目　录

第一章　地理信息系统基本概念 …… (1)
§1-1　概述 …… (1)
§1-2　与信息有关的概念 …… (3)
§1-3　地理空间数据与地图 …… (5)
§1-4　地理信息系统 (GIS) …… (6)
第二章　GIS 的基本构成及运行环境 …… (14)
§2-1　GIS 的硬件设备 …… (14)
§2-2　GIS 的软件模块 …… (17)
§2-3　GIS 地理空间数据和 GIS 组织管理人员 …… (19)
§2-4　系统运行环境 …… (25)
第三章　GIS 数据结构 …… (30)
§3-1　GIS 的数据 …… (30)
§3-2　矢量数据结构及其编码方式 …… (34)
§3-3　栅格数据结构及其编码方式 …… (41)
§3-4　矢量数据与栅格数据的转换 …… (51)
§3-5　GIS 数据模型 …… (60)
第四章　GIS 的基本功能 …… (66)
§4-1　数据输入功能 …… (66)
§4-2　图形与文本编辑功能 …… (70)
§4-3　数据存贮与管理功能 …… (73)
§4-4　空间查询与空间分析功能 …… (76)
§4-5　数据输出与表达功能 …… (85)
第五章　GIS 工程设计 …… (90)
§5-1　设计模式概述 …… (90)
§5-2　系统目标 …… (96)
§5-3　系统分析 …… (99)
§5-4　系统设计 …… (104)
§5-5　系统实施和系统评价 …… (114)
第六章　GIS 工具软件及其应用 …… (121)
§6-1　ARC/INFO 系统概述 …… (121)
§6-2　ARC/INFO 系统基本操作 …… (131)

§6-3 ARC/INFO 系统应用 …… (139)
§6-4 其它 GIS 工具软件简介 …… (147)
附录 国外八种主要 GIS 软件的基本情况与基本功能一览表 …… (152)
主要参考文献 …… (158)

第一章　地理信息系统基本概念

§1-1　概　述

随着信息产业的形成和发展并日益受到人们的重视，计算机技术和系统分析方法的广泛应用为现代科学技术的发展展现了广阔的前景。信息时代是以信息源的科学管理和充分利用为特性的。进入信息时代的地学，对地学信息的采集、管理、分析提出了更高的要求。可以说，信息时代地学的发展水平，取决于对有关地学信息以及其它与之有关各类信息的采集获取和分析技术水平。因此，地理信息系统（Geographic Information System，简称GIS）作为一门介于信息科学、计算机科学、现代地理学、测绘遥感学、空间科学、环境科学和管理科学之间的新兴边缘学科应运而生，并且迅速形成一门融上述各学科及其各类应用对象为一体的综合性高新技术。

地理信息系统是在计算机软硬件技术支持下采集、存储、管理、检索和综合分析各种地理空间信息，以多种形式输出数据或图形产品的计算机系统。它的外观表现是计算机软硬件系统，但是其内涵却是由一些计算机程序和各种地学信息数据组织而成的现实空间信息模型。通过这些模型，可以从视觉、计量和逻辑上对现实空间从功能上进行模拟；通过计算机程序的运行和各类数据的变换还可以对各类信息变化进行仿真。具有一定地学知识的用户可以在地理信息系统支持下提取现实空间模型各不同侧面、不同层次的空间和时间特征，快速地模拟自然过程的演变或思维过程的结果，取得预测或“实验”的结果，选择优化方案。而这种快速的信息模拟对于现有的GIS来说几乎是没有什么代价的，并且能够避免错误的决策带来的损失。因此可以说，GIS就是综合处理和分析空间数据的一种技术系统。

地球是一个客观存在。这一客观实体与人类存在着广泛的、密不可分的信息联系。人们所认识的地貌、地物、土壤、植被和城市、区域等等都是地理客观实体向人类发送的各种信息。人们正是通过对这些信息进行分类、评价、分析后得到用来指导决策规律性的信息或知识，从而可以完成诸如资源调查、区域规划、土地管理、环境保护等利用和改造客观世界的任务。所以可以说，没有信息就没有统一的世界，没有信息传递作为地理客观实体与人们感受到的现实信息世界之间相互作用的方式，人们就无法真正认识自然，更无法改造自然。基于此，地理信息系统就是要以地理信息世界模拟表达地理现实世界，通过信息联系反映出客观实体之间的联系，对客观世界中各种具有空间特征的事物、关系和过程进行描述、分析和仿真。因而，GIS不仅在地球科学，而且在社会经济和规划管理等社会科学领域里也得到了越来越广泛的应用，受到越来越明显的关注。

这方面的例子有很多，例如：在城市规划过程中，对城市中救护车、救火车的分布位置以及行车路线和控制的规划；如何安排多通路警车交通路线，以保证在紧急时刻，在任

意地方应至少能有一辆警车在事发后最短时间内赶到出事地点；在环境保护方面，对水土流失导致土地资源的破坏进行评价；在区域环境质量现状评价工作中，对整个区域的环境质量进行客观地、全面地评价，以反映出区域中受污染的程度以及空间分布状况；在国土整治与管理规划中，建立国土资源信息系统，作为GIS的一个分支，用现代化科学技术服务于国土规划和国土资源综合开发研究；在灾害预测和防治方面，应用GIS支持下的空间模拟分析技术，可以模拟受灾范围，估算受灾损失，评估抗灾方案实施的先验效益等；应用网络分析方法可以对灾民撤退路线，救灾物资的合理调配进行安排，选择最佳通道，并可模拟险情发生时人员撤退方案的实施；在城市公用设施的规划与管理方面，应用GIS技术可以处理大量复杂的地下管线数据和各种地形图、竣工图数据以及与之相应的大量属性数据和各种市政与测绘资料数据；还可建立全球人口数据库以及与之配套的空间分析模型等等。总而言之，人们在对各种自然资源进行开发、规划、分析和管理的过程中，使用以GIS现代化手段建立起的管理模式对总体规划、合理布局、综合平衡、协调发展和科学决策等都起着十分重要的作用。

60年代中后期，一些世界先进国家开始研究和开发GIS。随着计算机技术的发展和数字空间数据使用价值的提高，GIS的研究和建设得到了飞速发展。在发达国家中，GIS经过从60年代中期到80年代中期，约20年左右的时间的发展过程，已经进入专业化应用与商品化生产阶段。例如，美国白宫、日本东京市政府和英国首相府等早在80年代中期就建立了基于GIS的高层宏观决策系统，为政府部门提供了有效的决策服务。可以预计，到本世纪末这一阶段将不断丰富成熟，大量GIS产品将涌入商品市场。GIS技术在不久的将来将进入全面推广实用阶段。

当前，西方各发达国家的政府和私人企业都在投入大量人力、物力资源以发展各自的GIS系统，致力于建立国家规模的、地区性的以至全球性的GIS系统。GIS将成为一个更广泛应用的综合产业化结构体系，并将拥有一个稳定的需求市场。

我国GIS的发展是从80年代初研究资源与环境信息系统开始的。虽然和世界先进国家相比起步晚了20年，但在短短的10多年里，特别是近5年里有了令人瞩目的长进。在我国，目前已经相继成立了一些专门的GIS研究实体，形成了具有一定规模的专业队伍，并且已建立或正在建设一批全国性、地方性综合或专题空间型信息系统。例如，自1984年开始由国家测绘局测绘科学研究所研究和建立的国土基础信息系统就是起步较早、投资规模较大、发展较快的一个。它将全国1∶100万比例尺地形图存入其中，建立起全国范围内的1∶100万比例尺数字地面模型、地形要素数据库和地名数据库，可做多种空间分析。另外还建有全国土地信息系统；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等，并且还开发了黄土高原信息系统以及洪水灾害预报与分析系统等专题地理信息系统。目前，这些系统已在多个领域中，如防洪实时监测洪水、快速灾情分析、水土保持、国土管理等方面都发挥了作用。除上述全国性的信息系统外，近年来我国的许多大中城市也在纷纷开发和建立城市信息系统的工作。例如，上海市于1995年建成的综合地理信息系统，不仅为市政府领导进行宏观分析决策提供多要素、多层次、多时态的空间信息，同时也为办公管理提供了先进、科学的分析、预测、规划、决策等技术手段。该系统的建立和使用

在上海市人民政府办公决策服务方面发挥了重要作用。据不完全统计，我国目前已建立具有一定规模的城市信息系统的城市还有：海口、深圳、北京、天津、北海、常州、沙市、洛阳、厦门等；此外，另有一些沿海及内地城市，如青岛、福州、武汉、烟台、佛山、大庆等，也正在积极筹备建立自己的城市地理信息系统。

综上所述，地理信息系统作为一门新兴边缘学科正在世界范围内日益广泛地得到应用，并且正在或已经形成为一个独特的行业而倍受用户的欢迎。可以说，GIS 不仅是进行资源普查、区域开发规划、国土管理规划、环境保护、灾害预测与防治以及城市规划和调度决策管理的根本基础，也是政府、企业、公司进行现代化管理及预决策的强有力手段，因而具有广阔的应用前景。

§1-2 与信息有关的概念

了解有关信息的一些基本概念对地理信息系统的认识是很必要的。信息产业作为一种新兴产业越来越受到人们的重视，信息革命的浪潮正在冲击着人类社会。人们正是通过对各种信息进行分类、评价、分析后得到指导决策规律性的知识，从而完成认识、利用和改造客观世界的目的。

信息是近代科学的专门术语，现在已经广泛地应用于社会各个领域。关于信息有多种定义，简单地讲，人们在进行现实世界管理时，将客观事物在人们头脑中的反映称之为信息。从这一角度来看，信息就是现实世界状态的反映，是向人或机器提供关于现实世界新的事实的知识，是用数字、文字、符号、语言等介质来表示事件、事物、现象等的意义和内容。信息不随载体的物理设备形式的改变而改变，它具有客观性、适用性、可传输性和共享性等特征。

——客观性，任何信息都是与客观事实紧密相关的，这是信息的正确性与精确度的保证。

——适用性：信息对决策是十分重要的，信息系统将地理空间的巨大数据流收集和组织管理起来，经过处理、转换和分析变成对生产，管理和决策具有重要意义的有用信息，这是由建立信息系统的明确性所决定的。

——传输性：信息可在其发送者和接收者之间传输，既包括系统把有用信息送至终端设备（包括远程终端）和以一定的形式或格式提供给有关用户，也包括信息在系统内各个子系统之间的传输和交换。目前，已有很多系统采用了网络传输技术。

——共享性：信息与实物不同，信息可以传输给多个用户，为多个用户共享，而其本身并无损失。这为信息的并发应用提供了可能性。

信息与数据有密切联系。数据是指对某目标进行定性、定量描述的原始材料，包括数字、文字、符号、图形、影像等形式。它是用以载荷信息的物理符号，在计算机化信息系统中往往与计算机系统有关。由于计算机只能处理数据化信息（即只能用字母、数字或符号表示），所以也可以说数据就是信息的符号表示。

由上所述可知，信息是用与物理介质有关的数据来表达，而数据中包含的意义就是信

息，所以信息与数据是不可分离的。数据是未加工的原始材料，是对现实世界中客观对象的表示，其具体形式是多种多样的，也可以由一种数据形式转换成其它数据形式，但其中包含的信息内容不改变。由此可知，数据是信息的载体，但并不等于信息，只有理解了数据的含义，对数据做出解释，才能得到数据中所包含的信息。例如从测量数据中可以抽取出地物的形状、大小和位置等信息；从遥感图像数据中可抽取出各种图形和专题信息；从实地调查数据中抽取出各专题的属性信息等。再例如，数字“1”，可以离开地理信息系统而独立存在，也可以离开地理信息系统的各个组成和阶段而独立存在；即它既可以回避实体是什么，也可以回避它本身能做什么，而且在计算机化的地理信息系统中，数据的格式往往和具体的计算机系统有关，随载荷它的物理设备的形式而改变。而信息则是对数据的解释、运用和解算，即使是经过处理后的数据，也只有经过解释才有意义，才能成为信息，就本质而言，数据是客观对象的表示，信息则是数据内涵的意义，只有数据对实体行为产生影响时才成为信息。同样是数字“1”，当用来标识某一种实体的类别时，它就提供了特征码信息；当用来表示某一种实体在某个地域内存在与否时，它就提供了有无的信息（1表示有，0表示无）；当在绘图矩阵中表示绘线与不绘线时，它就提供了抬落笔的信息等等。可见，信息是用数字、文字、符号、语言等介质来表示事件、事物、现象等的内容、数量或特征，以便向人们或系统提供关于信息世界新的事实的知识，作为生产、管理和决策的依据。

