 2 位数（线型）：1—实线；2—虚线；3—点线；4—加固坎；5—陡坎；6—坡。

第 3 位数（方式）：0—闭合；1—不闭合；3—平行。

第 4 位数（种类）：0—直接连接；>0 表示加绘不同符号。

5. 统一的数据记录格式

数据记录格式是指地理信息系统的原始数据和输出数据在磁性介质内的记录方式，对不同来源（遥感、地图、统计数据）和不同形式（点、线、面）的数据，都必须按照标准的记录格式记录，以保证系统对各种数据信息的接纳、处理和共享。

GIS 通常采用的数据记录格式有：多边形数据格式、栅格数据格式和影像数据形式。多边形数据记录格式采用弧段为基本的逻辑单元，其数据文件由结点文件、弧段拓扑文件、弧段坐标文件和多边形文件组成。为了加速多边形信息的检索，可以建立类型或区域位置检索表。栅格数据记录格式一般采用同质数据串作为基本的逻辑单元。而影像数据一般采用计算机兼容磁带来记录。首先是磁带头文件，它包含了一些影像的描述信息，其次是图像文件，图像的记录有两种方式，即交叉方式和顺序方式。图像处理系统一般都能识别这两种记录方式。

6. 统一的测试标准

这是针对地理要素属性数据提出的。例如：大量的环境、土壤等数据，由于测试的标准不统一，造成其量纲的不同，使数据间没有可比性，则数据完全不能共享。因此，统一测试标准和量纲是属性数据与其它系统交换的必要条件之一。

各种属性数据中的物理量的测试应按国家计量标准的测试方法和量纲来进行和表示。

7. 国际上 GIS 数据标准化简介

在国际上，先后出现了多种数据交换标准，例如：

(1) 美国地质调查局的 DLG (Digitel Line Graph) 数据交换格式；Antodesk 有限公司的 DXF 数据交换标准；Intergraph 股份公司的 ISIF。加拿大的 MDIF；英国的 NTF (National Transfer Format)。

(2) 美国国家地图精度标准。

(3) 特征分类、分析及标准特征编码方案，如加拿大的地理信息委员会 (Canada Council on Geomatics，CCDG) 制定的标准。

(4) 专题分类、分析与编码方案，如新西兰的土地利用分类标准 (NILUC，1984)。

(5) 数据质量与可靠性报告标准，如美国开发的 DCDS Task Force，1988。

(6) 美国于 1992 年 7 月 29 日批准“空间数据转换标准”(The Spatial Data Transfer Standard，SDTS)，或称“联邦信息处理标准 (FIPS) 出版物 173”，以解决不同数据类型、结构方式的地理和制图空间数据之间的转换问题。

(7) 多国数字地理信息工作组 (DGIWG) 设计了“数字地理信息交换标准”(DIGEST)，以建立数字地理数据交换的统一方法。

表 7—16　国家规范研究组建议的数据分类和项目总表

项　目	内　容	基本数据来源
Ⅰ社会环境		
1. 城市与人口	①城镇人口、分县人口总数；	人口普查办公室
	②自然村密度（大小、数目、按第Ⅱ级格网）；	1∶1万～1∶10万地形图
	③人口分布（按第Ⅱ级格网）	人口普查办公室
2. 交通网	①铁路（双轨、单轨、车站、专用线、长度、运输能力，与省界、公路等的交叉点）；	铁道部
	②公路（省级、县级、公社简易公路、桥梁载重限制、与省界、铁路和主要河流的交叉点）；	交通部
	③航运（港口、泊位、船舶吨位、通航路线、水深、季节变化）；	交通部
	④航空（航线、航班、航空港、运输能力）	民航总局
3. 行政区划	①国界、省、市、县级界线与面积（多边形）；	外交部、民政部、国家测绘局等
	②省、市、县级管辖区（按Ⅴ级格网点）；	
	③城市规划区（按Ⅴ级格网点）；	建设部等
	④自然保护区管辖范围；	林业部等
	⑤工矿区（油田、禁区、饲养场、旅游点，名胜文物保护区）	建设局、林业部等
4. 地名	①城市名称及其中心坐标；	地名委员会
	②各县名称及县城中心坐标；	地名委员会
	③主要河流、湖泊、山峰、港湾名称及坐标；	地名委员会
	④自然地理单元及其区域坐标（山脉、流域、盆地、高原）	地名委员会
5. 文化和通信设施	①学校、医院等；	文化部、教育部、卫生部等
	②科学试验站网点（气象、水文、地震…台站）；	
	③邮电通信网点	邮电部
Ⅰ自然环境		
6. 地形	①海拔高程（按Ⅴ级格网点）；	国家测绘局
	②山峰高程、水库、湖面高程；	
	③湖泊、水库水深、大陆架以及海深；	国家测绘局
	④地形图与遥感资料检索	国家测绘局
7. 海岸及海域	①分县海岸线长度、线段坐标；	海洋局
	②分县岛屿岸线、面积、长度、坐标；	海洋局
	③基本海况：滩涂面积、潮汐台风、常年风向、底质、温、盐、海浪等	海洋局
8. 水系及流域	①流域划分界线及面积（100平方公里以上与省界交点、控制站点、水库坝址及坐标、分段节点）；	水电部
	②流域辖区（按第Ⅱ级格网）；	水电部
	③水系交汇点（坐标、面积）及干、支流等级、长度（交叉点坐标）	水电部

续表

项　目	内　容	基本数据来源
9. 基础地质	①地表岩类或沉积层及其时代；	地矿部
	②断层性质（特别是活动性质）；	地矿部、地震局
	③地球物理观测点（重力、地磁、地震等）；	地矿部、石油部地震局、中国科学院
	④人工地震（浅层、中层和深部、包括海上）；	石油、地矿、地震局、煤炭、中国科学院等
	⑤地球化学观测点及其特性；	
	⑥环境地质（地盘沉降、土壤承压力、滑坡、泥石流、崩塌等）；	地矿部、中科院
	⑦地震裂度区划	地矿部、中科院
		地震局
Ⅱ资源与能源		
10. 土地资源	①地貌类型（包括海岸和浅海）；	中国科学院
	②土壤类型（包括土壤肥力等）；	农牧渔业部、中科院
	③土地利用类型；	国家测绘局、农牧渔业部、林业部等
	④灾害（风沙、盐碱、台风、雪害、水土流失、旱涝、霜冻、寒潮）	气象局、水电部、农牧渔业部、中国科学院等
11. 气候和水热资源	①辐射量、日照量和云量（按第Ⅱ级格网）；	国家气象局
	②热量资源（年最高温、最低温、年均温、月均温、积温等）；	国家气象局
	③降水（年最高、年最低、年、月平均、积雪等）；	国家气象局
	④风能；	国家气象局
	⑤陆地水文（最高、最低流量、年、月平均流量、含沙量、洪峰、污染等）；	水电部
	⑥冰川、雪坡、冻土；	中国科学院
	⑦湖泊、水库、港湾；	交通部、水电部、中国科学院
	⑧地下水	水电部
12. 生物资源	①主要农作物，分年的耕作面积、亩产、灌溉面积等；	农牧渔业部
	②森林类型、面积、树种、蓄积量、采伐、更新面积；	林业部
	③草场类型、面积、产草量、载畜量；	农牧渔业部
	④淡水养殖与渔业（种类、面积、产量等）；	农牧渔业部
	⑤病虫害、减产频率和程度；	农牧渔业部
	⑥野生植物、野生动物资源	林业部
13. 矿产资源	①煤炭、泥炭（类型、储量、矿区矿点、生产能力）；	地矿部、煤炭部、石油部、冶金部、有色金属总公司等
	②石油、天然气、油页岩（类型、储量、油田、生产能力）；	地矿部、煤炭部、石油部、冶金部、有色金属总公司等
	③黑色金属（分类、储量、矿山、生产能力）；	地矿部、煤炭部、石油部、冶金部、有色金属总公司等
	④有色金属（分类、储量、矿山、生产能力）；	地矿部、煤炭部、石油部、冶金部、有色金属总公司等
	⑤稀土元素（分类、储量、矿山、生产能力）；	地矿部、煤炭部、石油部、冶金部、有色金属总公司等
	⑥非金属（分类、储量、矿山、生产能力）	地矿部、煤炭部、石油部、冶金部、有色金属总公司等
14. 海洋资源	①海洋能源；	
	②海洋养殖与水产；	国家海洋局等
	③海底矿产资源；	
	④海涂资源	

在制定地理数据的标准或规范时，都要涉及四类相互关联的问题，或者说要同时协调处理这四类问题。它们是：硬件与通讯、软件、数据格式与结构，以及数据集。其具体内容和关系如图 7—63 所示。

为了加速地理信息系统在矿区的应用，应解决“矿区资源与环境信息系统”中数据分类、编码和格式体系问题。此项研究应在全国统一规范的指导下进行，既要与国家规范接轨，又要与兄弟部门协调，立足现在、面向未来。

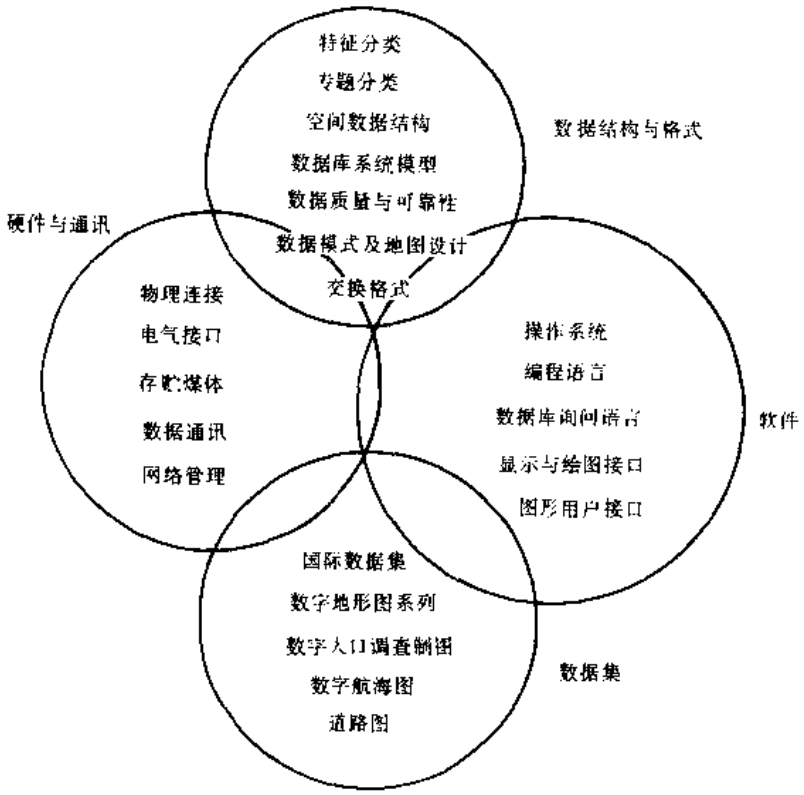


图 7—63　地理信息标准化、规范化中的有关内容

第八章　地理信息系统的应用

第8—1节　概　　述

人类正在进入以信息技术和空间技术应用为特征的新型信息时代。作为地理信息产业重要组成部分的地理信息系统，从80年代末期以来已成为最活跃的信息产业之一。它涉及到GIS软件、硬件、数据、遥感与航空摄影、制图、野外数据采集、数据交换、系统集成和咨询等内容。其中，软件是GIS的技术核心，而从事软件开发和系统集成的公司或机构则往往成为GIS产业的主体。

产业形成和发展的重要标志是GIS应用的人力、资金投入的增加和应用广度、深度的扩展。据统计1991年世界范围内GIS的总收入为70亿美元。其中，应用服务为46亿美元，软件为2亿美元，硬件为22亿美元。而据国际信息产业权威咨询机构Dataquest公司的调查，GIS的年产值为30亿元，并且近年来以16%～40%的年增长速度在发展。有人预计，按此速度发展到2000年，GIS软件系统将达200～500万个用户。

早期GIS主要应用于自动制图，设施管理（AM/FM）和土地信息系统（LIS）。后来逐步扩展到资源与环境管理、森林清查、城市规划、市政管理、灾害监测与预测、科学研究和军事战略等众多领域。随着GIS技术的成熟，数据积累和应用环境的改善，它的应用范围不断扩展，应用程度不断深化。现在，GIS已进入工程设计与管理、经济规划、政治分析决策、农业耕作、商业布局、金融决策、交通运输、卫生防疫、体育竞赛等几乎所有涉及空间信息的行业。GIS市场已由过去的技术驱动发展到应用驱动。从应用水平看，也已经跨上了一个新的层次。从过去简单的提供数据到现在的提供分析模型和解决问题办法。即从数据入到数据出的地理数据库系统到具有空间分析功能的地理信息系统。

在我国，GIS产业正在形成之中。“七五”期间在国家计划的指导下，采取引进、试验和开发的方针，为GIS的开发和应用积蓄了力量和经验。“八五”以来，GIS技术已进入实用阶段，在一些行业，如城市、土地和环保等的应用发展较快。据统计，截止1994年底，国外商业化的GIS软件在中国大陆的销售数量约为500套，其中以ARC/INFO、GENAMAP、Intergraph MGE、MAPVIEW等几种软件最多。国内软件SPACEMAN，Citystar，Geostar等的用户逐渐增多。从硬件平台支持看，微机略占多数，而工作站版本有增加的趋势。

第8—2节　GIS的工程建设

GIS的开发建设和应用是一项系统工程，涉及到系统的最优设计、最优控制运行、最优管理，以及人、财、物力资源的合理投入、配置和组织等诸多复杂问题。需要运用系统工程、软件工程等的原理和方法，结合空间信息系统的特点进行实施建设。GIS的用户不要认为购置了GIS硬软件就已大功告成。其实，在GIS中最有价值的是“信息”，而不是计算机等硬件。国际上的统计数据表明，一个信息系统如果硬件投资为1，则软件开发为5，而信

息的采集、整理和加工为 10。因此，在建立 GIS 过程中确定应用目标是什么，选用哪些数据源、什么数据入库，以及数据的质量、精度如何等一系列重大问题是至关重要的，直接关系到系统的有效性和实用性。也即 GIS 工程的成败及效益，取决于 GIS 工程的总体规划和设计、技术力量的组织、工程的建设实施和数据源的组织等问题。

一、技术力量的组织管理

GIS 工程建设首先要得到领导的支持，形成领导决策层。他们应该远见卓识，富于组织领导才能，不能急功近利。GIS 高级技术人才是系统建设不可缺少的。他们应当具有良好的测绘、地理、遥感、计算机和应用专业的知识基础，处理有关技术问题的能力和经验，较高的组织、决策才能。他们对 GIS 系统建设的顺利进行，并取得成功往往是至关重要的。一定要避免在总体设计或战略上发生失误，切勿小事精明，大事糊涂。其他技术人员包括数据采集、分析处理、程序编制、制图、应用技术、软硬件维护和业务管理等各类专业人员，各类人才要力求少而精，注意对他们进行不断培训和知识更新。

二、GIS 工程建设

GIS 工程的建设是一个较长时间的过程，涉及到许多环节。一般包括立项与可行性论证，用户需求分析，系统设计，系统实施，运行和维护等内容或阶段。

1. 立项与可行性论证

立项工作是 GIS 工程建设的起点，应在对用户深入调查基础上明确系统的总体目标、基本功能和要求。调研的内容包括系统建设的目的，数据源状况，经费支持能力，技术力量和工作环境条件，效益分析与评价，应用前景分析等。用户调研报告和可行性报告产生之后，最好组织有关专家进一步论证，最后由主管部门领导批准并取得人、财、物等条件的支持。

② 系统设计

工程项目较大时系统设计往往又分为总体方案设计和实施方案详细设计两部分或两阶段。总体方案设计的内容有：系统建设的目标，环境和原则，系统的配置和构成，子系统的划分及其接口设计，经费预算，实施总体步骤，运行管理办法等。

实施方案详细设计一般是在系统实施之前进行，有些内容也可在实施过程中进一步明确或细化。其内容包括子系统的划分及其功能，模块设计，空间数据库的内容和数据结构，数字化作业方案，应用模型，菜单、界面和图形显示设计，输入/输出设备、方式和精度设计，系统安全性设计等。最后产生一系列作业流程规范和技术说明文件。