对数据进行处理，如采集、筛选、运算、排序、归并、转换、存储、编码、分类、检索、计算以及分析、模拟和预测等操作，就是为了得到数据中所包含的信息。这些操作的目的就是：

（1）把数据转换成便于观察、分析、传输或进一步处理的形式；

（2）把数据加工成对正确管理和决策有用的数据；

（3）把数据编辑后存储起来，供以后不断使用。

数据处理是为了解释，而这种解释需要人的智慧、知识和经验作用到数据上才可以得到信息，获得信息量的多少，则与人的知识水平高低有关。对同一数据，每个人的解释可能不同，其对决策的影响也可能不同。而不同的解释则往往又来自不同的背景、目的和应用。

数据处理工作由来已久。随着计算机技术的发展，数据处理进入了电子数据处理系统时代。由于采用磁介质和半导体等存贮技术，数据管理主要是在提高数据独立性、降低数据冗余度、数据共享、提高数据的安全性和完整性等方面得到了迅速发展。对数据结构的研究就是对数据的组织形式的研究，通常可分为两种形式：数据在物理存贮设备上的组织称之为数据物理组织或物理结构，一般以文件形式进行组织，根据其特点可分为顺序文件、索引文件和随机文件等；数据在使用者面前所显现的组织方式称之为数据逻辑组织或逻辑结构。对于一种数据的逻辑结构可以用不同的物理结构来实现。所以数据的逻辑结构是从概念上对数据间的关系进行组织，具有普遍性；而数据的物理结构则是在硬件上对某一逻辑结构设计的一个具体实现，具有特殊性。

地理信息是指与所研究对象的空间地理分布有关的信息。它是表示地表物体及环境固有的数量、质量、分布特征、联系和规律等的数字、文字、图像和图形等的总称。从地理

实体到地理空间数据，再到地理信息的发展，反映了人类从认识物质、认识能量到认识信息的一个巨大飞跃。地理信息属于空间信息，地理空间数据所包含的意义中除了对实体特征的属性描述以外，还包括对其位置的识别，因此地理信息具有区域性并且具有多维结构特征，即在同一位置上具有多个专题和属性的信息结构。例如，在一个地面点位上，可以得到该点处的高程、土壤类型、污染、噪声、降雨量等多种信息。此外，地理信息还具有明显的时序特征，即随时间而变化的动态特征。这就要求及时采集和更新它们，并根据多时相的数据和信息来确定出随时间变化的分布规律。

系统是具有特定功能的相互有机联系的许多要素所构成的一个整体。如计算机系统就是为实现某些特定功能，由必要的人、机器、方法或程序按一定的相关关系联系起来进行工作的集合体；计算机系统内部要素之间的相互联系通过信息流（数据流）得以实现。

所谓信息系统是指具有对数据进行采集、存贮、管理、分析和再现功能并且可以回答用户一系列问题的系统。在辅助决策过程中，信息系统可提供有用的信息。从计算机科学角度看，信息系统是由计算机硬件、软件、数据和用户四个主要要素组成的。在计算机时代，大部分重要的信息系统都是部分或全部由计算机系统支持的，如目前流行的图书情报信息系统、经营信息系统、企业管理信息系统、财务管理信息系统、人事档案信息系统、空间信息系统和其它一些信息系统等。其中，空间信息系统是一种十分重要而又与其它类型信息系统有显著区别的信息系统，因为它所要采集、管理、处理和更新的是空间信息。因此，这类信息系统在结构上也比其它一般信息系统复杂得多，但功能上也较其它信息系统强得多。

§1-3 地理空间数据与地图

一、地理空间数据

空间数据是指以空间位置为参照的数据。地理空间数据是指人们通过观测地球上的某些位置上地物、地貌的空间数据。而地图是地理空间数据的一种表示形式，也可以说是地理信息载体和地理语言。

空间位置通常是用空间实体与某参考坐标系统之间的关系来表示。这种参考坐标系统可以是有限区域内的局部坐标系，也可以是纳入国家坐标网系统的坐标系或者是属于国际性投影的横轴椭圆柱正形投影坐标系统。根据空间实体与参考坐标系之间的联系，通常采用绝对坐标或相对坐标来表示空间位置。各种地理实体空间位置的表示在形式上是多种多样的，但是通过分析研究，可以将其抽象为最基本的实体：点、线、面。从理论上讲，任何一种地理实体都是可以用点、线或面这三种基本实体再加上说明注记来表示的。这种方法可以表示：

(1) 用点实体表示不同尺度空间实体。如：一口油井，可以用一个点实体的坐标 (x, y) 和注记“油井”来表示。另外，如城镇、高程控制点、交通网络结点等也可以用点实体表示。

(2) 用线实体表示具有线性特征的地理实体。例如：河流、铁路、公路等，可以用

个具有起点坐标（x_1，y_1）和终点坐标（x_2，y_2）的线段以及有关注记来表示一段道路。

（3）用面实体表示各种规则与不规则的多边形地理实体。例如：土壤分布、植被类型、空间影响范围等。

空间数据表示了各种空间现象之间的空间关系，如邻接、距离、重叠、包含等。空间现象的其它属性也可以用数据表示。地理空间数据可以用来描述地球上可更新的与不可更新的自然资源，人类赖以生存的自然环境，人文经济与劳动力资源的有关信息，包括：地形、地貌、土壤、岩石、矿藏、植被、水、土地利用、行政界线、交通网络、政治分区、公共设施位置、土地界线、土地价值、土地所有权、人口普查分区、人口分布、收入分配、经济区划、环境污染、疾病影响范围、自然灾害分布等等。

地理空间数据可以由位置组合变量的表格形式表示，也可以由相对位置在地图或图像上表示。目前最常用的就是以地图表示的地理空间数据。

地理空间数据是GIS中的重要组成部分，通常可以用数字化仪、扫描仪、键盘、磁带机或其它系统通讯输入GIS。它们是系统程序作用的对象，是地理信息系统所表达的客观世界的经过抽象的实质性内容。一般应以下述三个方面内容构成其最基本的三维结构数据：

（1）由某个参考坐标系确定的位置；

（2）地理实体之间的空间拓扑关系；

（3）与几何位置无关的属性。

GIS特殊的空间数据模型决定了它采用特殊的空间数据结构和编码方法，也决定了它具有特殊的空间数据管理方法和系统空间数据分析能力。

二、地图的发展

在人类社会的发展历史上，采集和获取地理空间数据一直是人类认识自然、改造自然活动的一个重要组成部分。长期以来，这项工作主要是由航海家、地理学家和测量学家来完成的，并且主要是以地图的形式来表示采集到的地理空间数据。

最早的地图是用于航海和领地测量。在古罗马时代，土地测量人员为帝国提供了丰富的地况地图资料，测量业在国家的支持下得到较大发展。罗马帝国衰败以后，测绘地图的发展趋于停顿。直到18世纪，随着欧洲文明的昌盛，地图生产才又开始复苏。此时，欧洲各国政府开始意识到系统地测绘出本国领土的重要性。因此，经典测量学在理论和实践上随之不断完善和发展直至今日。

随着人类对地球认识的不断丰富，需要进行测绘的内容也不再仅限于对地球表面的地物进行量测。从19世纪开始，地质学、地貌学、土壤科学、生态学以及土地管理等关于地球资源学科的出现，使得需要进行测量与绘图的内容大大增加了。一般的地形图所含的内容已不能再满足人类生产的特殊需要。为解决这一矛盾，所谓的专题地图由此应运而生。这些为特定目的而测绘的专题图应用于很多行业。例如，在土壤学中的土壤分布图、土地酸碱性分布图；在生态学研究中，对某地区出现某种传染病的影响范围分布图；气象学上的气压变化分布图等。一般是在经某种简化的地形底图上，由不同用户添加上不同的专题信息而得到这些专题图。通常，采用等值线（或等高线）或晕线的方法绘制，这种方法不仅能对新采集的地理空间数据进行定性描述而且还可以进行定量分析。因此专题图的出现与

使用为人类在地球资源方面提供了十分有力的工具。进入 20 世纪，航空摄影测量的应用以及后来出现的遥感技术，使得人类采集地理空间数据的技术大大地提高，由此，人类可以在更大的范围内，以更高的精度获取更大量的地理空间数据。随着已获取的数据量不断增加，人们的注意力开始从简单地追求数据采集获取过程转到寻求更为合理有效的使用这些数据，以及如何从已有数据中提取更多的信息。

到了本世纪 60 年代初，随着电子计算机的广泛应用，在计算机图形学的基础上出现了计算机化的数字地图。计算机辅助制图和辅助专题地图分析技术的发展对地图进行综合分析和制图输出的系统日益增多，数据自动采集、存贮、分析和表示技术也得到极大的发展，地图作为被地学工作者普遍使用的一代地理语言在技术上开始具备向新一代地理语言—GIS 的转变。

三、地图的性质

地图包括表示地表形态和地表地物分布（水系、植被、土质、居民点、交通网、国界和行政区划等）的普通地理图和地形图，还包括详细客观地表示某种自然要素或社会要素的专题地图。严格地讲，地图应定义为：由数学所确定的经过概括并用形象符号表示的地理表面在平面上的图形及用其表示各种自然现象和社会现象的分布、状况和联系，根据其具体用途对所表示现象进行选择和概括而得到的图形。

这就是说，地图包括三个方面的内容：

1. 由数学决定的结构。即必须首先确定地球表面上点的地理坐标或其它坐标与这些点在地图平面上的坐标系统之间的严格的映射函数关系。确定了这些函数关系，就可以利用地图来研究它所表示的地物的空间关系。

2. 特定的符号系统。地图采用便于空间定位的形式表示各种现象及其性质和相互关系的专用图解符号，用于记录、转换和传递各种自然和社会现象的知识，从而在地图上构成客观实际的空间现象。

3. 现象表示的取舍和概括，或称地图综合。由于地图图幅比例的限制和图解符号表示能力的有限性造成对某些碎部地物的取舍和概括。有时，地图的用途也常需要对自然景观进行取舍和概括表示。

实际上，地图是一种图解图像，是由点、线、面组成的集合，这些对现实世界抽象化后的点、线、面实体相对于某一参考坐标系的空间位置以及它们的非空间属性，如注记、标志、符号等决定了该集合的性质。从地图上我们可以获得一个地区或整个地球表面的同一时间的空间表象，它们建立起地物形状、尺寸和相对位置的视觉形象，从而可以从地图上得出地表景象的空间尺度，即空间坐标、长度、面积、高度和体积。除表示地物的数量和质量方面的特征以外，地图还可表示出地物之间的空间联系和其它联系。如：邻近性、包含性、叠加性、相对距离、从属关系、经济和交通方面的联系等等。

对于能够理解地图符号的使用者来说，地图是自然界的一种模型，是制图人根据对自然世界的认识，用概括和简化的可视形式表示客观实际某些方面的结构，而地图用户则通过对地图符号的理解，建立起地图所表示的客观实际的空间模型。这是一种思维模型，它的建立依赖于制图人和用图人对地图符号的认识的一致性。因此可以说，地图是制图人与

地图用户之间交流自然空间信息的图像载体。

地图的表示范围也是相当广泛的。地图表达的对象可以是具体的（如居民地、植被、土壤等），也可以是抽象的（如人口密度、工农业产值、影响范围等）；可以是现实的（如道路、河流、城市、土地利用等），也可以是预期的（如规划的道路、规划的水利灌溉网、土地利用规划等）。此外，地图还可表示自然现象的发生发展过程，表达其空间上的变化和时间上的发展，如表示城市化过程的城市扩散图、土地利用变化图、土壤侵蚀速率图等等。由上所述，我们可以看到地图具有表达空间地理信息精确、简洁、丰富、灵活等特点。

但是，当地学科学进入信息时代以后，从传统的制图与用图角度看，将地图以及有关注记说明资料等作为表达地理空间数据的一种形式，实际上仅相当于一个地理空间数据库的作用。从成图过程中的数据采集、编码成图，直至应用阶段，传统的制图技术已不适合当今地理空间信息变化的要求，这可以从以下几点看出：

1. 在地图成图过程中，对地物进行取舍和概括，将不可避免地会忽略或删除某些局部碎部地物，从而使原始数据获取量减少。

2. 从传统地图的制图过程来看，从地图的数据采集、编辑直至印刷出版是一项耗资耗时的过程。

3. 绘到地图上的地理空间数据，因为难以实现直接对其进行空间查询和空间分析的功能，所以不易于其它空间信息数据结合使用。

4. 对于已印制好的地图来说，它仅是对所绘区域的一份过去时态的、静态、定性的地理文件，缺乏实时动态地反映现实地理空间数据变化的功能。

在当今人类获取地理空间数据的手段，如航测像片、GPS 卫星数据、遥感影像等大大地发展的时代，人类可以实时动态地、全面地、快速地、大量地采集获取数据，人们今天所面临的地理空间数据不仅在数量上远远大于过去，更为重要的是在数据结构上更为复杂多样，难以用单一的点线面结构形式完整描述；数据间联系紧密，相互关联，并且有涉及区域广泛、精度要求高、变化复杂、要素众多等特点。因此，为了使这些地理空间数据所表达出的地理信息模型能够贴切等价地反映出地理现实世界，并且能够真实、快速而又毫无损失地模拟各种自然的和思维的过程，地图作为被普遍使用的第二代地理语言将逐渐被 GIS 所取代。而地图本身作为一种图解图像、一种地理模型，一种现实世界空间信息的高效载体和一种地理思维工具，也成为 GIS 中最主要的数据源。

§1-4 地理信息系统（GIS）

GIS 是一门属于高技术领域里的新兴的交叉学科。在 §1-1 中，我们已经给出了 GIS 的定义，它是以地理空间数据库为基础，采用地理模型分析方法，适时提供多种空间的和动态的地理空间信息，用于管理和决策过程的计算机技术系统。本节将要对 GIS 的特征、类型及其研究体系和任务进行讨论。

一、GIS 的特征

随着计算机绘图和空间分析技术的发展，在很多相关专业领域里，尤其是地籍测量、地

形绘图、专题图制图、土木工程、地理、土壤科学、空间变量的数学研究、测量与摄影测量、城乡规划、公用设施网络以及遥感和影像分析等行业中，在诸如自动化数据采集获取、空间数据分析、数据输出显示等应用技术方面亦得到很大发展。GIS 的出现则是将这些在不同行业中对空间数据处理过程的不同应用有机地联系起来所形成的一种技术系统。GIS 的出现正在成为人类对客观世界中各种具有空间特征的事物、关系和过程进行描述、分析和模拟，进而根据所得规律指导人类利用和改造客观世界的强有力的工具。换言之，GIS 为人类由客观世界到信息世界的认识、抽象过程以及由信息世界返回客观世界的利用改造过程的发展和转化，创造了空前良好的条件和环境。

综上所述，地理信息系统具有如下特征：

1. 具有采集、管理、分析和以多种方式输出地理空间信息的能力，具有空间性和动态性，GIS 的数据必须具有空间分布特征，具有一个特定投影和比例的参考坐标系统，基于共同的地理基础，并且是多维结构的。

2. 为管理和决策服务，以地理模型方法为手段，具有区域空间分析、多要素综合分析和动态预测能力，产生决策支持信息及其它高层地理信息。

3. 由计算机系统支持进行地理空间数据管理，并由计算机程序模拟常规的或专门的地理分析方法，作用到空间数据之上产生有用信息、完成人类难以完成的任务。计算机系统的支持使得地理信息系统具有快速、精确并能综合地对复杂的地理系统进行空间和过程的动态分析。

所以，GIS 的功能决不仅仅限于对现实世界中地理空间数据的采集、编码、存储、查询和检索，而是现实世界的一个抽象模型，它比由地图表达的现实世界模型更为丰富和灵活，用户可以按应用的目的观察提取这个现实世界模型各方面的内容，也可以量测这个模型所表达的地理现象的各种空间尺度指标，更为重要的是可以将自然发生的或者思维规划的动态过程施加在这个模型之上，取得对人为和自然过程的分析和预测信息，从而有助于做出正确决策。同时又能有效避免和预防不良后果的发生。因此，也可以说 GIS 又是一个地理空间数据试验场，无需方案的实施和结果的发生，就可以快速、灵活又毫无损失地进行数字模拟实验，取得科学依据。

二、GIS 的类型

从 GIS 发展的历史上看，对 GIS 系统按其内容进行分类，在很大程度上是由用户不同的应用目标或任务要求所决定的。各种 GIS 应用系统的出现就是从相应专业或学科以及应用范围的角度出发，对现实世界进行深入分析研究，借助计算机系统软硬件的支持，经过对现实世界的认识和抽象完成从现实世界到地理信息模型的转化。可以说，各种 GIS 系统的应用均反映了人类对现实世界利用改造的能动作用。这同时也是 GIS 技术产生社会经济效益的关键所在。因此各种 GIS 应用系统的发展在 GIS 领域中都占据十分重要的地位。

GIS 按其内容可以分为三大类：

1. 专题信息系统——

具有有限目标和专业特点的地理信息系统，系统数据项的选择和操作功能设计是为特定的专门目的服务，如森林动态监测信息系统、水资源管理信息系统、矿产资源信息系统、

作物估产信息系统、草场资源管理信息系统、水土流失信息系统等等。

2. 区域信息系统——

主要以区域综合研究和全面的信息服务为目标，可以有不同的规模，如国家级的、地区或省级的、市级和县级等为各不同级别行政区服务的区域信息系统，也可以按自然分区或以流域为单位的区域信息系统。其特点是数据项较多，功能齐全，通常具备较强的开放性。如加拿大国家信息系统、美国 Oakridge 地区模式信息系统、我国黄河流域信息系统等。

3. 地理信息系统工具（GIS-TOOL）——

是专门为 GIS 的建立和开发而研制的通用软件系统，是一组高效率系统化的具有图形图像数字化、空间数据存储管理、查询检索、分析运算和多种输出等 GIS 基本功能的软件包，为用户提供基础的地理信息管理与分析功能。GIS-TOOL 可以是专门设计研制的，也可以是从实用地理信息系统中抽取掉具体区域或专题的地理空间数据后得到的。它具有对计算机硬件适应性强，数据管理和操作效率高、功能强，且具有普遍性并易于扩展、操作简便、容易掌握等特点。GIS-TOOL 作为 GIS 支撑软件以建立专题或区域性的通用型信息系统，也可作为教学软件。只要将其用于与特定任务相关联的地理空间数据，与相应的应用模型相结合，就可形成各种不同规模、不同专题的实用 GIS。目前在国内较为流行的 GIS-TOOL 如美国环境系统研究所研制的 ARC/INFO 系统，美国耶鲁大学森林与环境研究学院的 MAP 系统等。

回顾 GIS 的发展历史，在 70 年代大多数 GIS 软件的开发都是针对某个应用任务的需要而进行的，如上述 1 和 2 类 GIS 系统。它们专用性强，但推广应用的范围有限。从上述对 GIS 的分类并根据已有的大量 GIS 应用系统来看，许多实际的 GIS 是介于上述 1 和 2 类之间的区域性专题信息系统。

进入 80 年代后，随着社会对 GIS 需求的日益广泛，GIS 工具的观念出现愈来愈受到人们的重视，而且在实际上也有了很大的发展。仅以美国 GIS World 公司 1990 年统计数字为例，美国就有近 90 个 GIS-TOOL 系统在市场上销售。这些工具的使用加速了 GIS 新兴产业的形成，产生了明显的社会效益。

当然，GIS-TOOL 的研制仍处于一个发展阶段，在很多方面还不是十分成熟的，在功能覆盖应用程序接口，硬件适应面和使用灵活性上还不能满足所有领域不同层次的需要。但随着社会对 GIS-TOOL 需求的不断扩大，人们对它的研制水平的进一步提高，GIS-TOOL 已成为加速 GIS 的发展和应用以及向商业化转变的关键。

三、GIS 的任务

1. GIS 的研究体系

GIS 的任务就是对地球表层人文经济和自然资源及环境多种信息进行综合管理与分析，从 GIS 的研究体系看，主要有三方面的任务：

(1) GIS 基本理论的发展。包括研究 GIS 的概念、定义和内涵；GIS 的信息理论研究；建立 GIS 理论系统；研究 GIS 的构成、功能、特点和任务；总结 GIS 发展历史；探讨 GIS 发展方向等理论问题。

(2) GIS 技术系统设计。侧重于计算机软硬件技术系统的逻辑设计和物理实现，包括

GIS 硬件设计与配置；地理空间数据结构及表示；输入输出系统；空间数据库管理系统；用户界面与用户工具设计；GIS 工具软件研制；微机地理信息系统的开发等。

(3) 地理信息系统应用方法研究。涉及有关地理信息系统应用问题的研究，包括应用系统设计和实现方法；数据采集与检验；GIS 空间分析函数与专题分析模型；GIS 与遥感技术结合方法；地学专家系统研究等。

GIS 技术是在各地学科学研究的实践需求中产生的，GIS 应用使技术系统不断完善，并逐渐发展了 GIS 理论；理论研究又指导着开发新一代高效的 GIS，并不断拓宽其应用领域，加深应用的深度；GIS 的应用又对理论研究和技术方法提出了更高的要求。因此，上述三个方面的工作是相互联系、相互促进的。这是由 GIS 的发展和自身的特点所决定的。

从上面对 GIS 整体研究体系的研究内容分析并结合人类对客观世界的认识、利用和改造过程来看，GIS 在实现从现实世界向观念世界转化中，需要通过人类对现实世界的认识和抽象，即通过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要的概念模型，具体包括对现实世界实体特征描述，关系分析和过程模拟等内容，这些内容被抽象地影射为数据结构的定义、数据模型的建立以及专业模型的抽象等。它们构成了 GIS 基础研究的主要内容和任务。

2. GIS 的任务

从具体的 GIS 技术的角度分析，GIS 应完成以下几个方面的任务：

(1) 地理空间数据管理。数据管理是 GIS 的初级任务，即以多种方式录入地理数据，以有效的数据组织形式进行数据库管理、更新、维护，进行快速查询检索，以多种方式输出决策所需的地理空间信息。

地理信息系统数据管理的目的是提高决策的效率和科学性，其优点表现在：

① 信息量大，存储效率高，体积小，保存时间长，利于信息的共享和交流。

② 检索效率高，可即时输出辅助决策信息，省时、省力，降低管理费用。

③ 数据精度高，比例尺精确，便于不同图件之间的匹配，数据精度不随时间、温度和湿度变化而改变。

④ 输出灵活，利于根据不同的特殊任务选择不同内容、不同形式输出精确定义的空间信息。

⑤ 采用数字化存储，支持各种分析、统计和模拟实验。

⑥ 采用计算机辅助制图技术，支持非专业制图人员制图，并可快速更新、选择和合并地图，改进传统制图工艺。

空间数据管理是 GIS 的最基本的功能。只有在有效的数据管理基础上，才能对空间数据进行各种测量、分析和模拟实验，达到辅助决策的目的。

(2) 空间指标量算。GIS 是以定量的数字地图方式存储地理空间信息，便于灵活、快速、动态地对与地表有关的各种空间指标进行精确的量测，是对传统的地图量算方法的质的改进。

(3) 综合分析评价与模拟预测。GIS 不仅可以对地理空间数据进行编码、存储和提取，而且还是现实世界的模型，可以将对自然世界各个侧面的思维评价结果作用其上，得到综

合分析评价结果，也可以将自然过程、决策和倾向的发展结果以命令、函数和分析模拟程序的方式作用于这些数据上，模拟这些过程的发生发展，对未来的结果作出定量的和趋势性预测，从而预知自然过程的结果，对比不同决策方案的效果以及特殊倾向可能产生的后果，以做出最优决策，避免和预防不良后果的发生。