3. 系统实施

这一阶段的工程量最大，主要任务和内容包括：硬软件设备的初安装和调试，数据准备和输入，数据库建立和子系统应用模式的研究开发，程序编制和调试，各子系统联网、运行测试，系统质量和水平的评判，测试报告、用户操作手册编写，总结。

4. 系统的运行和维护

系统的生存和效益不仅取决于上面所述的系统设计和实施，系统交付使用后的运行、维护也极其重要。其主要任务和内容包括日常管理、运行，数据更新和增删，系统功能的扩展和完善，硬件设备的更新或扩充等。

第 8—3 节　GIS 的应用模型

所谓 GIS 应用系统，实际上是根据某种或某几种具体的应用目标或任务要求，从相应专业或学科角度出发，对客观世界进行深入分析研究，并借助 GIS 软件工具的支持，使观念世界中形成的概念模型具体化为计算机环境或信息世界中信息系统的产物。或者说，某种 GIS 应用系统是客观世界中相应系统经由观念世界到信息世界的映射。这种系统的应用，反映了人类对客观世界利用改造的能动作用，并且是 GIS 技术产生社会经济效益的关键所在。因此，各种 GIS 应用系统的开发在 GIS 技术中占有十分重要的地位。

一、应用系统的数学模型

应用模型中的数学模型是在对系统所描述的具体对象与过程进行大量专业研究的基础上，总结出来的客观规律的抽象或模型。这种模型不仅是建立一个应用系统的主要内容之一，而且也是该系统解决实际问题的能力、效率及实际效益的关键所在，因此日益受到重视。由于各种系统的应用目标、复杂程度差异很大，故着力研究应用系统及其数学模型的个性十分必要。

（一）模型的作用和特点

（1）应用模型是联系 GIS 应用系统与常规专业研究的纽带。模型的建立绝非是纯数学或技术性问题，而必须以坚实而广泛的专业知识和专门研究为基础。对专业研究的深入程度决定着所建模型的质量和效果。

（2）应用模型是综合利用 GIS 应用系统中大量数据的工具。在系统中存贮有大量形式不同的数据源。它们的综合处理、分析应用主要是通过系统中模型的作用来实现的，也即系统中数据的使用效率与深度，很大程度上取决于模型的数量和质量。

（3）应用模型是 GIS 应用系统解决各种实际问题的武器。一般来说，应用模型能有效地帮助人们从各种因素之间找出其因果关系或内在规律，促进问题的解决。然而，由于许多问题十分复杂，完全依靠数学模型的定量方法往往很难完满解决，所以要结合使用人机交互的方法，采用从定性到定量的综合集成法。

（4）应用模型是 GIS 应用系统向更高技术水平发展的基础。大量模型的研究、开发和应用，凝聚和验证了许多专家的经验和知识，无疑也为一般 GIS 应用系统向专家系统发展打下基础。

（二）模型的分类

由于 GIS 属于空间信息系统，因此根据模型的空间特性，可分为空间模型和非空间模型两大类。然后各自再细分。这是由于空间和非空间两类模型在运算方式、所用数据种类、结果形式和管理方法等方面都有较大差别的缘故。见图 8—1。

非空间模型主要是对系统中的各种属性数据进行运算。常用的方法有投入产出、计量经济、经济控制论及系统动力学等。这些模型多用来解决社会经济领域中的一些问题，如评价、预测与规划等，有时也可用于生态环境及自然资源等领域。各种模型的形式、特点、局限性和主要应用范围见表 8—1。

空间模型需同时对系统中的图形和属性两类数据进行运算，一般要比非空间模型复杂些，且是 GIS 研究的重点和主要发展方向。表 8—2 对空间模型中常用的方法及其形式、特点、局限性和应用范围作了概括。在社会经济和自然科学两大领域中都常用到空间模型。

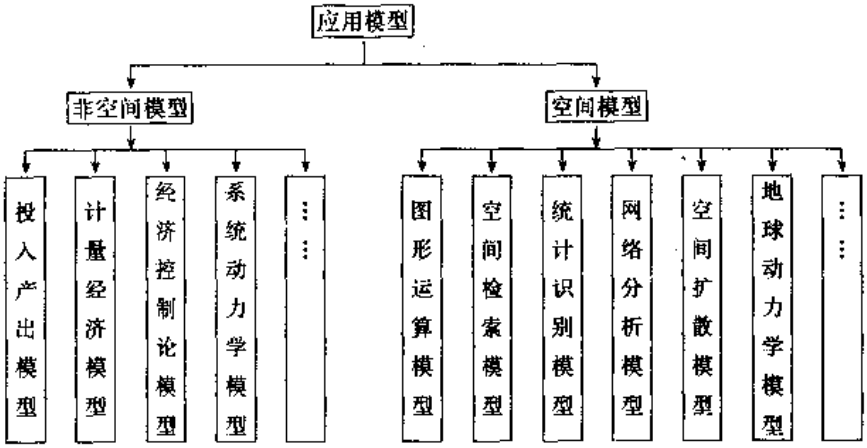


图 8—1　应用系统模型的分类

表 8—1　GIS 中常用的非空间模型

模型类别	模型形式	主要特点	局限性	主要应用范围
投入产出	以矩阵投入产出表为核心	能清晰地反映出各部门间的生产联系，模型简明	(1) 较难处理经济活动中的动态问题 (2) 没有考虑最终需求的确定以及生产与收入之间的反馈联系	(1) 是搞好综和平衡的重要工具，主要用于生产系统的平衡问题 (2) 经济—环境预测的重要方法之一
计量经济	线性或非线性联立通用方程	将经济理论、统计学，数学和计算机仿真技术有机地结合在一起	(1) 高质量的经济数据难于获取 (2) 当经济结构等变化较大时，模型难以及时作出反应	(1) 经济结构分析 (2) 经济政策评论 (3) 经济预测，但预测期不长
经济控制论	以离散形式的状态空间模型为主	能反映国民经济中的调节控制机制，可充分利用控制论的现代化成果	目前实用的模型还处于开发阶段	(1) 能进行各种期限的预测 (2) 能进行战略研究和制定最优经济政策 (3) 能作为制定五年计划的辅助工具
系统动力学	可带有时滞的一阶微分方程	以反馈控制理论为基础，便于处理非线性和时变现象，能作长期的、动态的、战略性的仿真与研究	预测精度不高，较适于研究分析系统的动态行为	(1) 在制定国民经济的远期发展规划时，作战略研究和政策分析 (2) 土地承受力、生态环境系统调控研究与仿真

二、模型的开发方法

一般的 GIS 软件都具有一定的空间分析、评价的功能模型。但对各种类别的用户来说，一些模型往往难以直接使用。象 ARC/2NFD 这样的通用软件，由于着眼于模型的通用性，必将难以适应用户多种多样的特殊要求，因此，许多具体的应用分析模型就得用户自己开

表 8—2 GIS 中常用的空间模型

模型类别	模型形式	主要特点	局限性	主要应用范围
图形运算	图内各点及图间同名点的算术或代数运算	反映图内各点以及图间同名点之间的空间关系，运算简单	以网格数据结构较易实现，矢量数据则较复杂	(1) 数字地形模型及应用；(2) 区位分析；(3) 各种较简单的回归模型
空间检索	搜索并检出符合条件的网格或图形	适用于目前还不能用数学公式描述，但却能根据专业研究给出的某些条件或阈值解决问题的模型	检索条件及阈值的确定要有专业研究的支持，如果它们不合理或不准确，会产生很大误差	(1) 大型厂矿及工程选址选线；(2) 根据成矿条件寻找靶区；(3) 土地适用性评价等
统计识别	统计识别分类及主成分分析等图像处理方法	移值图像处理中的有关方法，可进行区域划分及综合指标确立等应用	仅适用于网格数据结构	(1) 进行专题或综合分区；(2) 建立综合评价指标及模型
网络分析	图论及运筹学方法	解决与网络有关的各种问题。如最短路径，运量分配等	实际情况较复杂，计算量很大	(1)最佳路线选择；(2)网络中流量的分配；(3) 货物的集散调度；(4) 城镇及生活服务网点布局
空间扩散	解扩散方程、数学—力学方程	描述了许多现象的物理过程	在实际计算时，参数的选择有较大的经验性。使结果的精度受一定影响	(1)大气和水环境质量预测；(2)热污染模型建立；(3) 人文地理问题的解决；(4) 地球动力学问题

发。就 GIS 应用的二次开发而言，可将开发方法归结为下列三类：

1) 内部空间模型法　这种方法又可称为完全（紧密）结合法。它又可分为两种，一是由 GIS 软件商直接提供的，二是用户自己利用 GIS 软件的宏语言，如 ARC/INFD 的 AML，System 9 的 ATP 开发各自所需的空间分析模型。这种建模法是把 GIS 软件支持的功能看作模型部件，按照分析目的和标准，对部件进行有机的组合。因此，这种建模法能充分地利用 GIS 软件本身所具有的资源，效率较高。

2) 外部空间模型法　这种方法基本上是将 GIS 当作一个空间数据库看待，应用空间模型在 GIS 环境外部利用其它软件或计算机高级程序语言来建立。根据数据共享与否，又可分为两种：一是不与 GIS 空间数据库共享数据，而利用一个中间文件作为数据交换媒介；另一是开发的空间分析模型与 GIS 空间数据库共享数据。这种建模法虽然运行效率受到很大影响，但实现了软件的嫁接，无需在 GIS 环境中重新编制分析软件，适应性较强。

3) 混合型空间模型法　这是上述两种方法的结合，即尽可能利用 GIS 提供的功能，最大限度地减少用户自行开发的工作量，又利用外部空间建模法的灵活性。例如，自然语言命令或地图代数法就是属于这种方法。常用的自然语言命令举例如下：

(1) AVERAGE EXPOSURE TIMES 2 PLUS STEEPNESS TIMES 3 FOR COST ANALYSIS（益本分析）；

(2) SUBTRACT V MINUS U FOR CHANGE ANALYSIS（动态分析）；

(3) RENUMBER LANDUSE FOR WATER ASSIGNING 0 TO1—3.5—6，AND 1 TO 4（聚合分析）；

(4) ADD OVERLAY 1 TO OVERLAY 2 FOR OVERLAY3（叠加分析）；

(5) SPREAD ROADS TO X METERS（缓冲区分析）。

上述三种方法如何选用，不能一概而论，这要取决于用户对软件系统的理解深浅程度，以及所要解决问题（应用目标）的性质和复杂程度。

三、应用模型开发举例

（一）工程方案的评价模型

工程设计的好坏将直接影响到能否达到预期的目的要求。对工程方案进行模拟分析评价，预测可能出现的问题，选择最佳的方案，是十分必要的。影响工程评价的因素很多，利用 GIS 具有空间数据叠加、综合定量分析、模拟预测等功能，通过建立应用评价模型，对多种方案进行评价，从而为最终决策提供科学依据。

1. 评价指标的选取

合理选取评价指标直接影响评价结果是否符合实际。在选择评价指标时，一般应遵循以下原则：选取对工程方案影响大的因子；选取区域变异大的因子。

2. 评价指标的量化

重大工程方案的评价指标一般涉及多个学科专业，在对评价指标量化时可先由各方面的专家确定量化规则，然后由计算机根据规则算出每个指标相应的得分或隶属度。指标的量化标准可采用百分制，如表 8—3 所示。

表 8—3 评 分 标 准 表

等级	1 极差	2 很差	3 差	4 较差	5 稍差	6 中等	7 良	8 良好	9 优	10 极优
分数	0～10	11～20	21～30	31～40	41～50	51～60	61～70	71～80	81～90	91～100

3. 加权计分法模型

对参与方案评价的各评价因子依其重要性赋予不同的权值，按各评价因子的加权得分总计来综合评价方案的优劣。计算方法如下：

$$S = E \cdot W$$

$$\begin{bmatrix} s_1 \\ s_i \\ s_m \end{bmatrix} = \begin{bmatrix} e_{11} \cdots e_{n1} \\ e_{1i} \cdots e_{ni} \\ e_{1m} \cdots e_{nm} \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_i \\ w_n \end{bmatrix} \qquad (8—1)$$

式中，s_i 表示第 i 种方案的综合评价得分，s_i 值越大，表示该方案越好；E 为各评价因子的标准分评价矩阵，按百分制给分；W 为各评价因子的权向量，其约束条件为

$$\sum_{i=1}^{n} w_i = 1 \qquad (8—2)$$

各评价因子权值大小，可由各评价因子的重要性通过层次分析法确定，或者由有关专家直接确定。

由式（8—1）可得每个工程模拟方案的总评分为

$$s_i = w_1 e_{1i} + w_2 e_{2i} + \cdots + w_n e_{ni} = \sum_{i=1}^{n} w_i e_{ii} \qquad (8—3)$$

由上式可见，方案评价中重要性越大的因子其权值越大，对总分的影响也越大。若一个方案中重要性较大的因子都满足要求，那么该方案的总评分肯定很高。反之，则总评分

低。这种方法比较符合人们通常的决策思维方式。

4. 多目标模糊评价模型

有些工程方案的评价目标常具有不确定的边界，有模糊性，可运用模糊数学方法将模糊信息数值化，以进行定量评价。

工程方案评价一般要涉及多个评价目标，多目标的模糊评价步骤是，先求各评价目标及加权系数的模糊评价矩阵，再根据模糊矩阵合成规律合成，求得该方案的综合模糊评价集。

设工程方案的模糊评价目标集 $F=\{f_1, f_2, \cdots f_n\}$

评价集 $X=\{x_1, x_2, \cdots x_n\}$

加权系数集 $A=\{a_1, a_2, \cdots a_n\}$

这里，$0<a<1$，$\sum_{i=1}^{n} a_i=1$。

某方案对 n 个评价目标的模糊评价矩阵为

$$R=\begin{bmatrix} r_{11}\ r_{12}\cdots r_{1j}\cdots r_{1m} \\ r_{i1}\ r_{i2}\cdots r_{ij}\cdots r_{im} \\ r_{n1}\ r_{n2}\cdots r_{nj}\cdots r_{nm} \end{bmatrix} \tag{8—4}$$

其中，r_{ij}为第 i 个评价目标。f_i 称为 x_j 级的隶属度。且 $0<r_{ij}<1$，$R_i=1$。于是考虑加权的综合模糊评价

$$B=A\cdot R=[b_1, b_2, \cdots b_j \cdots b_m] \tag{8—5}$$

模糊矩阵的合成有多种模式，在工程方案评价中通常采用乘加运算（$M(\cdot, +)$）进行矩阵合成

$$b_j=\sum_{i=1}^{n} a_i r_{ij} \qquad j=1, 2, \cdots, m \tag{8—6}$$

这种运算方法综合考虑了各项隶属度和权重的影响，因此评价的效果较好。对方案优劣定级时，遵循最大隶属度原则。方案之间比较时，以同级中隶属度高者为先，级别相同时可用本级更高的隶属度之和作比较。

（二）城市温室效应要素分析模型

利用遥感资料提取的植被和城市用地类型等信息，通过构建的模型分析它们与地面温度的相关关系，以定量描述影响城市温室效应的主要因素，从而为城市规划决策中缓和温室效应提供依据。该分析模型的结构如图 8—2。对该模型的建立进一步说明如下。

1. 土地利用分类模块

该模块主要应用 GIS 软件提供的图像处理功能，首先读进（TM）图像 1～5、7 波段，然后分别对它们作几何纠正、选择训练区、监督分类、精度校核，如有必要则重新选择训练区或增加训练区，直至分类精度符合标准，以使城市用地分类数据的精度满足要求。