四、GIS 与其它计算机系统的区别

GIS 是在地学学科与数据库管理系统（DBMS）、计算机图形学（Computer Graphics）、计算机辅助设计（CAD）、计算机辅助制图（CAM）等与计算机技术相关学科相结合的基础上发展起来的。GIS 与其相关学科的关系可以用图 1-1 形象化地简单表示。

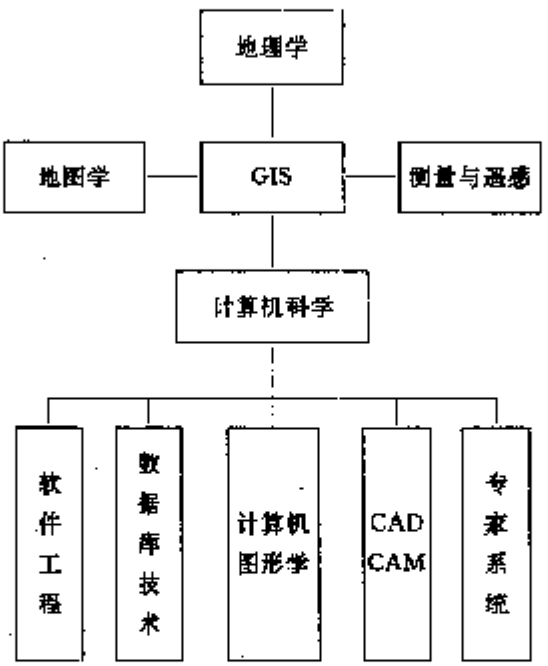


图 1-1　GIS 的相关学科

GIS 应用了许多计算机图形学的技术，但又与计算机图形学不同。计算机图形学主要是利用计算机处理及显示可见图形信息以及借助图形信息进行人机通讯处理的技术。通常计算机图形学所处理的图形数据是不带地理属性的纯几何图形，是地理空间数据的几何抽象。而对地理空间数据进行空间分析过程中，地理属性是不可缺少且十分重要的因素。因此，应用计算机图形学只能完成 GIS 底层的图形操作，可以说，它是 GIS 算法设计的基础。而 GIS 是随计算机图形学的发展而不断发展完善的，它除了能对图形信息数据进行显示和处理以外，还能完成数据的地理模型分析和许多具有地理意义的数据处理。

CAD 是使用计算机技术来辅助设计人员进行设计，以提高设计的自动化程度，节省人力和时间的一种计算机系统；CAM 是使用计算机技术进行几何图形的编辑和绘制的计算机系统。CAD 和 CAM 的使用使众多行业和部门中的设计人员从制图板上和手工劳动中解

放出来，因而受到广大设计人员、制图人员的青睐。GIS 与 CAD 和 CAM 有许多共同点，例如它们均可以处理非图形的属性数据；它们也都可以对空间数据建立空间相关关系，对所描述对象的拓扑结构进行处理等。GIS 与 CAD 和 CAM 的主要区别在于：

(1) CAD、CAM 不能建立地理坐标系和完成地理坐标变换。

(2) GIS 的数据量要比 CAD 和 CAM 的数据量大得多，数据结构、数据类型亦更为复杂；数据间联系紧密，这是因为 GIS 涉及的区域广泛、精度要求高、变化复杂、要素众多、相互关联，单一结构难以完整描述。

(3) CAD、CAM 不具备 GIS 具有地理意义的空间查询和分析功能。

从技术的角度分析，GIS 的基本任务之一就是对地理空间数据进行管理。因此，GIS 在数据管理上借鉴了数据库管理的理论和方法。数据库管理系统 (DBMS) 是操作和管理数据库的计算机系统。它提供可以被多个应用程序和用户调用的软件系统，具有对数据库的建立、更新、查询和维护功能。有时，对非几何属性数据也可采用通用数据库管理系统或在其上开发的软件系统进行管理。然而，在对地理空间数据的管理上，与 GIS 相比较而言，数据库管理系统 (DBMS) 尚存在两个明显的不足：

(1) 缺乏空间实体定义能力。目前流行的数据库结构，如网状结构、层次结构和关系结构等，都难以对地理空间数据结构进行全面、灵活、高效地描述。

(2) 缺乏空间关系查询能力。目前通用的 DBMS 的查询主要是针对实体的查询，而 GIS 中则不仅要求对实体的查询，还要求对空间关系进行查询，如关于方位、距离、包容、相邻、相交和空间覆盖关系等。因此，通用 DBMS 尚难以实现对地理数据空间查询和空间分析。

我们知道，数据是信息的载体，对数据进行解释可以提取信息。目前通用数据库和地理数据库都是针对数据本身进行管理，而 GIS 则在数据管理基础上，通过地理模型运算，产生有用的地学信息。取得信息的多少和质量是与所建立的地理模型水平密切相关的。

因此，GIS 是一种全新的区域研究支持系统，最适合于区域的地理特点，在思维方法上也有质的改进。GIS 不仅具有地理意义明确的空间数据管理能力，更重要的是可以通过地理空间分析产生常规方法难以得到的分析决策信息，并可在系统支持下进行空间过程演化的模拟和预测，只有在地理信息系统支持下，才能高效率、高精度、定量地实现真正地理意义上的区域空间分析和过程模拟预测。

第二章　GIS的基本构成及运行环境

从计算机科学角度看，完整的GIS至少应由四个主要部分组成，即GIS硬件系统、GIS软件系统、GIS地理空间数据以及系统的组织管理人员。这其中，硬件和软件系统是GIS的核心部分，空间数据库可以用来表达地球表层的地理数据，而GIS的管理人员和用户则决定系统的工作方式和信息表达方式。此外，由于GIS不同于一般的管理信息系统，它是一个集空间型、管理型于一体的信息系统，具有数据存储量大、功能繁多、处理复杂、规模宏大等特点。因此，根据GIS的特点选择能满足其功能要求的合适的运行环境，也是决定该GIS能否充分发挥其效率的关键问题之一。

§2-1　GIS的硬件设备

GIS的硬件设备构成GIS的物理外壳。系统的规模、精度、速度、功能、形式、使用方法甚至软件都与硬件设备的配套有极大的关系，受到硬件指标的支持或制约。GIS由于其任务的复杂性和特殊性，必须由计算机与其外围设备连接形成一个GIS的硬件环境。其硬件配置一般包括四个方面，即①计算机主机；②数据存储设备；③数据输入设备；④数据输出设备。其主要硬件设备构成如图2-1所示。

一、计算机主机

计算机是GIS的核心，是用作数据和信息的处理、加工和分析的设备，可以组成网络也可以单独使用。它的主要部件由中央处理器和主存储器构成。目前能运行GIS的计算机包括大型机、中型机、小型机、工作站和微型机。各类计算机的主要性能可以用处理速度、字长和内存容量来描述。

衡量计算机处理速度的方法有许多种，较为流行的有MIPS和时钟频率。MIPS的意义为百万条指令每秒，这里是指执行机器指令，是衡量计算机主机处理速度的指标；时钟频率则是指每秒变换内存中某一单元内容的次数，即指内存的访问效率。在微机上通常采用时钟频率作为处理速度指标，而工作站则多采用MIPS为指标。例如，微机486 DX/66表示时钟频率为66 M；而最近新推出的Sun-ultra系列工作站的处理速度可达160MIPS以上。

字长和内存也是影响计算机效率的二个重要指标。字长由比特（bit）位数来定义，而内存容量则通常是以千字节（kB）或兆字节（MB）为单位。现在市场上大多数微机的字长是16位或32位个人计算机，其内存容量从640 kB到128 MB以上；而有作为当前GIS主流机趋势的工作站，其字长一般为标准32位或64位，内存容量是从8 MB发展到128 MB，甚至可扩充到512 MB以上。

从目前形势看，由于工作站的处理速度、内外存容量、工作性能已接近或达到小型甚

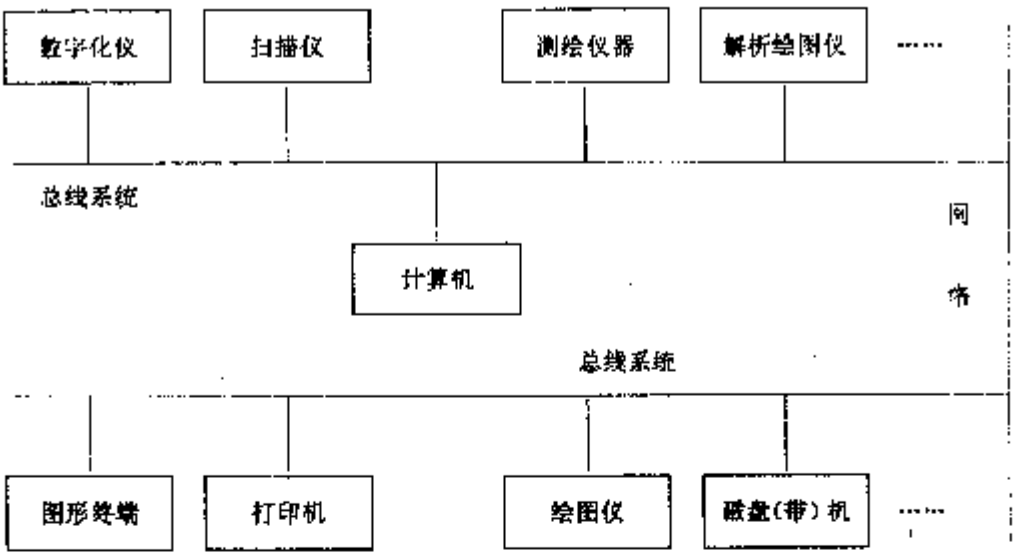


图 2-1　GIS 的主要硬件设备

至中型机的水平，特别是工作站上还配有大屏幕、高分辨率图形终端，很适用于 GIS 中的图形图像处理。而工作站体积、价格与中小型机相比，则大大地缩小和降低。因此，从最佳性能价格比考虑，采用工作站作为 GIS 主流机将成为趋势。但是也应该注意到，随着微机性能迅速提高，如最近推出的 PENTIUM PRO 微机的时钟频率可达到 200 MB，内存可达到 32 MB 以上，也可配置可读写光盘，虽然数据传输率和图形处理速度还较低，但与工作站之间的界限已经不很明显。因此，对于一个局部小区域，或某种专题应用，微机也是很有可为的。

工作站在 GIS 中的应用将改变以往大型计算机集中式配置的模式，而采用工作站的网络配置模式。将许多较小的、专门化的计算机联成一个网络，在网络内，不同程序可同时互相独立地在不同的计算机上运行，而数据和资源又可以供网络上多用户共享，从而大大地扩展了 GIS 功能的应用。

二、数据存储设备

数据存储设备包括：软盘、硬盘、磁带、光盘等及其相应的驱动设备。

磁盘是一般计算机必备的存储装置，分为硬盘和软盘两种。硬盘的存取速度和存储容量比软盘大得多，现在一般微机用的硬盘容量从几十 MB 到上千 MB 以上；而工作站的硬盘已超过 4 GB（1GB＝1000 MB）的容量。一般的 5 英寸* 高密软盘容量为 1.2 MB，3.5 英寸软盘可达 1.44 MB 到 2 MB。硬、软盘的驱动器以往通常安装在主机的机箱内，现在也推出了外接硬盘。此外，激光技术的应用已使计算机的存储容量有了较大突破，光盘的出现使存储容量大大增加了。目前市场上已见到的光盘，可读光盘和可擦写光盘，容量可

* 英寸（in）为我国现行国家计量标准应废除的单位，1 in＝2.54 cm，下同。

达到 1 GB 以上；光盘驱动器（CD-ROM DRIVER）的数据传送速度也已接近硬盘。

GIS 是数据密集型系统，例如：一张中等密度的地形图一般需要 1～2 MB 的存储容量。所以，如果没有高密度存储介质和较快的传输速度，GIS 就很难充分发挥其应有的性能。

应该说明的是，由于 GIS 的许多操作并不经常涉及特大数量的运算，但需要十分频繁地访问大型而复杂的数据库。因此这种对磁盘访问的频繁性和复杂性就要求 GIS 的磁盘存储器的速度和容量都较大。这一点在评定 GIS 系统效率高低标准时，甚至比实际运算速度还重要。

三、数据输入设备

GIS 基本的输入设备除键盘、鼠标和通讯端口外，还包括：数字化仪、扫描仪、解析和数字摄影测量仪以及全站型速测仪、GPS 接收机等其它测量仪器。

数字化仪是 GIS 中最基本的一种输入装置。使用数字化仪可以对各种已有线划图、地形图进行数字化。现今通常采用的跟踪式数字化仪对地图数字化产生矢量结构的数字地图，其输入数据的精度受数字化仪本身分辨率、数值记录字长、操作者经验和技能等多因素影响。常见的数字化仪的幅面有 A3、A1、A0 和超 A0 等规格，用于制图的数字化仪一般是 A0 规格的，幅面尺寸为 914 mm×1219 mm。当然对于一般应用而言，A1 幅面的数字化仪也能满足要求，因为它能包含一幅 0.5 m×0.5 m 尺寸的常规地形图。A3 幅面的数字化仪通常用于试验或培训工作。

扫描仪也是 GIS 图形、图像数据输入的一种重要工具。按不同的划分方法，可以对扫描仪进行多种分类。例如，按照辐射分辨率划分，有二值扫描仪、灰度值扫描仪和彩色扫描仪。二值扫描仪每个像元占一位，取值 0 或者 1，用于线划图和文字的扫描；灰度值扫描仪每个像元占 8 位，分为 256 个灰度级，可扫描图形、影像和文字等；彩色扫描仪是通过滤光镜将彩色图件分解为红、绿、蓝三个波段，每个波段占 8 位，再加上黑白图像，共占 32 位。按照扫描仪结构划分，又分为滚筒式、平台式和 CCD 摄像扫描仪。滚筒式扫描仪将扫描图件装在圆柱形滚筒上，然后使扫描头沿 x 轴方向移动，滚筒沿 y 轴方向转动，从而实现对图件的扫描；平台式扫描仪是将图件平放在扫描平台上，扫描头可以在该平面内沿 x 和 y 两个方向移动对图件扫描；CCD 摄像扫描仪是一种装有电荷耦合器件（CCD）且具有高分辨率镜头的电视摄像扫描装置，通常将低功率激光源与它组合在一起，扫描过程中 CCD 将摄取的影像分解成离散的像元。按扫描方式划分，又可分为按栅格数据形式扫描的栅格扫描仪和按线划扫描并直接产生矢量数据的矢量扫描仪。通常，扫描仪的分辨率是用像元大小（10 μm～100 μm）或每英寸的点数（dpi）表示，一般扫描仪的分辨率要求在 400dpi 以上。

解析和数字摄影测量仪器也可以作为 GIS 的输入设备。解析测图仪可用于数字测图，在机械绘图桌上装上电子机械 x、y、z 记录器，可以得到三维坐标。第三代的解析测图仪就是为 GIS 图形数据采集而设计的，这类仪器不仅能记录三维坐标，还能通过连接的微机处理比例尺变形和其它制图变形，处理后的数据可以直接记录到磁盘、磁带或活动硬盘上。

全站型电子速测仪是一种可以自动地以数字形式记录测量数据的仪器。

GPS 接收机，即全球定位系统接收机，可以通过卫星定位，将地面静止或移动目标，如

勘测车船、测量仪器车辆的空间坐标实时地输入计算机。

四、数据输出设备

GIS 的输出设备主要有图形图像显示器、矢量式绘图仪、栅格式绘图仪、行式打印机、点阵式或喷墨打印机和彩色喷墨绘图仪等。

通常，GIS 有许多输出并不要求硬拷贝，而只是在屏幕上显示出图形、图像。因此，图形显示器在数据输入编辑和检索等阶段都要用到。高分辨率大屏幕的图形显示终端，尤其是近期出现的立体显示终端，使 GIS 的图形输出取得良好效果。绘图机是 GIS 的主要图形输出设备。矢量绘图机有滚筒式和平台式两种，其图形输出质量主要取决于笔控马达的步进量。对制图而言，以不超过 0.05 mm 为好。绘图的灵活性和绘图速度则很大程度上取决于绘图软件的功能。

§2-2 GIS 的软件模块

GIS 的软件是指 GIS 运行所必需的各种程序，它们构成 GIS 的核心部分，关系到 GIS 的功能。这些软件通常是由两大部分组成的。一是计算机系统软件，它包括与计算机硬件有关的操作系统、汇编程序、系统库、编程语言、库程序等以及一些标准软件，如图形处理程序、数据库、Window 系统等；二是 GIS 系统软件和其它 GIS 应用软件，如 GIS 与用户的接口通讯软件、GIS 应用软件包和 GIS 基本功能软件包。按照 GIS 对数据进行采集、加工、存储管理、分析查询、显示再现和与用户接口，可将 GIS 软件系统中与用户有关的软件分为如下几种软件模块（图 2-2 所示）。

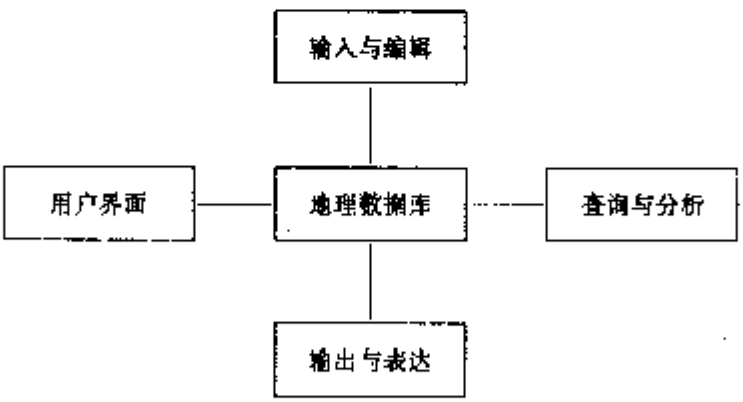


图 2-2 GIS 的主要软件模块

一、数据输入与编辑

通过各种数字化设备将现有地图、外业观测成果、航空像片、遥感数据、文本资料等转换成计算机兼容的数字形式的处理转换软件。也可以通过通讯或读取磁盘、磁带的方式录入已存在的数据。图 2-3 所示为针对不同仪器设备，数据输入软件配备示意图。

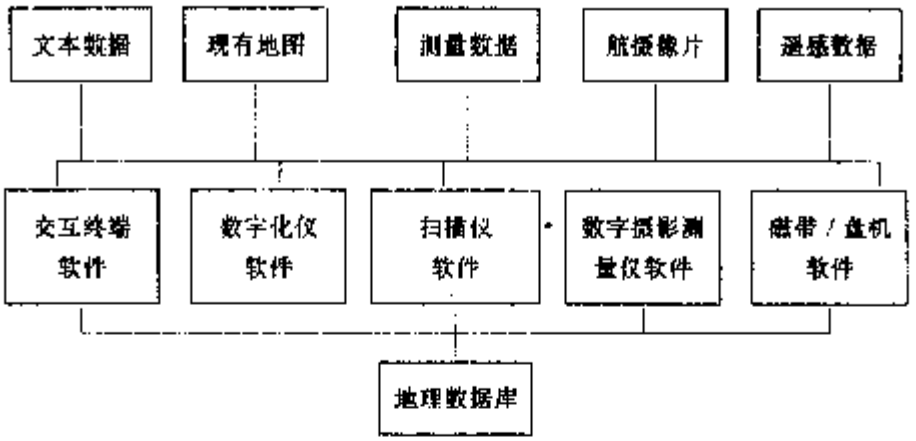


图 2-3 数据输入子系统

对原始输入数据进行观察、统计分析和逻辑分析，检查数据中存在的各种错误，通过编辑修改予以改正。另一方面，通过编辑还要对图形进行修饰，建立拓扑关系以及组合复杂地物。其功能如图 2-4 所示：

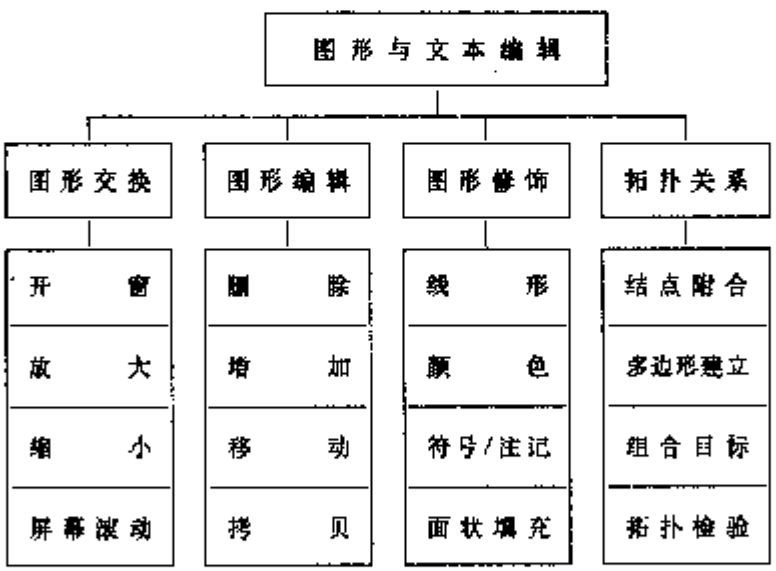


图 2-4 编辑软件模块

二、地理数据库管理系统

数据存储和管理内容包括空间景物，如地物的点、线、面、体的位置，空间关系以及它们的地理意义如何结构和组织，使其便于计算机处理和系统用户理解。此外，还要处理诸如数据格式的选择和转换、数据压缩编码、数据的联接、查询、提取等内容。

目前，通用数据库的模型一般采用层次模型、网状模型或关系模型。如前节所述，由于一般数据库在空间实体定义和空间查询、空间分析方面存在缺陷，致使现在许多 GIS 软

件在处理空间数据时采用拓扑数据模型，仅对属性数据采用关系模型。近来在计算机技术中出现的所谓“面向对象”数据模型，既可以表达图形数据又可以有效表达属性数据，因此，面向对象数据模型可能成为GIS发展的一个方向。

三、空间查询与空间分析

通常指对单幅或多幅专题图件及其属性数据进行分析运算和指标量测。在这种操作中，以原始图为输入，而查询和分析结果则是以原始图经过空间操作后生成的新图件来表示，在空间定位上仍与原始图一致。因此，也可将其称为空间函数变换。例如查询某县所有地块的土壤类型；或查询某铁路沿线周围2公里的居民点就要具有拓扑叠加分析功能和建立缓冲区分析功能。空间指标的量测包括对面积、长度、体积、空间方位、空间变换等进行计算。

四、用户接口

该软件模块是用于接收用户的指令和程序。系统通过菜单和命令解释方式接收、解释并运行完成用户要求任务的系统程序。用户自行编制的应用程序可以是调用系统功能的批处理程序，也可以是处理系统数据的分析程序。用户接口模块可接纳用户开发的应用程序，并提供系统与用户程序的数据接口。该模块还随时间向用户提供系统运行信息和系统操作帮助信息，从而使GIS成为人机交互的开放式系统。

五、数据输出与表示

输出与表示是指将GIS内的原始数据或经过系统分析和转换后重新组织的数据以某种用户可以理解的方式提交给用户，可以用地图、表格、图表、文字、数字、影像等多种形式表达，也可以将输出结果记录于磁存储介质设备或通过通讯线路传输到用户的其它计算机系统。

§2-3 GIS地理空间数据和GIS组织管理人员

一、GIS地理空间数据

地理空间数据是GIS的研究、作用对象，是指以地球表面空间位置为参照的自然、社会和人文经济景观数据。它们可以通过图形、图像、文字、数字、表格等形式表示，可以通过各种数字化设备，以及键盘机、磁带机或其它系统的通讯接口输入GIS。GIS正是通过对这些数据的采集、管理、分析，提取其实质性信息内容构成信息模型，根据用户的要求再现出现实世界。