2. 植被指数计算模块

该模块采用标准植被指数（*Normalized Vegetation Index*）进行计算，结果可反映地表植被量的多少。具体方法为

$$NVI=\frac{NIR-VIS}{NIR+VIS}\cdots\cdots \tag{8—7}$$

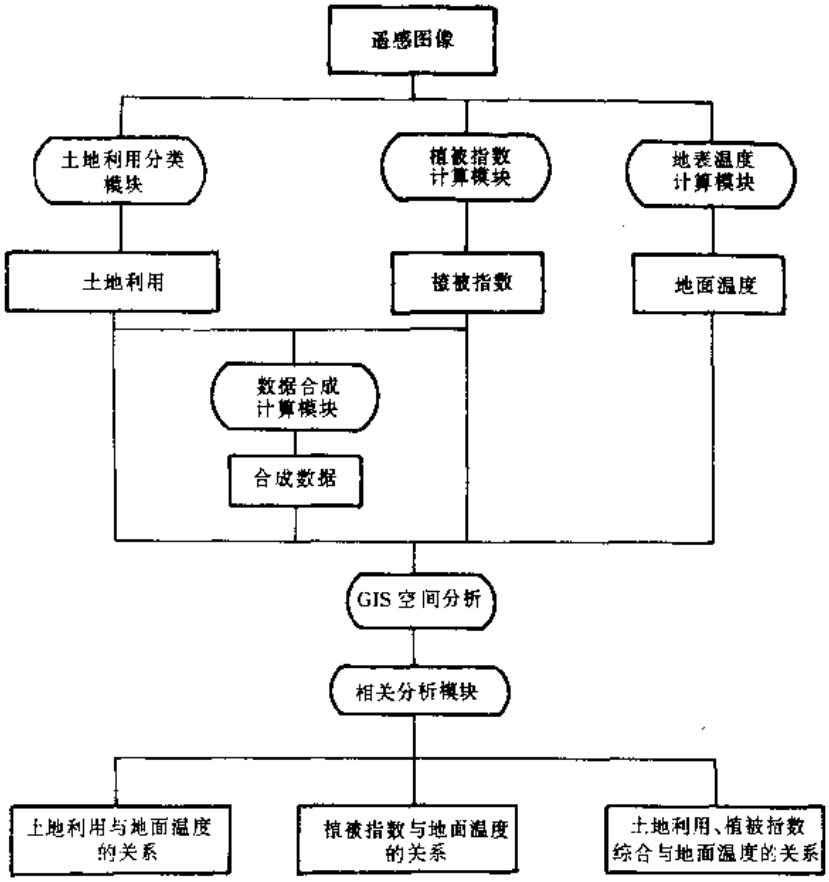


图 8—2　温室效应要素分析模型结构图

式中，NIR 为近红外波段的灰度值；VIS 为可见光波段的灰度值。

这里，计算出的 NVI 值域为−1.0～1.0。为便于显示输出，可将它转换成 0～255。这样式（8- 7）就转换为

$$NVI=\frac{NIR-VIS}{NIR+VIS}\times 127+128 \qquad (8—8)$$

这样，通过对（TM）图像的 4 波段和 3 波段的计算，便可得每一像元的 NVI 值。

3. 地面温度计算模块

（TM）图像的 6 波段为热红外波段，它所记录的信息主要反映地面的温度。放射辉度与温度的关系式为

$$N=(5.1292\times 10^{-5}T-0.017651)\ T+1.6023\cdots\cdots \qquad (8- 9)$$

式中，N 为放射辉度；T 为绝对温度。

同时，补正后的灰度值与放射辉度有如下关系式：

$$N=V/255\ (N_{max}-N_{min})\ +N_{min}\cdots\cdots \qquad (8- 10)$$

式中，V 为灰度值；N_{max} 为 340K 时的放射辉度；N_{min} 为 200K 时的放射辉度。

根据式（8—9）和（8—10）可求出灰度值与地面温度的关系，结果为

$$1.40654/255\times V=(5.1292\times 10^{-5}T-0.017651)\ T+1.47852\cdots\cdots \qquad (8- 11)$$

于是，可从 6 波段图像中的灰度值计算出每个像元的地面温度值。

4. 数据合成计算模块

该模块将城市用地类型数据与植被指数统一考虑，计算两者对地面温度影响的程度，数据综合方法为

$$T_C = aT_{LC} + bT_{VI} \tag{8—12}$$

式中，T_C 为综合影响数值；T_{LC} 为根据城市用地类型推定的温度；T_{VI} 为根据植被指数推定的温度；a 为根据城市用地类型推定的温度与地面温度的相关系数；b 为根据植被指数推定的温度与地面温度的相关系数。

通过这一模块的计算，可得到每一像元的综合影响数据，用于与地面温度进行相关分析。

5. 相关分析模块

该模块可以采用一元或多元，线性或非线性相关的方法分析两变量之间的相关性及其相关程度。回归直线方程及相关系数为

$$y = ax + b,\quad r_{xy} = \frac{\delta xy}{\delta x \delta y} \tag{8—13}$$

根据最小二乘法的原理，可求得

$$a = \frac{\delta xy}{\delta^2 x},\quad b = \bar{y} - \frac{\delta xy}{\delta^2 x}\bar{x}$$

其中，$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$，$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$，$\delta^2 x = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$，$\delta^2 y = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2$，

$\delta xy = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$

通过这一模块的计算，可定量地反映地面温度与植被、用地类型等信息的相关程度、变化趋势及确定它们关系的回归方程。一元线性回归方程分析最为简单，一般也能满足要求。有时也要求采用多元回归或非线性回归的分析方法。

应该指出，城市温室效应的影响因素往往是较复杂的，多种多样的。除了上述因素之外，水体面积的大小、植被的种类与分布状态、高层建筑区的阴影等，也与地面温度有密切关系。

第 8—4 节　国家基础地理信息系统

国家基础地理信息系统（NFLIS）是目前我国最大的全国性 GIS。它是国家测绘局的专业信息系统，也是国家经济信息系统网络体系中的一个基础子系统。中国测绘科学研究院从 1984 年开始研究、开发它。

NFLIS 是以形成数字信息服务的产业化模式为目标，以国家基本比例尺地形图、测绘控制点、航空和航天遥感影像等为主要信息源的一种 GIS。

NFLIS 的地形数据库是空间型 GIS 数据库。它是将国家基本比例尺地形图上各类要素包括水系、交通、境界、居民地、地形、植被等，按照《国土基础信息数据分类与代码》国家标准（GB13923—92）进行分类编码，并按照一定的规则分层，对各要素空间位置、附属属性信息及互相间的空间关系等数据进行采集、编辑处理建成的数据库。考虑到数字形式数据处理的特点及数据综合的局限性，全国范围地形数据库的比例尺大体分为三级，即 1∶100 万、1∶25 万和 1∶5 万（或 1∶10 万）。

地名数据库是与地形数据库配套的空间定位型关系数据库。它是将国家基本比例尺地

形图上各类地名注记进行分类编码，连同其标准地名、所说明实体的空间位置（经纬度、投影坐标）及其它有关数据项录入数据库中。它与地形数据库之间具有技术接口，以便互相连接、互相访问。

NFLIS 的主要数据源及逻辑结构如图 8—3 所示。这些不同的数据源通过数据格式转换、几何配准和接口处理，构成统一的空间数据库。

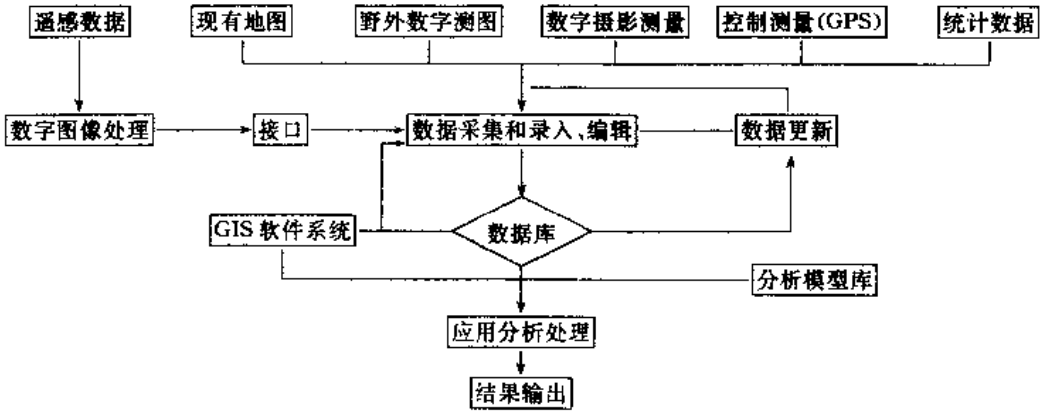


图 8—3 NFLIS 逻辑结构框图

NFLIS 地形数据库具有如下特点：基础性强、数据量大；数据质量控制严格；数据规范化、标准化程度高；具有多种数据格式相互转换功能；图形数据与属性数据之间有良好的接口；能保持数据的现势性；具有明确的产业化方向；试验性与实用性并举，达到实用化目标；用户界面友好；具有广泛的应用前景，社会经济效益显著。

NFLIS 是为适应多层次、多方位用户需要而设计的。该系统的应用大体有三个方面，即快速查询检索与分析，提供专题信息空间定位基础，以及加快地图成图和更新周期。

系统应用＝系统功能＋基础数据＋专题（含统计）数据＋分析应用模型。此外，系统应用的能力还取决于：良好的数据库设计，数据规模和覆盖范围，数据质量和标准化程度，以及数据的维护更新能力。该系统应用的基本模式大体包括：

（1）提供快速查询检索与分析数据；

（2）统计分析与统计制图；

（3）专题信息叠加分析和专题制图；

（4）为各种专业信息系统提供统一的基础数据；

（5）快速成图和更新地图；

（6）分析应用模型或专家系统；

（7）提供标准数字产品；

（8）为用户设计和开发各自的系统提供技术服务，推动 GIS 技术的发展和普及。

第 8—5 节　土地资源管理信息系统

土地资源管理信息系统也可简称为土地信息系统（Land Information System，LIS），是 GIS 的一个分支，是一种基于宗地（以宗地（地块）为单位）的计算机管理信息系统。

随着国民经济建设的发展，日益要求对土地资源和房地产进行科学的管理；同时随着计算机技术应用的日益广泛，将计算机技术用于土地管理已成为必然趋势。在西方发达国

家如德国、瑞典、荷兰、加拿大、美国等很早就实现了土地管理的自动化。在亚洲的日本、泰国以及香港、台湾等地区也逐步开发了LIS。这些系统是以土地管理为中心，为其它行业提供及时的、准确的土地和相关信息。该系统不仅可以完成土地登记，还可进行土地估价、税收和规划等方面的支持工作，可创造显著的社会和经济效益。在中国，首先在一些沿海城市已着手研建LIS，但大都还不够完备，只有一些子系统开始运行，比如土地利用详查和成图系统，地籍图自动制图系统等。因此，建立一个兼容土地管理、估价和规划为一体的综合性LIS是我国许多省区、城市的重大任务。这里，需要有一系列土地管理的政策、法规和社会主义市场经济体制完备性的配合。

一、土地信息系统的必要性和可能性

虽然手工操作管理方式可以完成许多日常的土地规划、管理工作，但经济建设的发展、土地有偿使用的实施和信息社会的发展，日益要求科学的现代化的管理方式和手段，而手工操作方式已越来越不适应当今的要求。主要表现在以下几个方面：

1）数据量巨大　数以千万计的宗地图形描述和权属描述，以及土地利用类别等的描述数据量很大，需要用高速的空间信息系统来处理。

2）要求地籍信息和地籍图的现势性　地籍信息的变更比较频繁，如土地利用类型、权属或宗地的重划、合并等，手工作业已难以满足要求。

3）科学地进行土地利用规划　城市规划建设的现代化、科学化要求土地利用与规划的最优化、最佳配置，只有LIS才能适应这一要求。

4）土地估价和税收、房地产交易　土地使用权可以转让、抵压、出租。土地的税收率不仅取决于土地利用类别，还取决于地价。房地产市场的规范化、国际化等，要求必须使用先进的计算机管理技术。

5）土地信息咨询和检索　各种土地信息和数据可提供各行各业客户使用，但是在种类众多、数据量巨大的情况下，询问和查找工作唯有LIS能够胜任。

那么，LIS具备些什么功能和优越性而成为土地资源规划开发和房地产管理的有力工具呢？

利用LIS，土地信息可以很快地进入计算机系统，且根据用户的要求可以快速从图形查得权属信息；反之，也可以根据权属情况查得具体宗地的位置、形状和相邻关系，以及面积等数据。

利用LIS，不仅可以快速地查询信息，还可以进行分析、评价。运用分析模型可以对土地信息、数据进行深加工，如土地估价、土地利用规划。

利用LIS，可以对宗地信息、权属关系进行更新、增删、查找，绘制新的地籍图。

土地信息系统与办公自动化结合，可以极大地提高土地管理的业务水平和工作效率。

二、LIS的基本硬软件配置要求

1．一般要求

由于LIS的专业性特点，对软硬件环境也有一定的要求，主要体现在：

1）微型化　为便于管理、维护以及低成本和易于更新的要求，一般采用微型机或工作站。

2）图形处理功能　土地利用图、地籍图、宗地图等的图形数据处理和输出是系统必备的功能。

3）大容量存储能力　为满足巨大的数据量要求，可用光盘或大容量硬盘存储方式，并具有高速度数据检索和图形处理能力。

4）汉字化　大部分地籍信息应以汉字表达，以便于使用。

5）数据标准化和网络化　为了便于不同地域的LIS实现数据共享和交换，应实现数据的标准化和网络支撑。

此外，还应具备足够齐全的适用的输入/输出设备，如数字化仪、扫描仪、打印机、绘图仪等。

2. 软件配置方案

根据不同的应用条件，软件选择可适当灵活。下列软件配置仅供参考：具有中文界面的操作系统X—WINDOWS，MS—WINDOWS等；具有GIS或LIS的基本软件；通用关系数据库，如Paradox、Foxpro、Oracle、Informix、Sybase等；网络软件$3C^{+}$、Novel、MS—LanManger等；最好还有可供二次开发的应用软件（模块）等。

3. 硬件配置方案

根据数据量大小和软件要求来选择硬件，下列配置仅供参考：

工作站机：486/66或更高档、8MB内存、VGA彩显、500MB硬盘；

网络服务器：以Intel Pentium(586) DEC Alpha、Power PC为主，2GB硬盘、16/32MB内存、只读光盘、网卡、备份磁带机；

A_1或A_0幅面的绘图仪、数字化仪。汉字激光打字机一台，24针汉字打印机若干台。

三、建立LIS需要解决的一些主要技术问题

（一）利用遥感图像资料进行土地类型研究

利用航空遥感或卫星遥感资料进行土地利用调查的方法，具有更新快、现势性好等优点，但存在图像判读、数字化、图形拼接、编辑、面积量算等技术问题需要解决。我们以航空彩色红外像片为信息源，利用VAX3100工作站为硬件平台，开发了土地利用调查与分析系统（CMLUS），主要功能有：数字化数据采集、几何纠正、拓扑结构的建立、多边形数据生成、面积统计、图形拼接、线划图和充填图的输出、报表输出等。该系统软件利用VAX FORTRAN、VAX汇编、图形核心系统（GKS）和VMS提供的有关系统软件编程，采用菜单式用户界面。CMLUS的逻辑结构如图8—4所示。其基本工作原理简介如下。

1. 像片数字化

像片数字化软件设计有两种功能：一是直接数字化多边形边界点，产生多边形顺序结构数据文件；二是按链段数字化多边形边界，产生链式结构的数据文件以及拓扑和属性文件。

含有链段的拓扑信息（即链段的左、右区和始终点编码以及属性左右区的土地类型）。对于图8—5所示的链段，其拓扑信息与属性如表8—4所示。

表8—4　链段的拓扑信息与属性

链段	区号		土地类型		结点		链段长度
	左	右	左	右	始点	终点	
1	0	(2)	0	521	2	3	2
2	(3)	(2)	15	521	7	6	8
3	⋮	⋮	⋮	⋮	⋮	⋮	⋮

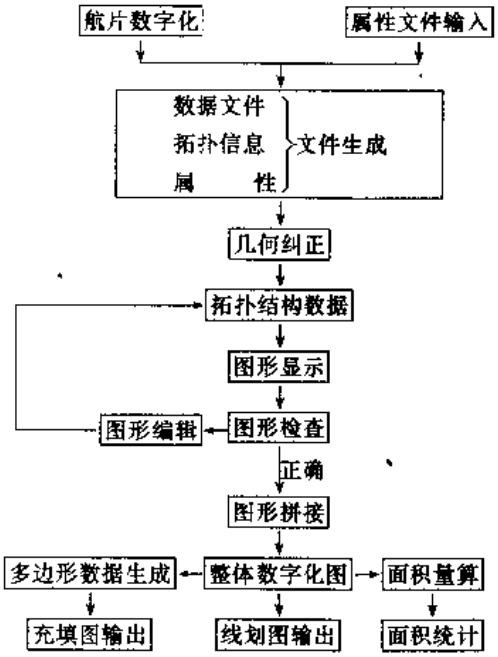


图 8—4　利用遥感图像资料作土地利用调查研究逻辑框图

2. 几何纠正

航片数字化数据的几何纠正，最好的方法是利用 DTM（地面数字模型）对航片进行正射投影改正。在进行几何纠正之前，必须建立该区的地面数字模型。但是，我国绝大部分地区没有 DT M 数据。建立 DTM 的方法之一是在地形图上采集离散的高程点，然后按移动加权平均（也可用其他方法，如 Kriging 法）进行内插，求得栅格 DTM。因建立 DTM 需增加不少工作，所以我们选用多项式拟合纠正。