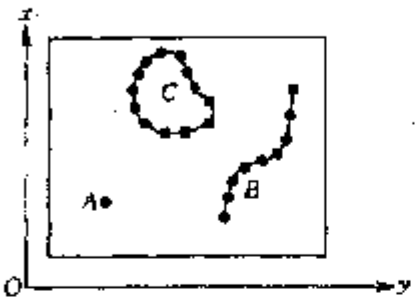
地理空间数据主要包括空间位置、拓扑关系和属性三个相互联系的方面。

1. 在某参考系（如已知坐标系）中的空间位置

几何坐标用来标识空间目标在自然界或包含某个区域的地图中的空间位置。例如用经纬度、空间直角坐标、平面直角坐标、极坐标等方式表示。在用各种数字化设备将其输入GIS中，通常应顾及参考系的一致性和转换关系。

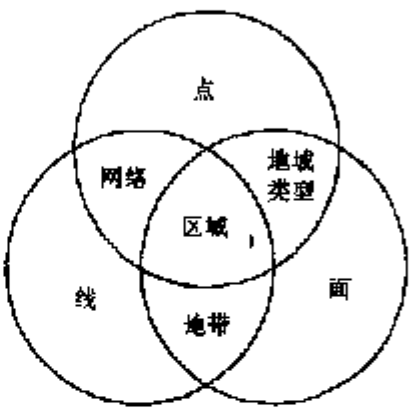
地理实体或景观表现可归纳为点、线、面等几种空间类型。点具有几何确定位置，可以用坐标对（x，y）表示；线具有一定的长度和走向，表示线状地物或点之间的地理联系，

可以用坐标对序列 (x_1, y_1)，(x_2, y_2)，…，(x_n, y_n) 记录；面具有确定的范围和形态，表示空间连续分布的地理景观或作用范围，通常可以用坐标对序列表示的闭合线段构成。图 2-5 (*a*) 所示为仅用坐标数据表示的无拓扑结构的点、线、面矢量数据。点和线的结合可组成网络；线和面结合可组成地带；点和面结合组成地域类型；点、线、面结合组成区域，如图 2-5 (*b*) 所示。



特性	数据	位置
点	*A*	(x_a, y_a)
线	*B*	$(x_{b1}, y_{b1})(x_{b2}, y_{b2})\cdots$
面	*C*	$(x_{c1}, y_{c1})\cdots\cdots(x_{cn}, y_{cn})$

(*a*) 无拓扑结构的矢量数据组织



(*b*) 点、线、面实体的空间组合

图 2-5

2. 空间实体的拓扑关系

无拓扑关系的数据结构虽然简单，但是在实际应用中因难于建立空间关系进行空间分析而受到限制。因此，现在许多 GIS 系统都尽可能地采用拓扑结构数据模型，建立地理数据的拓扑关系。

拓扑关系是指表示地理实体间的空间相关性，即确定点、线、面等实体之间的空间联系。如结点与弧段形成的枢纽关系（图 2-6 (*a*) 所示），弧与面之间的构成关系（如图 2-6 (*b*)）以及面与“岛”或内部点形成的包含关系（图 2-6 (*c*)）等。

利用空间实体之间的拓扑关系可以解译出更多的地理信息。例如按照城市街道图找出从机场到旅馆的路线，可以识别出彼此相邻的两个特征物，并确定它们位于哪条街道上；可以通过识别道路上的连线确定这些连线所封闭的区域以及识别相邻的区域来解释这些相互关系等。所以采用拓扑关系作为数学工具来描述地理实体之间的空间相关性是十分必要的。空间拓扑关系涉及到对于地理空间数据的编码、录入、格式转换、存储管理、查询检索和模型分析等诸多环节。但是涉及到用于存储地图特征时，可简单地归纳为三种空间关系，即

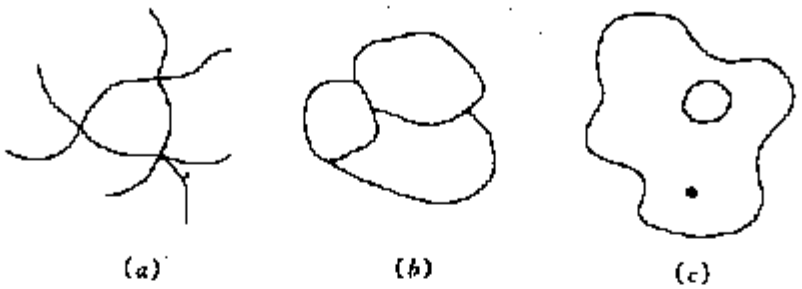


图 2-6　空间实体的拓扑关系

区域定义、邻接性和连通性。下面分别给以说明。

(1) 区域定义

用拓扑结构定义区域是采用一组封闭线来定义多边形，而不必列出封闭线上的所有点的坐标。如图 2-7 所示，图中的阴影区是由一组线定义的，即线 1、3、4、5 加上由线 2 定义的多边形“岛屿”。多边形的拓扑定义是一组线，并作为一个简单的表存贮在计算机中.

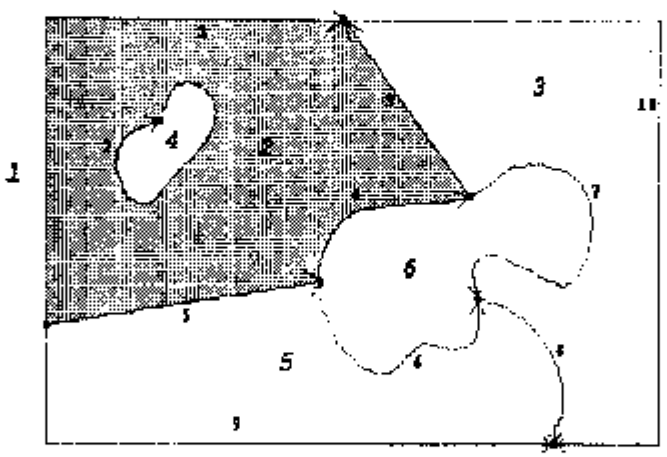


图 2-7　区域定义

应该说明的是，这些定义了一个封闭区域的线段是有序排列的，然后排列组成岛屿的线段。通常岛屿是由一条或多条线段组成的，一个多边形可能包括多个多边形岛屿。也就是说，图上的多边形并不都是同一等级的多边形，而可能存在多边形内嵌套小多边形（次级）的情况。一般采用所谓“包含性”来描述多边形之间的嵌套关系。

如图 2-7 所示，阴影区可通过定义该区的一组线段存贮在计算机中，而这些线段又可作为定义每条线段的坐标的指针。

这种结构的优点是，对于每条线段来说，只需列出一次线段的坐标即可，这不同于那种用封闭线上的每个点的坐标来定义多边形的方法，从而大大减少了所需要的存贮空间。其次，这种结构还克服了定义多边形的坐标数量受到限制的缺点。

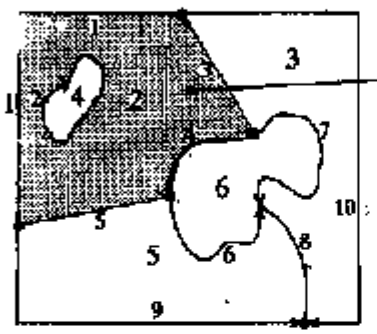
这种拓扑数据结构通常被称为线段—结点。每条线被称为一个线段，每个结点被认为

是线段相交的位置。

(2) 邻接性

另一种可以用拓扑描述的空间关系是邻接性，即确定多边形相互之间的邻接关系。邻接性对于许多应用来说都是非常重要的。例如，野生生物学家可能想确定一块与湖泊相邻的土地覆盖区，用于对栖息环境做出评价；城市规划人员可能要在两个相邻地区之间确定一下潜在的地区冲突如寻找与新发展的居民区相邻的工业区。

相邻的多边形可用它们公用的一组线段来定义，每个线段的左右两侧的信息可用于对多边形进行判别。例如在图 2-8 中，线段 4 的一侧是多边形 2，而另一侧则是多边形 6；每个线段的坐标列表确定了线段的方向，如图中箭头所示。



多边形/线段表

多边形号	线段号
2	1,3,4,5,2
3	3,10,8,7
4	2
5	5,6,8,9
6	4,7,6

线段坐标表

线段号	坐标 x, y
1	5,3 5,5 8,5
2	4,4 4,5 …
3	9,4 6,5
4	8,3 9,4…
5	…
6	

线段号	左多边形	右多边形
1	1	2
2	2	4
3	2	3
4	2	6
5	2	5
6	6	5
7	3	6
8	3	5
9	5	1
10	1	3

线段号	坐标 x, y
1	5,3 5,5 8,5
2	4,4 4,5 …
3	9,4 8,5
4	8,3 9,4…
5	…
6	…
7	…
8	…
9	…
10	…

图 2-8 邻接性的表示

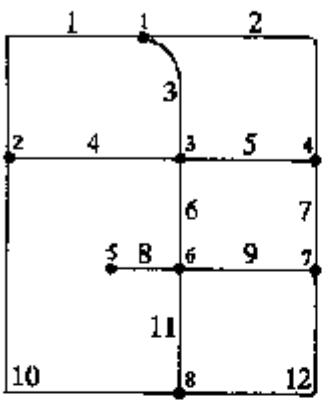
利用这种显示线段方向的表示法，可以明确指出位于该线段左侧和右侧的多边形。对于所有线段的相邻区域，图 2-8 中的结构列表给出了其邻接性表示。

(3) 连通性

连通性的拓扑关系与邻接性相似，连通性指的是对线段连接的判别。这种关系对于路径搜寻以及其它网络应用，如最佳路径计算和全网络流程分析都是非常重要的。

连通性可以用在每个结点上汇集的线段的列表来表示。例如图 2-9 所示，结点 2 上汇集了三个线段，即线段 1，4 和 10。每个线段的起点和终点都是结点。对于每个线段都可以列出它的起始结点和终止结点。例如，在线段坐标的列表中，如 $x1$，$y1$；$x2$，$y2$；……；x

10，y 10。在这个列表中第一个结点和最后一个结点都是标有序号的结点。如果在另一条线段上具有与上面结点之一相同的 x，y 坐标，则该线段上的那个结点将具有相同的结点序号。在图 2-9 中，线段 4 的起始结点为 2，终止结点为 3，另外还有三条线段汇集在结点 3 上。



线段—结点表

线段号	起始结点	终止结点
1	2	1
2	1	4
3	1	3
4	2	3
5	4	3
6	3	6
7	4	7
8	5	6
9	6	7
10	2	8
11	6	8
12	8	7

线段坐标表

线段号	坐标 x, y
1	5,3 5,5 8,5
2	8,5 11,5 11,3
3	8,5 9,4 9,3
4	5,3 9,3
5	11,3 9,3
6	9,3 9,1
7	……
8	
9	
10	
11	
12	

图 2-9　连通性的表示

在网络分析中应用连通性可以进行路径确定。这如同从多条路线中如何选择一条行车路线，或沿着城市街道从一个地方到达另一个地方去一样，当你到达一个交叉路口时，有好几个方向可以选择。这种情况就可以用在每一结点上有多条线段汇集在该结点上的模型来描述。

线段—结点拓扑结构还有一些其它重要性质。例如，每个线段的起始结点与终止结点表示了线段的方向，这样使得判别每一线段的左侧和右侧更加容易。另外，起始结点是每个线段的起始坐标，终止结点是每个线段的终止坐标，这就是说，结点坐标不需要在两个地方存贮（一次是作为线段坐标，另一次作为结点坐标存贮），从而大大减少了数据冗余。

3. 与几何位置无关的属性

这种属性即常说的非几何属性或简称属性（Attribute），是与地理实体相联系的地理变量或地理意义。属性通常分为定性和定量两种。定性数据包括名称、类型、特性等，定量数据包括数量和等级。比如，定性描述的属性有：岩石类型、土壤种类、土地利用、路面材料、行政区划等；定量描述的属性有土地面积、道路长度、土地等级、人口数量、降雨量、水土流失等。非几何属性数据一般是经过抽象的概念以及通过分类、命名、量算、统计得到的。任何一个地理实体至少拥有一个属性，而 GIS 的分析、检索和表示也主要是通过属性的操作运算实现的。因此属性数据的分类、量算对 GIS 系统功能有重要影响。

举例来说，对于一个道路实体可以将其数字化为一个矢量表示的线实体或者是栅格表示的连续像元链，并可用一定的颜色、符号或数据定位等，作为 GIS 的空间数据表示出来；道路类型可以按常规制图符号表示；而用户还希望知道的，如道路宽度、路面材料、建筑

日期、施工方法、特殊交通规则、车流量等则是用属性数据描述。显然，这些属性数据是与道路这一空间实体有关的，对属性数据进行有效地存储和处理也是十分必要的。通常的做法是给予每个属性数据赋予一个公共识别符并使其与空间实体连接起来。

二、GIS 的管理人员

人是GIS 中重要构成因素。GIS 不同于一幅地图，而是一个动态的地学模型。所以，仅有系统的软硬件和数据还不能构成完整的地理信息系统，需要人进行系统组织、管理、维护和数据更新、扩充完善、应用开发，并灵活采用地理分析模型提取多种信息，为研究和决策服务。

显而易见，任何先进的现代技术的引进和开发应用，都必须拥有掌握该技术的人才。地理信息系统的建立和应用是实现地理分析、环境分析、土地及城市规划与管理，合理利用国有资源的先进现代技术手段。要使 GIS 有效地运行，必须对人力资源进行投资，使有关人员得到全面的技术培训。另一方面还必须促使从事 GIS 工作的人们使用新的思维方法和使用与以往传统手工操作不同的方式进行实际工作。

GIS 的技术人才是GIS 系统建设不可缺少的。他们应当具备很好的测绘、遥感、地理、计算机技术和应用科学的知识基础，处理各类技术问题的经验，果断的判断决策能力和较强的组织指挥才能。如果一个单位或部门没有既懂本行业专门知识和技术，又懂得 GIS 知识和技术的熟练的科技人员，那么要充分发挥 GIS 的功能就必须对现有人员进行有关培训或调入新的技术人员。同样地，任何部门也都还存在对现有科技人员的知识更新和再教育问题以及裁减不适合在本行业工作的人员，进行合理的技术人才流动之类的问题。总之，拥有 GIS 的理论知识和操作技能的人员是开展 GIS 工作的先决条件之一。

通常可以将从事 GIS 工作的人员分成两类，即一般技术人员和高级技术人员。

对一般技术人员而言，并不要求对 GIS 的理论和方法有精深研究，在应用 GIS 系统时也不必一定要知道其工作原理，而只要求知道如何使用和操作 GIS 系统即可。通常，他们的日常任务是进行数据输入并保证把结果输出来，操作计算机以及对打印或绘制的输入/出结果的检查工作等。

高级技术人员又可分成四类：业务管理人员、业务联络员、技术人员和科学研究人员等。业务管理员主要处理 GIS 的日常运行，进行系统与其他单位或部门之间的对话和业务商谈等；业务联络员的任务是建立并维护与用户的联系，负责对外介绍 GIS 的功能，使用户能放心地将需要完成的任务交给 GIS 系统拥有者处理；技术人员是指计算机程序编制、计算机制图、数据分析处理模式、系统开发和精通 GIS 技术工作的人员；科研人员包括环境科学家，以 GIS 进行科研和应用的科学家，计算机科学家等。这些高级技术人员具有建立一个新的 GIS 系统的能力或对现有系统进行大规模开发的能力。

综上所述，为充分发挥 GIS 系统的应用功能，还应该认真考虑部门内部的组织状况以及 GIS 与部门内各单位的相互关系，使建立的 GIS 既能适应多方面服务的要求，又能与现有的计算机及其他设备相互补充和配套。如果仅是买一些计算机硬件设备和一些 GIS 有关软件，由几个热心人管理和使用，然后就想立即得到收益，是远远不能如愿的。理想的 GIS 应适合于不同类型的数据采集方法和数据处理方法，适合于生产多种 GIS 产品，就是说整

个组织必须全盘规划数年内希望完成的各种任务。

§2-4 系统运行环境

地理信息系统不同于一般的管理信息系统，它是一个集空间型、管理型于一体的数据存贮量大、功能繁多、处理复杂、规模宏大的系统。根据系统的特点，必须选择一个较为适合的运行环境来满足系统的功能要求。系统运行环境一旦被确定，地理信息系统将在此环境下进行开发、建设和运行，并针对系统的近期目标、中期目标和远期目标逐步完善系统的各项功能。所以，系统运行环境选择的正确与否，将意味着GIS系统建设的成败。

一、系统运行环境确定的依据

系统运行环境的确定应该遵循以下基本原则：

——最大限度地满足用户的工作需求；

——在保证实现系统功能的前提下，尽可能地降低资金投入；

——考虑一定时间内技术的相对先进性以及软硬件之间的相互兼容性。

在具体进行选择时，还需要根据以下几点要求制定选择条件：

(1) 满足大数据量处理和存贮的要求；

(2) 满足图形的采集精度，图形的处理（编辑）、显示、查询速度，图形多种输出的要求；

(3) 满足系统的空间数据多种分析功能的要求；

(4) 满足系统的地形分析功能的要求；

(5) 满足数据库管理系统的功能要求；

(6) 为系统提供先进的开发环境和开发工具；

(7) 操作系统和用户界面的要求；

(8) 网络管理和远程通讯的要求；

(9) 满足系统的开放性和系统的可升级性的要求。

二、运行环境的选择

系统的运行环境包括三个内容，即系统的硬件运行环境、系统的软件运行环境和系统的网络运行环境。通过系统分析、子系统和数据库的划分，明确了各级系统的功能，各级系统的运行环境主要是依据系统的功能来确定的。

1. 硬件运行环境

从计算机技术的发展来看，客户机/服务器（CLIENT/SERVER）体系结构是90年代的主要潮流。这种结构的工作方式是：处理控制逻辑、显示和数据处理及用户交互的部分在网络上的CLIENT（客户机）端完成，这些客户机可以是网络上的工作站、微机或图形终端；控制数据存取安全、完整、容错和并发的部分在中心共享的计算机——SERVER（服务器）上完成，服务器可以是微机、工作站服务器、小型机甚至大型机。这种体系结构的优点是：

(1) 处理能力强。通过网络分布工作任务，可提高整体事务处理吞吐量和由此环境支持

的同时用户数量。

(2) 扩充能力强。当系统须增加新的应用子系统时，只须在网络上添加客户机，而不必对服务器做任何改动。

(3) 保护硬件投资。这种体系结构提供混合和匹配多厂家硬件产品的灵活性。所以，现有硬件可用在网络上或做客户机或做服务器，从而保护了用户现有的硬件投资。

因此，对于大区域应用的GIS系统而言，应选择客户机/服务器的体系结构。这种结构的选择应根据系统的功能选用不同档次、数量的工作站、服务器、微机或图形终端。

对于一般地区，系统可以采用局部网络方式，也可采用所谓WINDOW工作群组方式(WINDOW FOR WORK-GROUP)，再低一级的小范围地区，如县级地区等，一般可采用微机局部网络甚至暂时可使用单机方式。

从总体上讲，GIS的硬件平台包括有图形输入、输出设备，数据存储、处理设备以及相关的辅助设备等。在选购时原则上还应该满足：

——性能/价格比较高。可维护性好，可靠性高；

——硬件的速度和容量方面能够满足系统和用户的要求，并且易于扩展；

——有较好的售后服务及技术实力。

2. 软件运行环境

软件运行环境的考虑包括三个方面：操作系统；数据库管理软件；GIS软件。

(1) 操作系统

在考虑操作系统时必须考虑其开放性。这里的开放性指的是系统在采用了一种操作系统后，能否很方便地在其上挂接各种必需的应用软件。在当前的技术条件下，操作系统的选择应与硬件的运行环境、GIS的数据库对操作系统的要求、网络环境有关。对于采用客户机/服务器结构的系统来说，目前多是采用基于UNIX操作系统，但局部网络环境下也可采用UNIX，DOS和WINDOWS的混合使用；对于微机局部网络的情况，也可采用UNIX和WINDOWS，或可采用DOS和WINDOWS。

根据当前的计算机市场资料介绍：WINDOWS NT的出现为操作系统带来了新的选择，WINDOWS工作群组方式可以代替网络功能，实质上是把网络操作提升到WINDOWS级。

(2) 数据库管理软件

在地理信息系统建设中，一是针对各子系统的业务过程进行设计，二是设计共享的数据库；在运行中，一是子系统执行的业务过程，二是对数据库的管理和建设。所以，数据库系统在地理信息系统中占有非常重要的位置。

GIS系统包含大量的图形数据和属性、统计数据。通常每一级系统中都有十几个数据库，图形数据一般可采用GIS软件进行管理；但大量的属性和统计数据的管理，必须采用功能较强的数据库管理软件，这样才能提高系统的运行效率和各子系统充分利用共享的数据库资源。特别是大区域范围的GIS系统，是以统计数据为主要内容的，而它的数据又来自于下面各级系统，不但数据量大，数据格式也可能不完全统一。因此，对于大区域范围的地理信息系统在考虑数据库管理软件时必须注意如下原则：

——与系统选用 GIS 的连接问题；

——是否能在多种主机、多种操作系统上运行；

——是否提供了先进的开发工具和具有良好的开发环境；

——是否基于 SQL 语言；

——对多用户联机事务处理的能力；

——能否支持多种网络协议；

——能否支持多媒体等大数据量；

——能否支持汉字系统；

——对众多数据库快速的联机处理能力；

——能否满足与其它系统的消息传递；

——是否具有良好的售后服务和对用户的技术支持。

基于上述问题的考虑，大区域范围的 GIS 系统以及采用客户机/服务器结构体系的一般地区级系统应使用诸如 SYBASE、ORACLE、INFORMIX、INGRES 等大型关系型数据库系统；在基于 WINDOWS(或 DOS)环境下的系统，可以采用诸如 FOXPRO、PARADOX、FOXBASE、dBASE 等微机数据库。

在选用数据库管理软件时还须注意一个问题，即系统建设的初期，如果选用小型数据库管理软件，应该考虑将来采用大型数据库管理软件时，这两者之间的数据转换问题。

(3) GIS 软件

GIS 软件可以大致分为系统基础软件和二次开发软件。系统基础软件是指构成系统基础软件平台的数据录入软件、编辑软件、管理软件、分析软件和数据输出软件等；二次开发软件则是指在系统基础软件平台上针对不同用户、不同功能需求、不同管理和运作方式而进行的二次开发的软件。虽然系统基础软件决定了系统的主要功能，但是，若要充分发挥系统功能的作用，通常需要进行二次开发。二次开发软件制作的越好，对用户就越友好，系统功能的作用就越大，系统使用就越简单、越方便。因此也可以说，只有二次开发软件做得好的地理信息系统才会是一个成功的实用的地理信息系统。

还应说明的是，GIS 系统在建设和运行过程中无时无刻不与图形打交道，各种图件是地理信息系统的基础数据之一，因此，GIS 系统的基础软件必须具备对图形的输入与输出以及对空间数据的分析和管理功能才能来完成 GIS 系统要求的图形处理工作。

当前的 GIS 软件已有几十种甚至上百种，地理信息系统采用的 GIS 软件主要考虑如下一些原则：

——所选择的 GIS 软件与所选用的主机之间的兼容性；

——国内有较强的技术力量支持，并且有一个较大的用户群体；

——GIS 与数据库的连接问题；

——根据工作需要是采用全套的 GIS 还是采用 GIS 的核心部分；

——所选择的 GIS 软件是否具有数据输入与转换、图形与文本编辑、数据存贮与管理、空间查询与空间分析、数据输出与表达这五大基本功能以及系统的运行效率；

——所选择的 GIS 软件是否提供了先进的开发工具和良好的开发环境；

——能否满足与其它系统的消息传递；

——能否支持汉字系统；

——版本的升级和技术支持问题。

3. 网络环境

网络环境包括各级系统的局域网环境和对外网络通信的广域网环境。

局域网（Local Area Network，简称 LAN）是指小区域内，如一座办公大楼、一个校园或一个工厂等的计算机互连网络。由于其通信距离较短，且采用共享通信信道和总线以及采用环形或树形拓扑结构，因而具有较高的数据传输速率和较低的误码率。局域网又可根据其传输带宽进一步分为所谓“基带局域网”和“宽带局域网”。前者是指其传输介质中只存在一条信道，即介质的频带全部为该信道所占有，且介质中传输的是数字信号，其数据传输率可达到 10～20 Mbps，一般可以采用以太网、剑桥网等结构；后者是指其传输介质中存在多条信道，即某频带可以为各信道所共有，介质中可以传输数字信号和模拟信号，从而能够实现多媒体在同一网络中的传送，如 PC 网、王安网等。