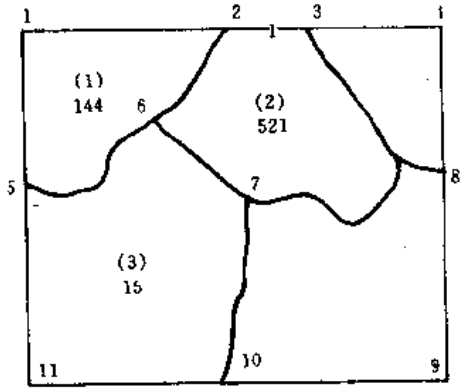


图 8—5　链段的编码

假设 (l, p) 为像片数字化后的坐标，其对应的地面坐标 (X, Y) 可表示为

$$X=f_X(l, p)=\sum_{k=0}^{n}\sum_{j=0}^{k}A_{jk}\cdot l^j\cdot p^k \quad (8—14)$$

$$Y=f_Y(l, p)=\sum_{k=0}^{n}\sum_{j=0}^{n-k}B_{jk}\cdot l^j\cdot p^k$$

式中，n 为多项式的次数；A_{jk}、B_{jk} 为多项式的系数。

如果在地形图上量得一组控制点坐标 (x_i, y_i) $(i=1, \cdots m)$，与这些点对应的像片数字化坐标为 (l_i, p_i)，用 (l_i, p_i) 代入式 (8—14) 可得 (X_i, Y_i)，根据最小二乘原理

$$\left.\begin{aligned}\sum_{i=1}^{m}(x_i-X_i)^2=\min\\ \sum_{i=1}^{m}(y_i-Y_i)^2=\min\end{aligned}\right\} \quad (8—15)$$

可求得两组多项式 (8—14) 的系数 A_{jk}、B_{jk} $(j, k=1, \cdots n)$。为了求得上述系数，控制点数目 m 应大于 $(n+1)(n+2)/2$。若用三次多项式，控制点点数应多于 10 个。求得系数

后，利用像片数字化坐标按式（8—14）求得各点的地面坐标，则完成像片数据的几何纠正工作。

3. 多边形数据的生成

像片数字化后，不同土地类别的多边形边界是以链段形式存贮的。为了便于充填显示和覆盖分析，必须将链段数据转换成多边形数据。由于每个节点都被数字化多次（若 n 条弧段相连，则数字化 n 次），数字化的误差使同一节点有不同的坐标。在组成多边形之前，必须对每个节点进行检查，若同一节点各次数字化的坐标的最大差值在限差之内，则取平均值作为该节点的坐标。这一工作可通过寻找拓扑与属性、文件中的节点号以及链段长度（相当于指针）来实现。全部节点坐标平差后，根据左右区码及始终点号按顺时针寻找各条链段，并使其相连组成多边形数据。

4. 面积计算

面积计算与统计是土地利用调查的主要任务之一。虽然根据多边形顶点坐标能够计算出多边形的面积，但它对多连通区域无效，即不能扣除中孔多边形的面积。为解决这一问题，我们采用拓扑方法计算面积。首先计算每一链段的始终节点投影至 Y 轴所包围的多边形面积（如图 8—6 所示的阴影线面积为链段 1 所包围的面积）

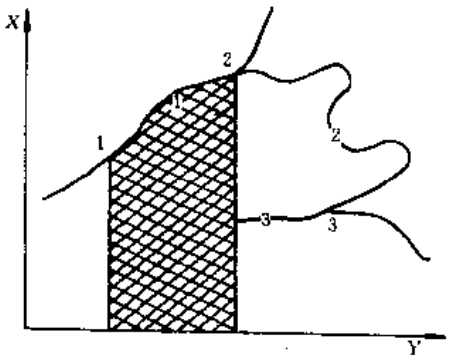


图 8—6　面积计算

$$R_k=\frac{1}{2}\sum_{i=1}^{u}(x_{i+1}+x_i)\cdot(y_{i+1}-y_i)$$

然后求由各链段组成的多边形的面积

$$S=\sum_{k=1}^{m}R_k\cdot\varphi(k)$$

式中 $\varphi(k)=\begin{cases}1 & \text{该多边形位于第 } k \text{ 条链段的右区}\\ -1 & \text{该多边形位于第 } k \text{ 条链段的左区}\end{cases}$

m 为组成该多边形的链段数；l_k 为第 k 条链段的长度，即数字化点数；(x_i, y_i) 为链段上各点之坐标。

计算出各个多边形的面积后，按分类号统计各类土地的面积。

5. 图形拼接

像片数字化及纠正工作是一张张像片独立进行的，由于判断误差、纠正误差和数字化误差使得相邻像片边界的位置和属性不吻合。此外，航片数字化的链段、节点等编号是独立体系，因而要将每张像片数字化的结果拼成整幅图并非易事。虽然有些GIS软件提供了拼图的功能（如 ARC/INFO 的 MAPJOIN），但未见文字描述，且有许多不足之处。“无分幅系统”（Sheetless System）的概念，可在相邻数字图之间建立整体联系。我们设计的软件能够解决拼图问题，有时适当加以人工干预，可起检查作用。

6. 土地利用图的输出

图形软件用 VAX FORTRAN 和 GKS（图形核心系统）编制。所谓语言独立（Language Independence）就是所有标准编程语言都能调用它的软件；而设备独立（Device Independence）就是可在不同型号、类型的设备上输出，且不必改动应用程序。GKS 的主要功能有：控制输出，输入、变换、图段（Segment）生成、图形元文件（Metafile）生成、询问、错

误处理等功能。

我们利用GKS开发了显示与绘制充填图和线划图的软件。该软件可以将图形显示在屏幕上检查、修改，也可在绘图仪或打印机上输出。要在不同的设备上输出，只需将工作站的类型参数定义为输出设备的连接符，而不必改动程序。此外，控制板、鼠标器、光笔也为图形编辑提供了很大的方便。

（二）编码方案设计

在LIS中对多种类型的数据都要给予适当的标识或分类代码。编码方案的优劣对系统的数据组织方法及系统的运行效率有重大影响。

（三）信息内容应主副兼顾

系统中的信息内容根据其服务对象会有所不同，但应以和土地管理相关的信息为主，兼顾一些其它信息，如供电、供水、通讯、人口，甚至矿产资源、地质信息等。

（四）分析模型的建立

为了对一些复杂问题进行分析、评价和计算，如土地估价，房地产交易等，需要研建一些分析数学模型。

（五）其它问题

如良好的用户界面，多媒体功能，历史性数据的存储，数据的安全性等。

第8—6节　城市规划管理信息系统

城市规划、建设与管理是一门社会、经济、环境、技术、艺术高度综合性的科学。主要研究对象是组成人类社会的人、人的活动、活动所依赖的环境和资源之间的关系，以及这三大要素的科学规划和管理。城市规划信息系统，有时又称城市地理信息系统（Urban GIS），以计算机硬软件为支撑，多学科知识为依据，空间和非空间数据为基础，集知识、模型、决策于一体，完成复杂的城市规划和管理任务。它是传统的手工作业方法难以胜任的。

城市规划管理信息系统，简称UGIS。它的建设和应用是一项涉及面广、工作量大、周期长的复杂工程，是GIS应用中较活跃的领域之一。由于城市的大小规模、功能、特点有很大差异，因此UGIS建设的一项重要任务，就是在GIS现有基本功能的基础上，根据具体的城市情况和应用要求，研究、开发有针对性的城市空间分析模型及应用功能。

UGIS的逻辑结构和基本功能见图8—7。它的主要功能和规划管理的分析模型包括：

1. 城市建设用地规划与评价

基于综合考虑用地的自然和社会经济因素，分析评定规划区域内建设用地的适宜性、开发的时序和地价等。

2. 交通流量预测分析

3. 城市环境质量评价

根据清洁、美观、舒适的基本原则对城市自然环境质量和社会环境质量进行综合评价分析，以确定环境质量分区、制定治理措施。

4. 景观分析

城市景观分析是根据视觉分析的原理和景观分析的视角过程，考虑视点、视角、视距、视场等基本视觉因素，在不同的空间层次和尺度上分析城市景观的视觉形象和布局结构，模

拟景观控制规划的视觉效果，以使城市的整体景观优美化。

5. 城市灾害的预防分析评价

根据城市各区位的人流、物流、能量流、信息流的负荷状态，生态环境平衡原理，应用增熵原理、界面原理等方法，对各种城市灾害（交通事故、火灾、水灾等）进行定量与定性相结合的预测评价，提出防灾减灾的规划、对策，以防止或减少灾害的发生和损失。

6. 建设项目选址分析

根据拟建项目的性质、规模和功能目的要求，利用 GIS 中的覆盖分析、缓冲区分析、空间查询方法等功能，基于建立的目标函数和约束条件选择最佳的布局点位。

7. 竖向规划分析

确定规划道路交叉口标高、纵堤转向点标高和场地平整控制标高，生成规划区内的控制标高，以适应排水、交通和城市景观的要求。

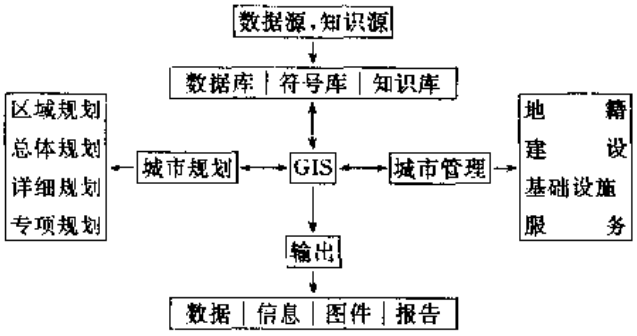


图 8 7 城市规划管理信息系统的逻辑结构框图

8. 市政管线规划管理与分析

城市中地面、地下的各种管道、线路很多，应进行科学的规划管理，对其空间和时间状态信息进行存储、分析和评价。

第 8—7 节 区域资源与环境信息系统

资源与环境信息系统，简称 REIS。此系统作为一种综合处理和分析空间数据的有效手段又通称为地理信息系统。90 年代以来，它同 GPS 和 RS 密切结合，共同发展，在全球和区域资源与环境动态监测、趋势预报，重大自然灾害监测与预报，以及灾情评估与减灾对策，城市、经济开发区的规划，开发与管理，全球环境变化等方面发挥了科技的先导作用。

一、自然资源的信息分类

在讨论资源与环境信息系统时首先涉及到的内容是资源，这里主要是指自然资源的含义和自然资源信息的分类问题。自然资源信息分类是自然资源实体分类的反映，因此信息分类应以实体分类为基础，并反映实体的特征及其相互关系。目前尚无自然资源分类系统的统一标准。袁朴在“关于自然资源信息分类的研究”一文中参照有关文件资料，提出了一种分类体系，见图 8—8。它所遵从的主要原则如下：自然资源在自然界中的空间位置；自然资源的属性；自然资源的用途；自然资源的开发利用条件。在具体分类时不能只简单孤立地考虑某一原则，而应综合考虑多项原则，进行多层次、多目标的分类。

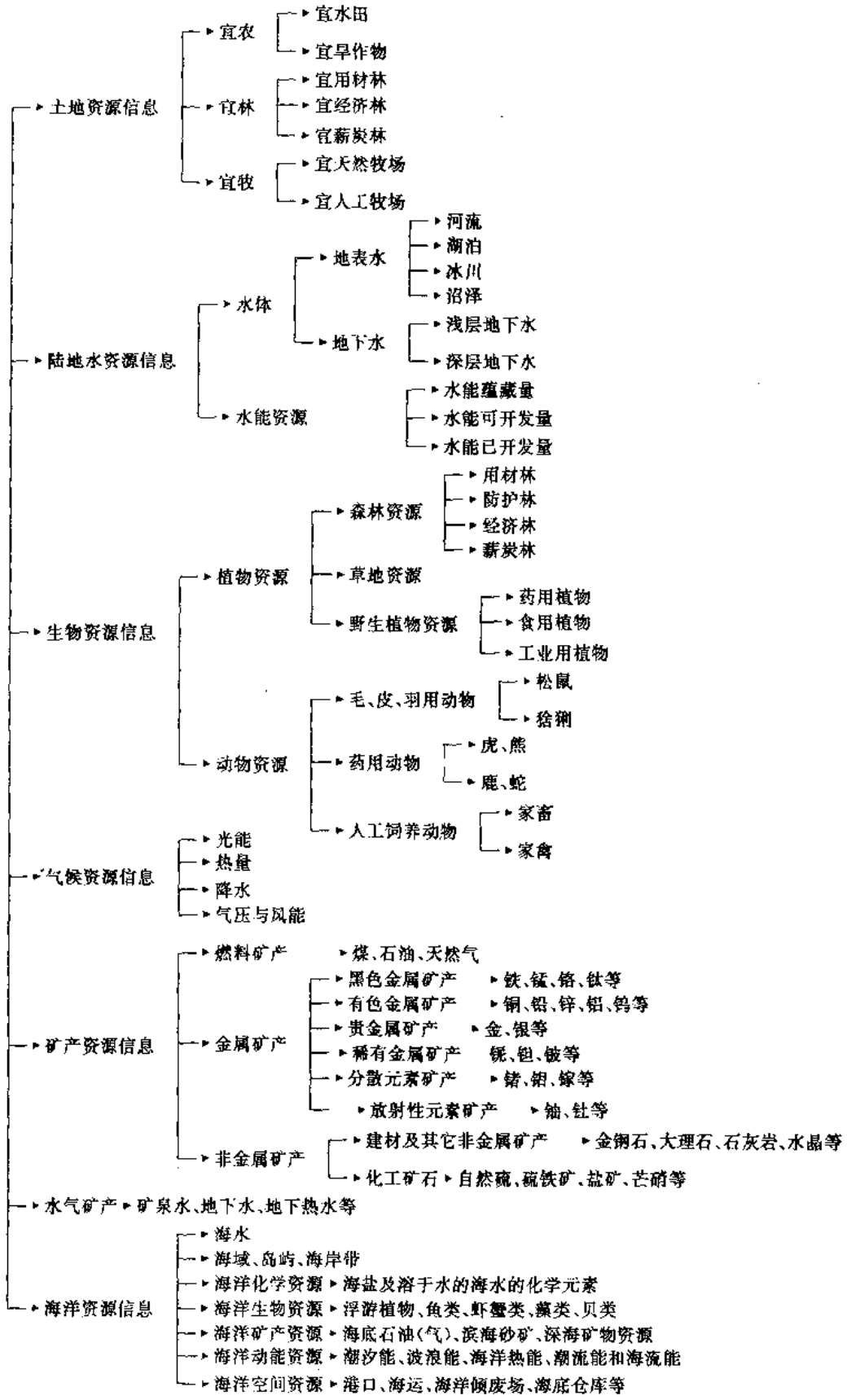


图 8-8 自然资源信息分类

由于自然资源和环境的内涵及种类十分丰富、复杂，因此相应的资源与环境信息系统的种类、功能和目标是千差万别的。这里只是介绍REIS的一些梗概，并且在上面几节中实际上已论及了土地、水及城市环境的信息管理问题。

二、全球资源与环境数据库系统

由于人类过量采伐森林，特别是被称作地球之肺的热带雨林被大面积采伐，造成全球气候变化，使非洲大面积干旱和沙化，亚洲和北美洪水成灾，人类生存环境在不断恶化。1987年联合国开始实施一项环境计划（UNEP），计划包括建立一个庞大的全球环境变化监测系统（GEMS），旨在利用现代高科技成果，如卫星监测、GPS和GIS，对全球环境变化实施监测。GEMS是靠全球资源与环境数据库（GRID）系统支持的，GRID包括十几种应用软件包，这些软件包是整个系统的核心。它完成所有数据的获取、输入、处理、分析和输出产品等任务，同时承担联接GIS网络。经过几年的操作运行，目前GRID可为全球的GIS服务网络提供实时的资源和环境动态变化信息，为联合国UNEP计划的顺利完成，有效地协调各国森林资源和环境的保护，提供可靠的数据，起投资导向作用。

GRID在瑞士的日内瓦和肯尼亚的内罗毕所设的两个中心，都装备了很强的数字图像处理和GIS设备，以确保完成数据处理并提供良好的服务。

GRID的硬件包括IBM微机、工作站，PS/ZS和DEX VAX微型机以及输入/输出外围设备。图像处理和GIS软件有IAX、GEIS、ARC/INFO、ERDAS、IDRISI、CHIPS、ILWIS和SPANS。目前已实现了与十几个国家联网，形成庞大的全球信息网。

从1980～1990年GRID主要开发了全球森林监测和森林生态变化的研究和软件开发项目。采用的遥感数据包括AVHRR—LAC、LANDSAT—TM和SPOT数据。1990年完成了利用AVHRR卫星遥感数据对西非森林覆盖面积填图和对亚马逊地区原始森林砍伐状况的调绘。经过处理、分析的数据按统一坐标系存入GRID数据库。近年来数据库中的数据得到进一步扩充，扩充的重点是全球热带雨林覆盖面积系列图件数据。此外，拟在西亚再建一个GRID中心，并与上述两个中心之间通讯和信息交流。

GRID中心还为用户提供下列五种服务：世界森林资源的年度评价数据；GRID数据库中不同数据分析和处理结果；AVHRR/LAC热带雨林指定地区实时图像数据处理和分析；GRID数据库中整个热带雨林地区数据分析和综合处理结果；提供设备和专家咨询服务。

三、水资源动态分析预测系统

合理开发利用陆地水资源是当代科学的重要课题之一。根据地球水均衡原理，高山融雪和大气降水是世界主要河流的补给来源。对于研究大面积涉及水均衡动态变化的问题，仅用常规的地面预测方法是远远不够的。GIS为解决这一复杂课题提供了强有力的技术手段。世界发达国家如美国、加拿大，发展中国家如巴西、中国和巴基斯坦等建立了国家级水资源信息系统。采用GIS技术可以大大提高水资源预测精度和节省大量的经费。下面对美国加州大学James等人建立的科罗拉多河流水量动态分析预测系统作些简介。

James等人经过对河流域644000km^{2}范围的地表、地下水动态监测资料分析研究后，建立了河水面垂向变化与融雪体积和降雪、水之间的最佳回归数学模型。为了完善模型还收集到36种数据，并将数据变换成栅格数据输入数据库。系统分析软件选用图像分析决策软件（DICS）。根据分析需要在DICS的基础上又开发了SNOWPAC智能性软件。该软件可自动识别云层和在最佳均衡模型基础上逐个计算单个像元的雪水均衡参数（SWE），新开发

的软件也适合于其它河流流域的水均衡分析与计算。

系统的动态分析数据库的数据项（层）如表 8--5 所示。建立最佳回归数学模型的基本思想是通过每个像元点上观测的 SWE 均衡值，来估算整个研究区融化雪水补给水源的体积。为了建立 SWE 与系统数据库中其它各项数据之间的相互关系，将 513 个观测点测得的数据输入 LOTUS 绘制扩展图。除去三个有白云的像元点外，以 510 个观测点测量的均衡数据为基础建立最佳回归数学模型。利用这一模型对系统数据库中所有数据项进行回归分析和误差曲线分析，确定这些数据与地面实测数据的最佳相关性。在对研究区所有水体计算叠加后制成综合信息图像，这些图像可用来预报雪融季节对河流的补给量和绘制丰水、枯水季节水资源综合利用图。

表 8—5　系统数据库的数据项

序号	数据项（层）
1	遥感：可见光，近红外、热红外数据
2	数字高程模型
3	坡度、坡面、坡角数据
4	阴影等照明数据
5	盆地及盆地分级数据
6	地面观测站位置
7	行政分区

四、多元地学信息找矿系统

GIS 技术为地质找矿的综合分析和提取多元地学信息深层内涵，建立综合信息找矿模式，开展理论找矿提供了有力的技术手段。GIS 技术在地质找矿中的应用有以下几个主要优点：数据输入简单，可按地理坐标系建立多项信息数据库；改变了传统的专题地学图件目视分析方法，可按用户的要求用计算机灵活地对系统数据库中的多项数据进行分析和综合；便于实现不同数据结构的相互转换；便于建立仿真三维找矿模型，提高视觉效果；能对前人的地质勘查数据进行二次开发和利用；可以综合遥感、地球物理，地球化学和地面地质信息来解决地质找矿问题。参见图 8—9。应该指出，利用这种找矿信息系统在分析、筛选、组合与成矿有关的特征信息过程中，都必须有经验丰富的地质找矿人员参与，才能取得最佳效果。

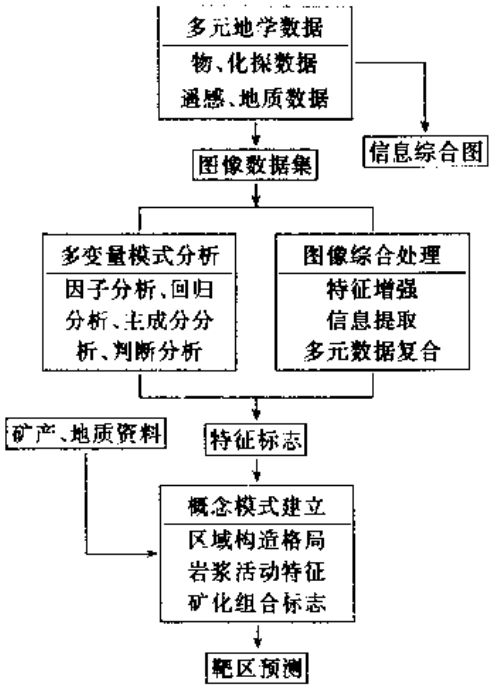


图 8—9　多元找矿信息处理及靶区预测流程图

在建立找矿地学信息系统时应注意下列问题：

（1）建立良好的找矿条件地质模型。不同矿床的成矿条件各不一样，建立地质模型时的多元地学信息目标也不同。

（2）选定合适的比例尺和精度要求。为了达到找矿的精度要求，地球物理、地球化学、地面地质和遥感数据都应有合适的比例尺。

（3）确定合理的遥感波段和图像信息采集环境条件，以利于揭示地质构造和地层结构信息。

（4）研究合理的数据组合模式、算法以及图像处理方法。

（5）采用哪一种数据结构取决于精度要求和原始数据类型。

第 8—8 节　矿区资源与环境信息系统

矿区是一种特殊的地理区域，其地理空间要素、资源环境信息和社会经济信息的内容广泛、综合、复杂、变化迅速。发展我国矿区资源环境信息系统及其应用十分必要，将产生重大的经济效益、环境效益和社会效益。

一、矿区资源开发及其环境问题

矿产资源是自然环境的重要组成部分，各种资源在不同时间和空间条件下按不同的比例和关系联系在一起，形成相互作用、相互联系和相互依存的统一整体。一种资源（不论是再生资源，还是非再生资源）的开发，必然导致与其相联系的其它资源的变化。自然资源之间的这种相互关系的影响，使得在自然资源开发时轻则会造成资源的浪费和破坏，重则会导致环境严重污染，生态恶化，甚至人类文明的衰落。

在实施矿产资源开发工程时，不仅只是涉及矿产这种资源，还不可避免地要波及土地、水、森林（绿被）等自然资源的综合开发和利用问题。与此同时，将产生矿区或工矿城镇的环境生态、甚至灾害问题。在世界能源消费构成中煤炭一直起着举足轻重的作用。与作为主要能源之一的石油相比，1972 年世界煤炭总产量与石油总产量之比为 66/115，1985 年为 115/216，预计 2000 年为 170/195，而到 2020 年将发展为 259/106。在中国，煤炭在总能源中的比例一直是 70%左右，并且在今后相当长的时期内不会有大的变化。在煤炭、冶金、化工、核能等采矿业中，就其规模和影响范围而论又首推煤炭资源的开发和加工、储运。我国有 600 多座大中小城市，而有煤炭资源蕴藏的城市近 80%，其中已有一百余座煤矿城市受到采煤工程的不同影响。

矿产资源开发、加工与利用过程中对其它自然资源和区域环境生态的影响往往相当严重、深远。主要表现为以下几个方面：

1. 对土地资源的影响

开发矿藏要占用和破坏大量的土地，露天开采必须剥离矿体上方的土层和岩层，使原来生长在这些土层表面的生物遭破坏，还将影响矿坑周围动、植物的生长。地下开采将造成地表变形、沉陷，对农田、道路、地面建筑物、地下管线产生危害或破坏。产生的大量固体废弃物，如排土场的土石、矸石、尾矿、煤泥等不仅要占用大量土地，并且对地表下面及附近的土地都有不同程度的污染。据统计，我国因各种人为因素造成的废弃地约 2 亿多亩，其中因采矿、烧砖、烧煤发电等生产活动造成的废弃地约 5000 万亩，并且 70%左右曾是良田沃土。预计，今后仅工矿废弃地将以每年 70 万亩的速度增加。因此，因地制宜地

进行综合治理刻不容缓。

2. 对水资源的影响

矿床开采和加工造成地表水、地下水污染，还可使地下水位降低，矿山周围水资源紧缺，旱灾增多，甚至土地沙化。例如，山西五阳煤矿周围 12 个自然村由于地下开采使水位下降 3～8m，严重影响居民用水。宁夏大武口洗煤厂煤泥池中沉淀后的废水漫流排放，造成约 20km^2 的煤泥沼泽，废水中悬浮物含量高达 10～30g/L，并且含有多种有害物质。

3. 对空气环境的影响

许多煤矿的矸石山和煤堆有自燃现象，有些煤层也有自燃现象，矿石冶炼、油气燃烧和燃煤排放大量烟尘、CO_2、SO_2 和其它有害气体，造成大气污染、温室效应和酸雨现象，从而直接、间接地危害人类健康。

4. 矿山动力地质和环境地质问题

采掘工程活动可能诱发一些动力地质现象，如冲击地压、矿震、岩土工程失稳、泥石流、煤与瓦斯突出、突水等。这些现象的产生既有其特定的地质条件，又与工程活动密切相关。现在，对矿山环境地质的研究日益受到重视。

总之，矿产资源开发和加工、利用，一方面促进了区域经济和社会的发展、进步；另一方面对自然景观、环境生态也有重大影响，将引起当地地形、地貌和景观的变迁，大气、土壤和水体的污染，水文地质和地球化学的变化，甚至局部微气候发生变化。这些影响和变化具有复杂的形成机制，多学科交叉的特性，以及调控、处置极其困难的特点。

二、矿区资源环境信息系统的基本内容和关键技术

人口、资源、环境、灾害是当今世界考虑人类生存与发展的四大基本问题。我国是这四大基本问题最为严重的少数国家之一。实施《中国 21 世纪议程》是我国对全球可持续协调发展事业的贡献。在实施过程中要不同程度地涉及和解决人地关系协调发展及优化调控体系等根本问题。在进行区域（如矿区或工矿城市）的可持续发展研究时，需要将人口、资源、环境与发展作为一个整体，研究它们之间的结构功能，相互作用机理，预测其发展趋势，提出不同类型区域可持续发展的优化模型，拟订调控和管理举措。

在信息时代，人类有可能全面而充分地利用地球资源、保护环境。但只有合理充分掌握信息，才能高效益地使用技术，避免各种不协调现象的发生。地球的任何资源都是有限的，但开发利用的方式却是无限的。信息和信息技术能给我们优化资源利用与保护环境提供依据和手段。

为使矿产资源开发与区域（矿区或矿区城镇）整体建设、环境保护和社会发展统一起来，将区域的整体经济效益、生态环境效益和社会效益统一起来，不仅需要应用现代先进的科技手段和方法解决矿山空间地理及资源信息的采集和加工、处理问题，还应该探索有效的矿区环境及灾害动态监测与分析技术，研究区域环境和灾害的动态变化与矿产资源开发、加工和储运的关系，与区域地质、地形地貌条件的相关关系。国内外和我们的研究实践表明，遥感技术配合常规的测绘和环境监测技术，能够比较有效地解决矿区资源与环境状况的调查、分析评价问题。而地理信息系统（GIS）的理论和技术方法是矿区多层空间（地面、地下和近地表大气层），以及资源和环境等动态四维时空信息的存储、处理、复合、分析与评价的有力武器。在此基础上，开发适用于不同矿区条件的矿区资源环境信息系统，并应用数学、物理和力学等理论进行定性定量的模拟，可望实现矿区资源的优化开发，环

境影响的动态综合评价与信息反馈和调控决策。从而为区域的可持续、协调发展服务。图8—10表示这种模拟、调控系统的原理框图。

（一）应用目标和系统构成

1. 系统的应用目标和任务

就矿业部门的资源环境信息系统或管理信息系统而论，可分为如图8—11所示的三级或四级层次结构。而本文讨论的矿区资源环境信息系统（MREIS）中的“矿区”，主要是指：一是某一个矿山，往往涉及几个平方公里的地面和地下范围；二是矿务局，它辖管由几个矿山组成的一个矿区；其三，也可以是一个矿山城市，因为在我国有不少中小城镇是以开发矿产资源而逐步发展起来的。这样，具体开发、应用的MREIS可以是针对上述三者之一的。

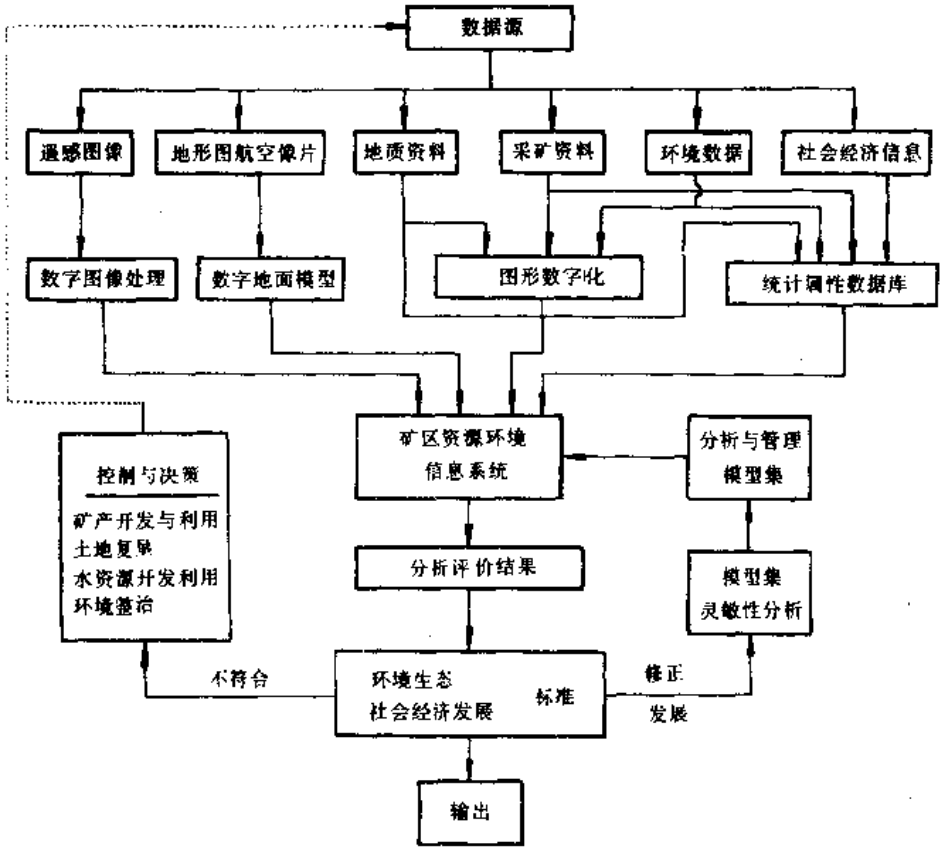


图8-10 矿山工程与区域环境影响的空间模拟、调控系统

矿区资源环境信息系统是一个复杂的开放的综合的动态空间信息系统。它具有与城市规划管理信息系统、土地信息系统等其它信息系统显著不同的特点。主要表现为

1）综合性、系统性 该系统要能够反映、处理从矿区的勘探设计、规划建设、生产经营到矿区衰老很长时间周期中的问题，要处理以矿山生产建设为主体的包括技术、经济、安全、环境和社会等多个组成部分，多级层次结构和多个应用目标。

2）三维空间特性 矿区空间是包括地面、地下和大气层的多层立体空间，它具有复杂

的内部结构，如复杂地形地貌中的悬崖峭壁，矿区地表的变形和塌陷，矿床中的褶皱、断层构造，呈不均匀分布的矿物成分，地球物理和地球化学信息以及采场，巷道周围的地应力场，大气的污染程度和污染源等。

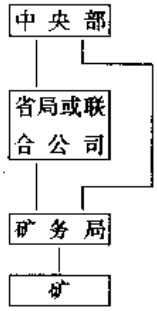


图 8—11 矿业管理信息系统的层次结构

3）动态性 它有两层含义：一是矿区开发和生产作业的地理空间地点时刻在变动中，地下巷道、采场和矿床存储状况日新月异，由此又将引起矿区生态环境、景观的不断变化；二是矿山开发的技术水平不断进步，对技术经济效益和环境效益的评价能力、标准也因时因地而不同。它们都影响着信息系统的设计和内涵。

4）差异性 各矿区、矿山的矿产资源地质条件、地形地貌形态、生产技术水平、社会经济情况以及人员素质、管理水平等存在着不同程度的差异性。

5）不确定性和随机性 由于地下矿藏赋存状况和地质构造的复杂性、不稳定性，并且受勘探技术手段和勘探工作量的限制，因此对矿体及其围岩地质特征的描述往往带有一定的推断性质，对开采作业对象不能完全确知，再加上井下作业环境较差，巷道、设备发生故障的可能性较大，诱发因素也较多，因而生产作业和管理工作完全按原定计划执行很难做到，而有一定的随机性。

6）后效性和滞后性 矿山建设工程量大、建设周期长，总体的效果滞后于生产，生产滞后于基建，基建滞后于勘察和设计。矿区开发、生产造成生态环境影响的后效性、滞后性也很明显。

7）流逝性 矿产资源是不可再生的自然资源，使矿产资源开采的生（产）—灭（报废）过程贯穿于矿区开发的整个时期。此外，在矿区资源储量一定的条件下，年产量与矿山寿命，近期效益与远期发展之间存在着相关关系和矛盾性。

在进行 MREIS 的可行性研究、系统设计和实施过程中要充分考虑以上特点或特性。

MREIS 的基本应用目标和任务可概括为如下内容：

（1）完成矿区或矿山城市的开发规划、生产管理和分析决策中带全局性、涉及多学科、多专业的任务。如矿山、井田的划分，矿山整体建设（生产、生活、环境及社会经济）规划设计等。

（2）完成矿区或矿山规划设计、生产管理和分析决策中的一些单项或专题任务。此时，往往只需要利用信息系统中的某一、二个子系统（或分系统）的功能。例如，矿山开采优化设计，服务于矿山生产现代化管理的矿床（煤层）地质信息建模，矿区测绘规划与管理，矿区水资源开发与管理等。

（3）完成矿区某种工程或问题的统计、规划、分析及评价。例如，矿床（体）开采技术经济合理性分析评价，开采沉陷预计与整治规划，矿区环境质量评价与治理等。

以上目标和任务概括地列于表 8—6 中。

2. 系统的构成

MREIS 的目标和任务众多，彼此关连性强，很难表述完整。因而系统的构成和内容也十分广泛和复杂，具有不同层次和交叉性质，需要多专业、多方面科技和管理人员通力合

表 8—6 矿区资源环境信息系统的基本应用目标、任务和要求

任务与要求 / 应用目标	基本任务	数据来源和基本技术要求
规划设计	1. 矿区开发规划与设计 2. 矿井设计与采区布置 3. 多种经营发展规划 4. 区域经济和社会发展规划 5. 各种工程设计 6. 环境保护规划 7. 测绘工程规划设计 ……	1. 矿区地形图，DTM 及测量资料，遥感资料 2. 矿区土地利用图 3. 矿区地质图件及文字报告 4. 矿床（体）产状图及文字报告 5. 采掘工程图 6. 井上下对照图 7. 市政工程图 8. 矿区开发设计报告 9. 环境监测资料 10. 技术经济与社会统计资料 11. 其它人文资料 …… 基本图件比例尺： 1∶5000～1∶50000，1∶2000 各种数据的质量、精度指标
生产经营管理	1. 生产调度与管理 2. 矿井通风安全计算和管理 3. 矿区测量与土地信息管理 4. 矿产储量管理 5. 水资源开发与管理 6. 环境监测与质量评价 7. 物资管理与矿产经销 ……	
分析评价决策	1. 开采技术经济合理性分析评价 2. 开采沉陷预计与防治 3. 环境生态调控决策 4. 工程选址，布置决策 5. 生产经营效益分析 6. 灾害防治预测、评价 ……	

作才能逐步建成，取得较满意成效。应该指出，拥有完整、准确、及时的数据，是任何信息系统能否完成各项任务，达到预期目标的基础。并且也是系统建设过程中需要付出极大人力、物力和财力的重要部分。否则，系统将成为无源之水，无本之木。其中测绘数据、资料，地质和矿产资源数据、资料，以及采矿生产数据资料则是 MREIS 的基础信息。MREIS 的系统构成如图 8—12 所示。由于矿山有基建阶段、生产阶段和衰老阶段的不同或侧重期，因此图 8—12 所示的子系统或模块数量和内容可以有所不同或侧重，并且也是逐步发展、建设起来的。不难理解，MREIS 及其中的许多子系统都是空间 GIS。

（二）系统的主要理论问题、关键技术和方法

在发展我国的地理信息系统（GIS）和地理信息产业中需要解决许多共同的理论问题、关键技术和方法。这里只论述发展具有矿业特色的 MREIS 中的重要问题。

1. 系统设计理论和方法

MREIS 是一个复杂的动态空间信息系统，是由一些相互区别、相互作用和影响的子系统有机地联结在一起，为某些共同目标（主要是生产矿产产品）而又具有各自功能的综合信息系统。在进行系统设计和建设时宜用系统动态仿真模型，对矿区开发、生产与自然资源、地理环境的内部结构和变量之间的因果关系，社会行为和过程进行仿真模拟，完成矿

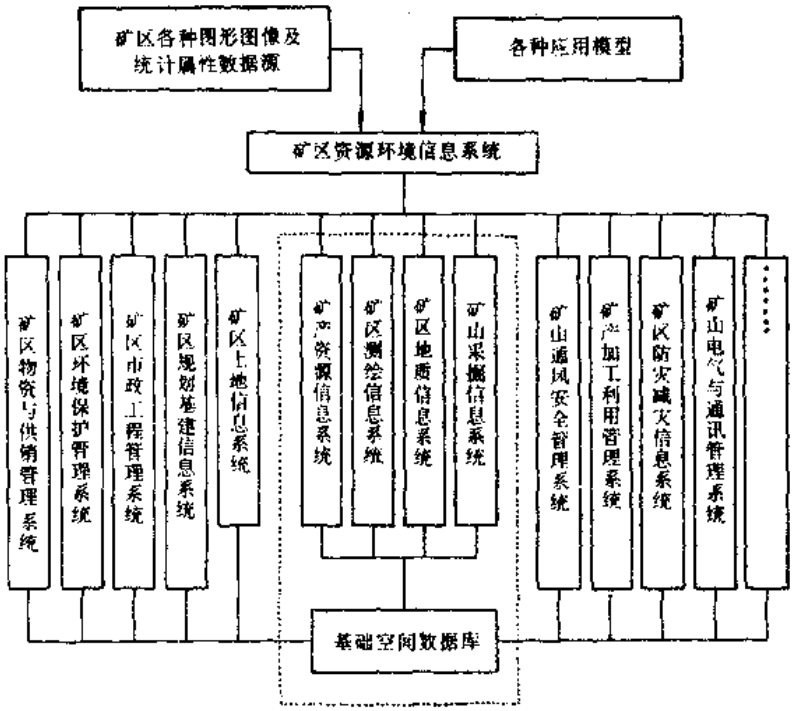


图 8—12　矿区资源环境信息系统的构成

区各种规划设计、生产管理、分析评价以及发展趋势预测，反馈和调控决策。

2. 数据的分类体系和标准化

信息时代的一个特征是实现信息或数据在社会中的通讯或共享。为实现 GIS 系统数据交换，要解决系统数据的分类、编码及格式的标准化问题。矿区地理信息系统数据分类、编码和格式体系的研究，应在全国统一规范的指导下进行，既要与全国接轨，又要与兄弟部门协调，立足现在、面向未来。

MREIS 的数据内容十分广泛、类型复杂，一般需要 6～7 级分类才能实用。表 8—7 给出了 MREIS 的数据高位分类体系（建议稿）。更详细（第 4～7 级）的分类由各子系统的研究开发人员制定、充填，从而完成一个完整的分类体系。

3. 数据采集及数据质量

输入 MREIS 中的各种类型数据的采集技术和处理方法，涉及许多学科，如测绘、地质、采矿、机电、环保、财经等。尽管对这些数据的采集和加工处理并非从事 MREIS 研究和应用人员本身的职责，但对它们的质量和误差大小则应该有确切的了解。因此，从事信息系统研究、开发和应用的人员应该了解输入并存储在 MREIS 中数据的质量。

MREIS 中处理的各种实体对象和目标不但表现为严格或不严格的确定性，而且普遍存在着某种不确定性（uncertainty）。例如，矿产的赋存状态、储量，断层的产状和大小，矿体品位高低，矿体围岩的岩性，巷道和采场周围的应力状态，开采沉陷规律，环境污染程度及其变化等等。对它们的衡量或评定不仅需要应用传统的数理知识，还要应用模糊浑沌理论、分形（fractal）理论、突变论、小波（wavelet）理论等新概念、新理论。由于研究实体本身的复杂性和特性，以及人们对客观世界认识的不足和局限性，人类自身条件（如人眼的分辨能力）的限制，技术水平和仪器设备的限制等，因此对所获取数据质量的衡量

表 8—7 矿区资源环境信息系统的数据高位分类体系（建议稿）

（以煤炭资源开发为主）

第一级	第二级	第三级
1 基础数据	11 区域境界	1101 行政区划
		1102 矿区区划
	12 交通运输	1201 铁路
		1202 公路
		1203 水路
		1204 管路
		1205 航空
		1206 海路
	13 地形地貌	1301 地形
		1302 地貌
		1303 遥感图像
	14 水土资源	1401 河流
		1402 湖泊
		1403 地下水
		1404 土地利用
		1405 土地权属
		1406 土壤类型
	15 矿产资源	1501 煤炭
		1502 油气
		1503 黑色金属
		1504 有色金属
		1505 贵稀金属
		1506 建材原料
		1507 化工原料
		1508 其它矿产
	16 地质基础	1601 基础地质
		1602 水文地质
		1603 工程地质
	17 气候和生物资源	1701 气温
		1702 降水
		1703 风力
		1704 森林
		1705 草场
		1706 野生植物
		1707 野生动物
2 矿山工程	21 矿山建设	2101 矿区开发
		2102 矿山设计
		2103 矿井建设
		2104 井巷工程
		2105 掘进

续表

第一级	第二级	第三级
2 矿山工程	21 矿山建设	2106 爆破
		2107 支护
		2108 露天矿建设
	22 矿山开采（煤矿开采）	2201 井田开拓
		2202 采区
		2203 采矿（采煤）方法
		2204 矿山压力
		2205 岩层控制
		2206 特殊开采
		2207 水力采煤
		2208 煤炭地下气化
		2209 露天开采
		2210 边坡
	23 地质与勘探	2301 成煤作用
		2302 煤岩
		2303 煤层（矿床、矿体）
		2304 含煤岩系
		2305 煤田（矿田）
		2306 勘探
		2307 储量
		2308 矿井地质
		2309 矿井水文地质
		2310 矿藏物探（煤田物探）
		2311 矿藏钻探（煤田钻探）
	24 矿山测量	2401 矿区控制测量
		2402 矿山工程测量
		2403 建井测量
		2404 矿井测量
		2405 开采沉陷
		2406 土地复垦
		2407 矿体几何
		2408 矿产资源保护
		2409 露天矿测量
	25 通风与安全	2501 矿山通风
		2502 瓦斯
		2503 粉尘
		2504 矿山灾害
	26 煤炭加工利用	2601 选煤
		2602 水煤浆
		2603 煤炭成型
		2604 煤化学

续表

第一级	第二级	第三级
2 矿山工程	26 煤炭加工利用	2605 煤炭分析
		2606 煤转化
	27 矿山机械	2701 采煤机械
		2702 掘进机械
		2703 钻孔机械
		2704 装载机械
		2705 矿山运输
		2706 矿井提升
		2707 液压支架
		2708 选煤机械
	28 矿山电气与通讯	2801 矿山供电
		2802 矿山电气安全
		2803 矿井监控
		2804 矿山通讯
3 矿山经营与管理	31 计划与调度	3101 基建计划
		3102 生产计划
		3103 资金计划
		3104 生产调度
		3105 经营调度
		3106 生产统计
	32 物资供应	3201 采掘设备
		3202 成套设备
		3203 配件
		3204 材料
		3205 供应统计
	33 产品运销	3301 产品销售
		3302 产品运输
		3303 产品质量
	34 财务与劳资	3401 成本核算
		3402 经济分析
		3403 财务劳资统计
		3404 工资
	35 人事	3501 人事统计
		3502 干部管理
		3503 劳动组织
4 社会经济	41 区域人口	4101 人口概况
		4102 劳动力
		4103 人口素质
	42 工农业经济	4201 综合经济
		4202 矿业经济
		4203 建筑业经济
		4204 化工经济

续表

第一级	第二级	第三级
4 社会经济	42 工农业经济	4205 其它工业经济
		4206 农业经济
		4207 乡镇企业
	43 商业服务业	4301 商业
		4302 服务业
		4303 饮食业
	44 运输邮电业	4401 客运
		4402 货运
		4403 邮电通讯
		4404 广播电视
	45 文教卫生	4501 教育
		4502 科研
		4503 文化娱乐
		4504 医疗卫生
5 环境与灾害	51 矿区环境	5101 矿山环境
		5102 地质环境
		5103 矿山地下环境
	52 环境污染	5201 污染源
		5202 污染监测
		5203 污染影响
	53 环境恶化与灾害	5301 地面塌陷
		5302 生态破坏
		5303 环境灾害
	54 环境保护	5401 环境规划
		5402 环境治理
		5403 生态恢复
6 区外信息	61 物资交流	6101 物资输出
		6102 物资输入
	62 能源交流	6201 能源输出
		6202 能源输入
	63 信息交流	6301 信息输出
		6302 信息输入
	64 市场信息	6401 国内市场
		6402 国际市场
	65 参考信息	6501 经济信息
		6502 科技信息
		6503 管理信息

指标不能象传统测量数据那样，只是精度或误差这一项，而需要采用多种指标，如利用七种指标：数据渊源（Source 或 lineage）、精度（包括位置精度与属性精度）、逻辑一致性、完整性（completeness）、可靠性（reliability）、现势性（Currecy）和安全等级（Security classification）。这样，作为 GIS 技术研究的一个重要领域，还应结合 MREIS 的具体情况，研究系统数据库中各类数据质量和精度的确定方法；系统中所用各种模型对数据误差或偏差的传

播规律，并在模型中设置误差控制机制；以及根据用户对输出结果的要求，帮助用户确定各种输入数据的误差限，实现系统的反馈调控。

4. 数据模型及数据结构

由上述不难理解，MREIS 需要能够存储、处理基于图形/图像和属性数据一体化的复杂或复合的空间实体数据，并且这些数据往往具有不同分辨率、不同精度、不同时间，还要求能够支持多媒体数据处理，如超长文本数据、图形图像数据、声音数据等，以及实现有效的数据查询和更新。这就要求有真三维的空间数据模型和数据结构，以及动态空间综合分析能力。

80 年代以后发展起来的面向对象技术已在数据库和软件工程方面得到广泛应用。近些年来也日益受到 GIS 研究人员的关注。初步研究成果表明，以面向对象模型方法（OOM）和面向对象程度设计（OOP）为基础的面向对象数据模型和数据结构，能够较好地模拟和处理现实世界中的复杂对象，克服传统 GIS 数据模型的局限性和缺陷，改善系统的性能。因此，设计和开发具有语义层次的面向对象的数据模型，已成为当前 GIS 研究的关键技术之一。然而，它能否有效地处理和分析 GIS 和 MREIS 中各种复杂类型的空间数据，甚至支持实时、动态和智能型的空间分析与决策，还有待实践的考验。

此外，以超图数据结构（Hyper graph－Based Data Structure，HBDS）为基础的超图理论(hypermaps)将成为设计动态空间数据库系统的理论基础，并且已在国内外的一些 GIS 软件中得到应用。以 Voronoi 图形和 Delaunay 三角网为基本图形要素的 Voronoi 数据结构具备一些特性，在建造 IAS（图像分析系统）/GIS 一体化的空间数据模型方面也不容忽视。体元结构（Volume Cell Structure 或 Cell Structure）是二维栅格结构在三维空间中的扩展，并在澳大利亚 MIMCOM 公司的 MINESCAP 软件，美国施伦贝克公司的 GEOQUEST 软件中得到应用。还有，近年来计算机三维图形显示技术中的三维实体造型技术也得到迅速发展，值得予以关注和研究。

总之，服务于千差万别、多彩多姿现实世界的 GIS，拥有各具特色、百花争艳的数据模型和数据库系统是十分自然、适应众多用户需求的事情，我们应为此而努力。

三、应用模型研究及举例

要使 MREIS 在解决矿区的各种实际问题中发挥效用，必须要有实用的应用（数学）模型及应用软件。应用模型及应用软件的研究、开发不仅是 MREIS 的重要内容之一，而且也是该系统解决实际问题的能力、正确性、效率和最终效益的关键所在。应用模型的建立必须以广泛、深入的专业研究为基础，具有数学和有关科技知识，并对所使用的空间信息系统的硬、软件系统有较深刻的了解。

MREIS 能够有效地处理矿区资源综合开发和利用，矿区优化规划，工程项目的决策或选址，采空区上方岩层及地表沉陷规律的预测和防治，采矿和矿产加工引起生态环境问题的分析与评价及其调控决策，采矿生产技术经济综合分析，各种信息的检索和统计分析，各种复杂长度、面积或体积的快速量算等等涉及空间数据和非空间数据的诸多问题。现举例如下：

1. 开采沉陷预计与分析的空间模拟

在采空区上覆岩层及地表沉陷中任一点的移动向量可用下式预计：

$$\vec{U}_1(x, y, 0) = \vec{u}_1(x, y, 0) + \vec{V}_1(x, y, 0) + \vec{W}_1(x, y, 0) \quad (8—16)$$

式中，$\vec{u}_1(x, y, 0)$ 为水平移动分量，$\vec{V}_1(x, y, 0)$ 为沿倾向的水平移动分量，$\vec{W}_1(x, y, 0)$ 为下沉分量。并且

$$\left.\begin{aligned}
u_1(x,y,0)=&\frac{-4M}{\pi^2}\int_0^{\infty}\int\frac{h_i\mathrm{sh}\rho h_i(\rho h_{i+1}\mathrm{ch}\rho h_{i+1}+\mathrm{sh}\rho h_{i+1})\sin\xi D_1\sin\eta D_2}{\{(\mathrm{sh}\rho h_i\mathrm{ch}\rho h_i+\rho h_i)(\mathrm{sh}\rho h_{i+1}\mathrm{ch}\rho h_{i+1}+\rho h_{i+1})}\\
&\frac{\sin\xi x\cos\eta y d\xi d\eta}{+\frac{G(1-\gamma_{i+1})}{G_{i+1}(1-\gamma_i^2)}(\mathrm{sh}^2\rho h_i-\rho^2\eta^2)\mathrm{sh}^2\rho h_{i+1}\}\eta}\\
V_1(x,y,0)=&\frac{-4M}{\pi^2}\int_0^{\infty}\int\frac{h_i\mathrm{sh}\rho h_i(\mathrm{sh}\rho h_{i+1}+\rho h_{i+1}\mathrm{ch}\rho h_{i+1})}{\{(\mathrm{sh}\rho h_i\mathrm{ch}\rho h_i+\rho h_i)(\mathrm{sh}\rho h_{i+1}\mathrm{ch}\rho h_{i+1}+\rho h_{i+1})}\\
&\frac{\sin\xi D_1\sin\eta D_1\cos\xi x\sin\eta y d\xi d\eta}{+\frac{G_i(1-\gamma_{i+1})}{G_{i+1}(1-\gamma_i)}[\mathrm{sh}^2\rho h_i-(\rho h_i)^2]\mathrm{sh}^2\rho h_{i+1}\}\xi}\\
W_1(x,y,0)=&\frac{4M}{\pi^2}\int_0^{\infty}\int\frac{(\mathrm{sh}\rho h_i+\rho h_i\mathrm{ch}\rho h_i)(\mathrm{sh}\rho h_{i+1}+\rho h_{i+1}\mathrm{ch}\rho h_{i+1})}{\{(\mathrm{sh}\rho h_i\mathrm{ch}\rho h_i+\rho h_i)(\mathrm{sh}\rho h_{i+1}\mathrm{ch}\rho h_{i+1}+\rho h_{i+1})}\\
&\frac{\sin\xi D_1\sin\eta D_2\cos\xi x\cos\eta y d\xi d\eta}{+\frac{G_i(1-\gamma_{i+1})}{G_{i+1}(1-\gamma_i)}[\mathrm{sh}^2\rho h_i-(\rho h_i)^2]\mathrm{sh}^2\rho h_{i+1}\}\xi\eta}
\end{aligned}\right\}\quad(8\text{—}17)$$

式中 h_i ——第 i 层岩层厚度；

h_{i+1}——第 i 层的下一层层厚；

M ——开采厚度；

D_1 ——矩形工作面之走向长度；

D_2 ——工作面倾向长；

G_i, γ_i ——覆岩第 i 层的剪切模量和泊松比；

G_{i+1}, γ_{i+1}——第 i 层的下一层之剪切模量和泊松比；

ξ, η ——积分变量，且 $\rho=\sqrt{\xi^2+\eta^2}$；

x, y ——开采工作面有效角点坐标。

应用这些或其他数学模型，针对具体矿山或矿区的地质、地形和采矿工程条件，能够利用 MREIS 的功能开发开采沉陷预计与分析的应用程序（软件）。

此外，还开发了有一定智能化的开采沉陷分析软件，能实现多种空间数据的输入、叠加、分析和评价。该系统的原理框图如图 8—13 所示。用户接口及数字化、数据输入子系统的功能显而易懂。模型化子系统是整个系统的关键部分，它包括模型库和驱动器两部分，开采沉陷预计模型也可存储在其中。该子系统是对知识库专家系统的模拟，其模型库类似于知识库，而驱动器是一种配套推理机。模型驱动器通过用户接口接受命令，以选择、启动适当的模型。被启动的模型可从数据库中检索所需的信息，并将模型化的成果再存储起来。将驱动器与模型库分开是便于增加新模型或修改旧模型而又不影响系统的其它部分。数据/成果表示子系统的功能是将模型化处理得到的中间成果或最终成果与系统用户通讯。

2. 矿山开采对环境影响的数学和时空模拟

矿产资源的开发和加工对区域环境有重大影响，将引起地形、地貌和景观的变迁，大气、土壤和水体的污染，水文地质和地球化学的变化，甚至局部微气候发生变化。

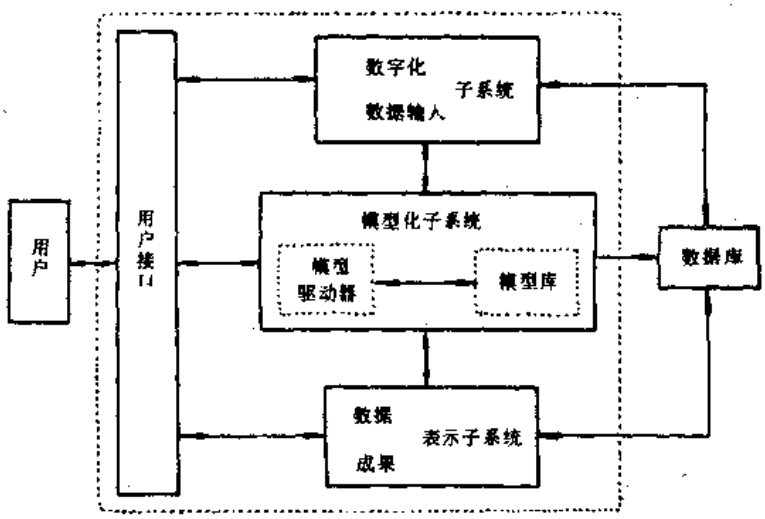


图 8—13　开采沉陷分析系统原理

为了预计、分析评价采矿工程引起的矿区环境状态和变化，应该针对不同矿（区）的地质、地形和采矿条件等的具体情况，研究各种子系统。目前，较成熟的子系统有大气污染、土地破坏和微气候变化等子系统。而象地下水体污染、水文地质和地球化学影响、变化等子系统，由于影响因素多、影响物质的迁运规律复杂，还很不成熟。

根据流体热力学的雷诺方程，可导出矿山工程对环境影响的数学模型的数字解。

$$\left.\begin{aligned}
&\frac{\partial u_i}{\partial t}+u_j\frac{\partial u_i}{\partial X_j}=-\frac{1}{\rho}\frac{\partial p}{\partial X_i}+\frac{\partial(-u'_iu'_j)}{\partial X_j}-\beta_T q\delta_{i3}+lu_2\delta_{i1}-lu_1\delta_{i2}+J_i\\
&\frac{\partial T}{\partial t}+u_i\frac{\partial T}{\partial X_j}=\frac{L}{C_p}\Phi+\frac{\partial V_{Tj}}{\partial X_j}\cdot\frac{\partial T}{\partial X_j}+Q_T+J_T\\
&\frac{\partial q}{\partial t}+u_j\frac{\partial q}{\partial X_j}=\frac{\partial V_{qj}}{\partial X_j}\cdot\frac{\partial q}{\partial X_j}-\Phi+J_q
\end{aligned}\right\}\qquad(8—18)$$

式中：t 为时间，u_i 表示坐标系统中单位质量流体的变化率，ρ 为密度，p 为压力，T 为相对于背景温度的温度差异值，q 为温度，β_T 为体积扩展的热系数，L 为扩散（蒸发）时产生的潜热，Φ 为凝聚（压缩）率，l 为互补系数，J_i、J_T、J_q 分别表示脉动、热和温度（水分）的湍流传输源，V_i、V_T、T_q 分别表示脉动、热和温度（水分）的湍流传输速率，Q_T 表示矿山工程产生的热源。

此外，湍流随机速度的散度等于 0，即 $\partial u_i/\partial X_i=0$，$-\overline{u_iu_j}$是一个二阶张量，且$-\overline{\rho u_iu_j}$通称为雷诺应力，表示湍流脉动引起的单位面积上的动量传输率，并且有如下关系：

$$-\overline{u_iu_j}=V\left(\frac{\partial u_i}{\partial X_j}+\frac{\partial u_j}{\partial X_i}\right)-\frac{1}{3}\overline{u_ku_k}\cdot\delta_{ij},\quad(i,\ j,\ k=1,\ 2,\ 3)\qquad(8—19)$$

通过土壤和岩层进行热传导的方程可表为

$$\frac{\partial T_s}{\partial t}=\frac{\partial k\ (X)}{\partial X_j}\cdot\frac{\partial T_s}{\partial X_j}+Q_T\qquad(8—20)$$

式中：T_s 为岩体的温度，K 为不同岩层的热传导系数。

危险元素在环境中辐射或迁移的规律可用如下数学模型推算：

$$\frac{\partial C_k}{\partial t}+u_i\frac{\partial C_k}{\partial X_i}-W_q\frac{\partial C_k}{\partial X_j}=\frac{\partial D_{kj}}{\partial X_j}\frac{\partial C_k}{\partial X_j}+W_k+J_k \qquad (8—21)$$

式中：C_k 表示第 k 种污染元素的浓度，$k=1\cdots n$（n 是污染元素的种类数）；D_{kj}为污染扩散（渗漏）率；W_q 为污染物沉漏系数；W_k 为化学反应产生的其它成分；J_k 是由于矿山工程产生的污染源。

湍流交换可利用半经验相关理论或封闭网络素流关系。边界条件取决于应用目标的具体条件。

在岩体（层）或土壤的交界处，可以平均湿度和污染物的平衡状态作为边界条件。数字高程模型（DEM）往往用来表示矿区地形或岩层的几何形态。有限差分法可用来近似表示“空气-岩体”结合系统。经上述数学模型处理后的数字输出形式可以多种多样，如等值线、三维曲面、矢量场，可根据需要选取。

这些不同的数学模型可用于评价：井下和地面采矿作业的通风及微气候状况；岩体温度或辐射元素与周围环境之间的交互作用；开采时环境条件的预计；按环境保护要求对矿山进行优化设计。

附录　世界主要GIS商品软件特征一览表

（据姚艳敏：《遥感信息》，1993～1994）

系统名称	软件类型									计算机										数据结构					
										操作系统					机型										
	GIS	设备管理	自动制图	DBMS	数字化	格式转换	遥感	CAD	其它	Unix	DOS	Macintosh	VMS	其它	PC及兼容器	Macintosh	工作站	微型主机	其它	栅格式	拓扑矢量式	非拓扑矢量式	TIN	3-D三维显示	其它
AGIS	×		×	×	×		×			×	×				×		×			×	×				
ALF					×					×							×					×			
ALICE	×	×		×				×		×	×				×		×			×	×		×		
ALK/ATKIS	×	×	×	×	×	×				×			×				×	×		×	×				
ALMAP，Microij					×				×	×	×	×	×	×	×	×	×	×	×						
ALPER	×	×	×	×	×	×		×		×							×			×	×				
ARC/INFO	×	×	×	×	×	×	×	×	×	×	×		×	×	×		×	×		×	×	×	×	×	×
ARGO	×	×	×	×	×	×		×		×	×	×			×	×	×				×	×			
ARIES	×						×			×			×		×		×	×		×					
ASAR3	×									×							×			×					
Atlas * GIS	×										×				×							×			
AUSTAD MAP	×	×	×	×	×	×				×	×				×		×	×		×	×	×		×	
Auto CAD								×		×	×	×	×	×	×	×	×					×		×	
AIIMUTH	×	×	×					×				×				×									
Byers Laptop Viewer	×	×	×					×			×				×							×			
Byers Plot Station									×		×			×	×					×	×	×		×	
Cable Cad	×	×	×	×	×			×		×	×			×	×		×			×	×	×			
CADCore/Tracer	×	×	×	×	×	×		×			×				×					×		×			
CAD Dverlay GS	×				×	×				×	×									×					
CAM-1＋	×										×				×										×
Carto PC							×				×				×					×					
Cartographer	×											×				×					×				
CEAI		×	×		×	×		×		×	×		×	×	×		×	×			×	×	×		
Census Windows：TIGER Tools	×		×								×			×	×						×				
CHIEF	×			×					×				×												
CHIPS	×						×				×				×					×					

续表

系统名称	软件类型									计算机										数据结构					
										操作系统					机型										
	GIS	设备管理	自动制图	DBMS	数字化	格式转换	遥感	CAD	其它	Unix	DOS	Macintosh	VMS	其它	PC及兼容器	Macintosh	工作站	微型主机	其它	栅格式	拓扑矢量式	非拓扑矢量式	TIN	3-D三维显示	其它
COMPASS									×		×				×										
CRIES-GIS	×			×	×	×					×				×					×	×			×	
Customized	×	×	×	×	×	×	×	×	×	×	×	×	×		×	×	×	×	×	×	×	×	×	×	
DanGraf	×	×	×		×	×					×			×							×				
DATANET Plus Mapping	×										×				×										
DataPix	×	×	×	×						×	×		×	×	×		×	×							
Delorme' s XMap	×	×	×	×	×	×				×	×				×		×			×	×	×		×	×
Delphi/GIS	×										×			×	×			×	×			×			
Desktop Mapping System	×		×				×				×					×					×	×	×		×
DGS-2000								×						×											
Digit-Aps	×		×	×	×	×				×	×		×		×		×	×		×	×	×	×	×	
DRAGON	×				×	×	×				×				×					×		×			
Earth One	×	×	×	×	×	×	×				×			×	×					×			×	×	
EARTH/INFO	×	×		×				×		×	×	×			×	×					×				
EASI/PACE	×						×			×	×	×	×		×	×	×	×		×		×		×	
EIS-Emergency Info System									×		×				×					×					
Electric Integrated	×	×	×			×								×			×	×			×	×			
Facilities Mgmt																									
Eng. Work Order System		×	×	×	×			×						×	×					×		×			
EPPL7	×										×				×					×		×			
ERDAS	×	×	×		×	×	×		×	×	×		×		×		×	×		×		×		×	×
Facet	×					×	×			×							×			×	×	×	×	×	×
FINNCAP	×	×		×	×	×							×				×	×			×			×	
FMS/AC	×	×	×	×	×			×		×	×				×		×				×	×			
GADS	×		×								×				×										×
Gas Integrated Facilities	×	×	×			×								×			×	×			×	×			
Mgmt																									
GENAMAP	×	×	×	×	×	×		×	×	×					×		×		×	×	×	×	×	×	×
Goo-Graphics	×	×	×	×	×	×				×	×				×						×				
Geo Axses	×			×		×					×				×					×	×	×			
GEODIS, INFOGIS,	×	×	×	×	×	×		×		×				×			×		×	×	×	×	×		
GEOMIP																									

续表

系统名称	软件类型										计算机										数据结构					
											操作系统					机型										
	GIS	设备管理	自动制图	DBMS	数字化	格式转换	通感	CAD	其它	Unix	DOS	Macintosh	VMS	其它	PC及兼容器	Macintosh	工作站	微型主机	其它	栅格式	拓扑矢量式	非拓扑矢量式	TIN	3-D三维显示	其它
Geo District									×		×				×							×				
Geog. Road Network	×		×	×					×					×					×							×
Database																										
GeoLink System	×	×	×	×	×		×		×	×	×				×		×			×	×	×	×	×		
Geo Query									×			×				×										×
GEOROUTE			×						×	×	×		×	×							×					
GeoSight	×										×				×						×					
GeoStation	×	×	×	×	×	×		×		×			×		×		×	×		×		×		×		
Geo Vision' s GIS	×	×	×	×	×	×				×			×				×	×		×	×	×	×	×	×	×
GIS-chains	×			×	×	×	×				×			×	×					×	×	×				
GIS Manager	×	×	×		×	×		×	×	×	×				×		×				×		×	×		
GIS Plus	×										×				×						×					
GIS-200	×		×		×	×	×						×					×		×		×				
GIST	×			×	×	×	×		×		×		×		×			×		×	×	×	×	×	×	×
GRADIS	×	×			×					×			×				×			×	×		×	×		
Graphics Design System	×	×	×	×	×	×		×	×				×	×			×			×	×	×	×	×		
GRASS	×				×	×	×			×					×	×	×	×		×	×			×	×	
GWN—GIS	×	×	×		×			×			×				×						×	×	×	×		
HI—VIEW	×					×	×		×		×				×					×						
IBERGIS		×	×		×								×				×				×					
IBM A IX geo GPG/6000	×			×										×			×			×						
IDIMS Product Family	×						×			×			×				×	×		×	×			×		
IDRISI	×						×		×		×				×					×		×				
IIS—Map	×													×				×				×				
IM²	×	×	×	×	×	×		×			×				×					×	×	×		×		
Imagis (Based on Auto CAD)	×							×							×						×	×		×	×	
INFOCAD	×	×	×	×	×			×	×	×							×	×		×	×		×	×		
INFOCAM	×	×	×	×	×	×			×				×				×		×		×	×		×		
Infomark for Windows			×								×			×							×					
Informap	×	×	×	×	×	×				×			×		×	×	×	×		×	×	×				
Interactive Surface Modeling			×		×				×	×			×				×	×				×		×		

续表

系统名称	软件类型									计算机										数据结构					
										操作系统					机型										
	GIS	设备管理	自动制图	DBMS	数字化	格式转换	遥感	CAD	其它	Unix	DOS	Macintosh	VMS	其它	PC及兼容器	Macintosh	工作站	微型主机	其它	栅格式	拓扑矢量式	非拓扑矢量式	TIN	3-D三维显示	其它
KARTO	×	×	×	×	×			×		×	×				×		×					×			
Kork Digital Mapping System			×		×	×			×		×		×	×	×							×		×	
Land Management System	×	×	×			×								×			×	×			×	×			
LandBase									×	×	×	×			×	×	×			×	×	×			
LISSY	×		×		×	×				×							×				×			×	
MacGIS	×											×				×				×				×	
Magellan	×			×	×	×	×				×		×		×			×		×		×		×	
Manatron Land Info System	×			×	×	×			×		×			×	×		×					×			
Map Designer	×		×	×	×	×								×			×			×		×			
Map Access									×	×	×				×		×	×			×				
Map Box	×									×	×				×		×			×				×	
Map Grafix	×	×	×	×	×	×						×		×		×						×			
Map Info	×	×	×	×						×	×	×		×	×	×	×								×
MAPIT			×								×				×						×				
Map Measure			×		×	×					×				×					×	×	×		×	
Map Power DeskTop GIS	×								×			×										×			
Map Station	×			×					×	×	×				×		×			×	×	×			
MARS	×										×			×	×						×				
Meta Map	×	×	×	×	×	×					×				×						×	×			
Micro Image Product Family	×					×	×				×				×					×		×		×	
MIMS	×	×	×	×	×						×				×										×
MIPS	×	×	×	×	×	×	×	×	×		×				×					×	×	×		×	×
Modular GIS Environment	×		×		×	×		×		×	×				×		×			×	×	×	×	×	×
Mota Site					×			×		×							×						×		
Mountain Top CAD & Info mgmt		×			×	×		×	×	×							×					×	×	×	×
Multiple: Systems Integrator	×	×	×	×	×	×	×	×		×	×	×	×		×	×	×	×	×	×	×	×	×	×	
Mun Map GEO/SQL	×	×	×					×		×	×				×		×				×		×	×	
Nucor	×	×									×				×					×	×				×
ORTHOMAP/H1—F1			×							×					×		×			×	×		×	×	
Pafec—GIS	×	×	×	×	×	×		×		×			×	×			×	×		×	×	×			
PFINDER, PFBasic, CBLOG					×	×			×		×				×										

续表

系统名称	软件类型									计算机										数据结构					
										操作系统					机型										
	GIS	设备管理	自动制图	DBMS	数字化	格式转换	遥感	CAD	其它	Unix	DOS	Macintosh	VMS	其它	PC及兼容器	Macintosh	工作站	微型主机	其它	栅格式	拓扑矢量式	非拓扑矢量式	TIN	3-D三维显示	其它
PICDMS	×			×			×							×	×		×	×		×	×				
PMAP	×				×						×				×					×					
Power Draw								×				×				×									
PTNM/GIS	×										×			×	×			×			×				
QUIKMap	×										×				×							×			
Roots					×					×	×	×			×	×	×				×				
Resource							×				×				×					×	×	×			
RVCS					×					×							×					×			
S600 Raster Geographic							×			×			×				×			×					
SICAD	×	×	×	×	×	×		×		×				×	×		×	×		×	×		×	×	×
Sirlin VIEW	×	×	×	×							×				×					×	×	×			
Site Comp 6	×		×		×			×		×				×	×		×						×		
Site Analysis								×		×	×	×											×	×	
Sokkia Software 4.0			×		×	×		×	×		×				×								×	×	
SPANS	×		×	×	×	×	×		×	×	×			×	×		×			×	×	×	×	×	×
SPASE									×					×	×					×	×	×			×
Stack Ware							×			×							×			×				×	
Statmap Ⅱ for Windows									×		×			×	×					×	×				
Strings/Landtrak/GeoBlocks	×		×		×					×	×		×		×		×				×	×		×	
SUS—25 & SUS—1	×								×		×				×						×			×	
SW—Image			×				×			×	×		×	×	×		×	×		×	×				×
SYNERGIS	×									×			×				×	×		×	×				
System 9	×									×							×			×	×	×	×	×	×
Tactician									×		×	×	×		×	×	×			×					
TADS			×					×			×		×		×			×				×			
TerraSoft	×	×	×							×	×				×		×			×	×		×		
"TIGER" GIS	×		×	×							×				×										×
Topologic	×													×	×					×	×				
Total Utility Mapping System		×		×	×	×				×			×				×	×			×	×			
VANGO	×		×	×	×	×		×					×								×		×	×	
Vortext	×		×	×	×					×						×	×			×	×	×	×	×	

续表

系统名称	软件类型									计算机										数据结构					
										操作系统					机型										
	GIS	设备管理	自动制图	DBMS	数字化	格式转换	遥感	CAD	其它	Unix	DOS	Macintosh	VMS	其它	PC及兼容器	Macintosh	工作站	微型主机	其它	栅格式	拓扑矢量式	非拓扑矢量式	TIN	3-D三维显示	其它
VTRAK	×		×		×	×				×			×				×			×	×				
Water Integrated Facilities Mgt	×	×	×			×								×			×	×			×	×			
Water, Waste Wates etc. Systems		×	×			×			×	×	×		×		×		×	×			×				
WDMS, WCMS, GPMS, EDTSMS	×	×								×	×		×	×	×		×	×	×						
Windows/on the World	×		×								×			×						×	×				

参 考 文 献

[1] 郭达志、余兆平：地理信息系统基本原理与应用(内部讲义)，1991

[2] 黄杏元、汤勤：地理信息系统概论，高等教育出版社，1990.9

[3] 兰运超等编译：地理信息系统原理，广东省地图出版社，1991.10

[4] 李德仁等：地理信息系统导论，测绘出版社，1993.9

[5] 龚健雅：整体 SIS 的数据组织与处理方法，武汉测绘科技大学出版社，1993.10

[6] 曹桂发等：城市规划与管理信息系统，测绘出版社，1990.12

[7] 马建文、阎积惠：地理信息系统及资源信息综合，地质出版社，1994.8

[8] 资源与环境信息系统国家规范研究组：资源与环境信息系统国家规范研究报告，1984.9

[9] 詹慈祥等主编：资源与环境信息系统软件研究，测绘出版社，1990.10

[10] 蒋景瞳等主编：全国性资源与环境信息系统研究，测绘出版社，1991.7

[11] 王能斌、董逸生：数据库设计与实现，华中理工大学出版社，1991.9

[12] 朱海滨：面向对象技术-原理与设计，国防科技大学出版社，1992.10

[13] 童颖、沈一栋：知识工程，科学出版社，1992.7

[14] 黄可鸣：专家系统导论，东南大学出版社，1988.12

[15] 徐庆荣、杜道生：计算机地图制图原理，武汉测绘科技大学出版社，1993.6

[16] 钟玉琢、李树青：多媒体计算机技术，清华大学出版社，1993.5

[17] 蒋景瞳等译：美国国家数字制图数据标准，测绘出版社，1990.11

[18] 资源与环境信息系统国家重点实验室年报，1992

[19] 中国 GIS 协会筹委会办公室：中国 GIS 技术与管理交流会论文选编，1994.3

[20] '94 地理信息系统学术讨论会：论文集，1994.11，北京

[21] 总参测绘研究所：GIS 面向对象的空间数据库译文集，1994.8

[22] 郭达志、余兆平：地理信息系统在矿山的应用，测绘学报，1989.No.4

[23] 郭达志、韩国建等：土地利用调查与分析系统 CMLUS 的原理及其应用，中国土地科学，1991.No.3

[24] 郭达志、杨维平：A Spatial and Temporal 4-D Data Model for a GIS，FIG 国际学术会议论文集，1991.5. 北京

[25] 韩国建、郭达志、金学林：矿体信息的八叉树存储和检索技术，测绘学报，1992，No.1

[26] 黄幼才、刘文宝等：GIS 空间数据误差分析和处理，中国地质大学出版社，1995.6

[27] 郭达志、韩国建等：地籍信息的自动采集与提取，中国土地科学，1993.No.5

[28] 蒋银明、郭达志、金学林：PC ARC/INFO 地理信息系统汉字注记的探讨，遥感信息，1993，No.4

[29] 郭达志、盛业华等：矿区岩溶陷落柱探测中遥感与地理信息系统技术的应用研究，测绘学报，1994.No.2

[30] 盛业华、郭达志等：地籍信息系统数据采集的质量和精度分析，中国土地科学，1994.No.5

[31] 方涛、郭达志：概括模型下的线性 2^n 叉树的自然数编码，环境遥感，1994，No.4

[32] 郭达志：地理信息系统的数据结构和数据模型及其矿山应用，全国 GPS、遥感与 GIS 技术在矿山应用学术讨论会论文集，1994.10。贵阳

[33] 郭达志：地理信息学的内涵与发展，全国 GPS、遥感与 GIS 技术在矿山应用学术讨论会论文集，1994.10. 贵阳

[34] 郭达志：发展我国矿区地理信息系统的基本模式和技术路线，地理信息世界，1994，No.2

[35] 盛业华、郭达志：地理信息系统在矿区环境综合评价中的应用，煤矿环境保护，1995.No.4

[36] 盛业华、郭达志：GIS 环境下空间要素的制图综合方法，测绘通报，1995.No.3

[37] 盛业华、郭达志：矿图扫描数据字符的提取与识别，'95 全国矿山测量学术会议论文集，1995.7. 太原

[38] 郭达志、蒋银明、盛业华：点阵汉字矢量化技术及其在 PC-ARC/INFO 地理信息系统中的应用，地图

(地图学与地理信息系统),1995. No. 3
9] 郭达志、郝庆旺:开采沉陷预计与分析的空间模拟,中国矿业大学学报,1995. No. 3
)] 郭达志、盛业华、金学林:矿产资源开发对环境影响的时空模拟,'95 城市与工程测量新技术应用学术研讨会论文集,1995. 9,哈尔滨
.] 黄杏元:地理信息系统讲座,江苏测绘,1994. No. 1. 2. 3. 4
2] 张家庆、张军:九十年代 GIS 软件系统设计的思考,测绘学报,1994. No. 2
3] 吴炳芳等:地理信息系统的发展,地理学报,1994 增刊
] 陈常松:我国 GIS 数据产品市场化问题初探,测绘科技动态,1995. No. 1
[45] 陈述彭、赵英时:遥感地学分析,测绘出版社,1990. 5
[46] 香港富融科技有限公司:ARC/INFO 中国通讯,1995. 1
[47] 姚艳敏:世界 GIS 软件特征调查表,遥感信息,1993. No. 2. 3. 4,1994. No. 1
[48] Won Kim:计算机科学,1994. No. 4
[49] 柯正谊、何建邦、池天河:数字地面模型,中国科学技术出版社,1993
[50] 美国环境系统研究所编写、中科院地理研究所资源与环境信息系统国家实验室译:ARC/INFO 地理信息系统,1989
[51] Robert Laurim, Derek Thompson, Fundamentals of Spatial Information Systems, Academic Press, 1992
[52] Star J. , John Estes, Geographic Information Systems, An Introduction, Prentice-Hall. Inc. 1990
[53] Antenucti, J. C. , Kay Brown etc. , Geographic Information Systems, A Guide to the Technology, Van Nostrand preinhold, 1991
[54] Newton P. W. , etc. , Networking Spatial Information Systems, Belhaven press, 1992
[55] Proceedings of the 4th International Symposium on Spatial Date Handling (Zurich, Switzerland, 1990), Vol. 1, 2
[56] Proceedings of the 5th International Symposium on Spatial Data Handling (Charleston, U. S. A. , 1992), Vol. 1, 2
[57] Advances in GIS Research, Proceedings of the 6th International Symposium on Spatial Data Handling (Edinburgh, U. K. 1994), Vol. 1, 2
[58] Goodchild, M. F. , Modeling Error in Spatial Databares, GIS/LIS 1990 Proceedings
[59] Moellering H. Spatial Database Transfer Standards: Current International Status, Elsevier Science Publishers Ltd, 1991
[60] Wortman K. and J. L. Morrison, Special Issue: Implementing the Spatial Data Transfer Standard, Cartography and Geographic Information Systems, Vol. 19, 1992, No. 5
[61] Mason, D. C. , M. A. O'conaill etc. , Handling Four-Dimensional Geo-referanced Data in Environmental GIS, I. J. GIS, 1994, Vol. 8, No. 2
[62] Coleman David J. and J. D. Mclaughlin, Standards for Spatial Information Interchange: A Management perspective, CISM Journal ACSGC, Vol. 46, 1992, No. 2
[63] Phillips, Magor M. , The Digital Geographic Information Exchange Standards, Proceedings of the Canadian Conference on GIS, 1993, Otawa
[64] TSAI, Victor J. D. , Delaunay Triangulations in TIN Creation: An Overriew and a Linear-Time Algorithm, I. J. GIS, 1993, Vol. 7, No. 6
[65] Advances in GIS Research, Proceedings of the Sixth International Symposium on Spatial Data Handling, University of Edinburgh, 1994, Taylor & Francis Publishers.