广域网（Wide Area Network，简称 WAN）又称为远程网，其特点是网络分布范围广，不受地区限制，范围可以覆盖一个或几个省、全国乃至全世界。例如，被人们称为全球信息资源网的 INTERNET 网络，目前联系着 150 多个国家和地区的近 400 万台计算机，入网户已达到 3200 万以上，用户不仅可以使用电子信箱与网上任何用户交换信息，还可以跨越地区甚至国界使用远程计算机的资源，查询网上的各种数据库内容以及获取希望得到的各种资料。广域网通常是租用公用通信网，如电话通道和卫星信道，接入方式可以采用专线（DDN）和/或拨号（电话网、ISDN 网）的方式，用有线和/或无线（卫星小站、微波等）进入国家公用数据网，把各级地理信息处理中心连接起来。同时，各个地理信息处理中心也可以通过公用数据网向社会提供信息服务。但是，由于广域网的布局不规则，使得网络的通信控制较之局域网复杂，并且传输速率通常也不超过 64 kbps，一般情况下，要求网上的任何用户都必须遵守当局所制定的各种标准和规程。

三、系统的技术标准

为保证 GIS 中的各个数据库和各子系统数据分类、编码及数据文件命名的系统性、唯一性，本系统与其它信息系统的联网以及实现系统的相互兼容性和信息共享，地理信息系统的设计必须充分考虑到工程的技术标准问题，对规范化和标准化原则应该予以高度重视。为此，在已有国家标准、部级标准、行业标准和地方标准的情况下，应严格遵循这些标准。此外，还应根据本系统工程建设的需要制定必要的标准、规则和规定。在 GIS 中，这种标准可分为如下几类：

——数据标准

——文档标准

——软件标准

——系统运行标准

(1) 数据标准包括：空间定位标准；数据分类标准；编码体系与代码标准；数据字典；各数据库与文件命名标准；汉字与符号标准；数据格式与交换标准；数据质量标准；数据

处理标准等。

(2) 文档标准包括：可行性分析报告；总体设计方案；数据规范化、标准化技术方案；子系统用户需求分析报告；子系统详细设计说明；用户使用手册；数据库建库作业规程、技术规定、验收标准；系统安装手册；程序开发日志等。

(3) 软件标准包括：用户界面；数据结构；数据模型；数据操作和分析；数据库建立；数据库管理；数据显示和产品生成；接口设计；汉化程度等。

(4) 系统运行标准包括：系统效率，系统利用率，操作的方便性、灵活性；安全保密性；数据准确性；系统可靠性；系统的可扩充性；系统的可维护性等。

第三章　GIS 数据结构

§3-1　GIS 的数据

一、GIS 的数据

地理信息系统所处理的数据与一般事务性信息系统如银行管理系统、图书检索系统不同。GIS 的数据处理不仅包括所研究对象的属性关系，还包括研究对象的空间位置以及空间拓扑关系等信息，数据量大，结构复杂。因此，人们对 GIS 中的数据结构和数据模型进行了大量的研究，并发展了一整套空间数据处理的算法。

1. 数据结构的概念

数据结构是指数据的组织形式，可以分为抽象数据结构（或称逻辑结构）和数据存贮结构（或称物理结构）来进行研究。所谓抽象数据结构是指人们仅从概念上描绘数据之间的排列和联系，而并不涉及数据和具体程序管理细节。这种结构通常具有普遍性并且宜于进行理论研究。数据存贮结构则是为实现某一抽象数据结构而具体设计的数据存贮管理方式，是依照任务的不同，软件系统和设计者的不同而改变的，具有一定的特殊性，是前者的一个具体实现。

地理空间数据在 GIS 中的流向可以认为经历了四个阶段，用户认识的数据结构输入 GIS 系统后转换成为 GIS 空间数据结构，然后，为有效地进行数据管理，将其转化为数据库结构，最后按某种特定程式以硬件结构写入存贮介质。上述流程即为数据的输入过程，而该流程的反向即构成 GIS 空间数据的输出，如图 3-1 所示。

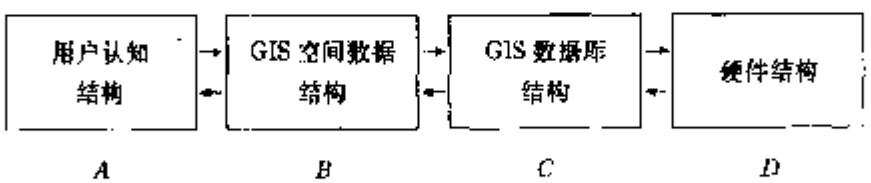


图 3-1　GIS 中的数据流向图

如前所述，GIS 的数据主要是二维或三维的空间型地理数据，包括空间位置、拓扑关系和属性三个方面的内容。GIS 中的空间地理数据结构就是数据经过上图的转化过程在系统内的不同组织、转换的编码形式。上图中的 A、B 和 C 三种属于抽象数据结构，是我们要研究的范围，D 则是数据存贮结构，属于计算机技术的范畴。

我们知道，地理空间实体可以抽象为点、线、面三种基本地形要素来表示它的位置、形状、大小、高低等。

——点（零维）：又称为元素或像元，是一个数据点，具有一对（x，y）坐标和至少一个属性，逻辑上不能再分。这里所谓逻辑上不能再分是指抽象的点而不是几何点。因为事实上抽象的点可以是实体线段或面块，对某个比例尺或图像分辨率而言，它们可以被抽象为以一对坐标表示的数据点，例如，在 Landsat 卫星影像上的一个点（像元）实际上为实地 79 m×79 m 的矩形区域；再如，在某些地图上城市、道路交叉点等被抽象为一个数据点等。

——线（一维）：是由一个（x，y）坐标对序列表示的具有相同属性的点的轨迹。线的形状决定坐标对序列的排列顺序，线上每个点有不多于二个邻点。地理实体、如河流、道路、地形线、公共设施走廊、区域边界、地质界线等均属线状地物，其特点是线上各点有相同的公共属性并至少存在一个属性。

——面（二维）：是以（x，y）坐标对的集合表示的具有相同属性的点的轨迹。面的形状不受各点坐标对排列顺序的影响。凡是面的内部点可以有多于三个的邻点，面内每个点应至少具有一个相同属性。土壤、植被、行政区划、岩石分类等地理实体属面状地物。

如果顾及平面位置与高程位置结合起来所构成的空间数据模型则还应考虑三维的体元素，作为点、线、面三个基本地形要素的一个外延。总之，从几何上讲，人们正是通过上述这些基本要素构成了对各种地理实体的认识结构。

2. 数据编码的原则

从前面的讨论中知道，地理信息系统数据库是一个空间数据库，它除具有一般事务性数据库中常见的数字、字符表示的数量和名称数据外，还必须管理具有空间定位和拓扑关系的地理空间特征数据。对于前种数据来说，因为它们是描述性数据且与几何空间参考系的选择无关，通常将其称为非几何属性数据，或常简称属性数据或属性；对于后者则是涉及地理实体在几何空间中的位置及其相互空间关系的数据，故称其为空间数据，或称为图形数据或几何数据。因此说 GIS 的数据是由空间数据和属性数据共同构成的。而 GIS 中对空间数据的处理及其对空间数据结构的研究也正是 GIS 有别于其它信息系统的特点之一。

二、GIS 属性数据结构及其编码

前面谈到的属性数据和空间数据是 GIS 中紧密联系的两部分内容。对属性数据管理的好坏与否直接影响到整个系统的运行，对于 GIS 这样一个复杂的系统，要求其处理具有较小的数据冗余，较好的灵活性以及较快的处理速度，这在很大程度上取决于对属性数据的管理和查询速度。因此属性数据管理对于 GIS 的运行效益有重要意义。

GIS 中属性数据主要是与各种专题地图有关的数量、类别、等级和描述性信息。除了通过统计、观测等直接产生的属性数据以外，还可以通过从地图图例中提取编码得到，也可以是由遥感影响分类提取后产生。此外，用于属性数据管理的属性数据库还可以管理非图形的各种文字记录、表格、说明等等。

在属性数据的支持下，图形数据（空间数据）就不再是仅仅具有几何意义的像元或图形，而是具有地理意义的地理实体。逻辑运算和地理分析、地理统计等空间操作都是通过属性数据与图形数据的结合与联系而得以实现的；此外，诸如景象编辑、景象分类、查询检索、模型应用、注记符号编辑、一致性检查、库结构定义等操作也都需要属性数据库的支持才能完成。

在GIS中，属性数据一般是通过相应的图素，如点（像元）、弧段、多边形的编号与图形建立联系的。属性数据的内容有时直接记录在图形数据中，有时则单独以某种结构存贮，通过指针或关键码与图形数据相关系。

属性数据结构记录的是与图形数据相关连的专题信息数据，通常采用数据库管理系统进行属性数据管理。为便于进行这种管理，可将各种形式的属性数据（如名称等级、数量、代码等）进行编码，使其成为计算机可接受的数值或字符形式，编码过程中，除了应顾及上述编码原则外，还应遵守如下三个原则：

(1) 管理效率高，具有较高的计算机录入、检校、存贮效率，具有较高的查找速度和查全率；

(2) 适用性好，符合专业化、标准化和系统化的要求，既要完整清晰地表示属性内容，又具有最小冗余，并与专业分类保持一致；

(3) 便于数据共事和扩展，可以满足设备和数据更新的要求。

对属性数据编码一般可包括如下三个方面的内容：

(1) 登记部分：用来标识属性数据的序号，可以是简单的连续编号，也可划分不同层次进行顺序编码；

(2) 分类部分：用来标识属性的地理特征，可以采用多位代码来反映多种特征；

(3) 控制部分：通过某些查错算法，对在编码录入和传递过程中的错误进行检查，特别是当属性数据量较大的情况下具有重要意义。

前面提到，在GIS中属性数据与图形数据有着密切的联系，在理论上应该对两者进行统一的管理和操作。但是目前具有这种功能的数据库管理系统还没有达到商品化的程度，现有GIS系统大都是将两者分别存贮：图形数据用文件系统管理，属性数据用数据库系统管理，然后建立起图形与属性的联接关系，从而才能实现两者的统一操作。

三、GIS空间数据结构的概念

这里所讨论的空间数据结构是指适合计算机系统存贮管理和处理的地学图形的逻辑结构。虽然这仅是对地理实体的空间排列方式和相互关系的抽象描述，但是它反映了对研究数据的一种理解和解释，并且构成了用户与GIS之间沟通信息的桥梁。

我们知道，人们根据自己认识的数据结构可以有效地辨识各种地理实体，而对于计算机而言，则需要用户或管理人员给以准确的指令、程序后才能正确处理和表示空间实体。不说明数据结构的数据是毫无用处的，不仅用户无法理解，计算机程序也无法正确的处理。另一方面，对同样的一套数据，如果按不同的数据结构去处理，得到的可能是截然不同的内容。因此，要正确有效地使用GIS就要充分理解GIS所采用的特定空间数据结构。GIS空间数据结构主要有两种类型：矢量结构和栅格结构。现代的一些地理信息系统结合了两种数据结构的特点，或采用混合数据结构或矢量栅格一体化的数据结构。如图3-2所示为用这两种数据结构来表示同一块由不同土壤结构构成的土地。

图3-2 (a) 中的土壤结构是由一组具有起终点坐标的线段和必要的连接指针构成。因为表示物体的线段有方向性，所以称之为矢量结构。线段端点的指针表明了这些线段应如何连接在一起才能形成相应地块。这种结构可以表述为：

地块→矢量组→连通性

图 3-2 (*b*) 中土壤结构是由格网中某一部分的像元或称栅格集合所构成，所以称之为栅格结构。在同一集合中的像元都具有同样的编码 "*a*" 或 "*b*" 或 "*c*" 等。实际上这些值本身并不一定显示出来，通常它们可能只代表某一符号或是某种颜色或是影像灰度。这种结构可以表述为：

地块→符号/颜色→像元

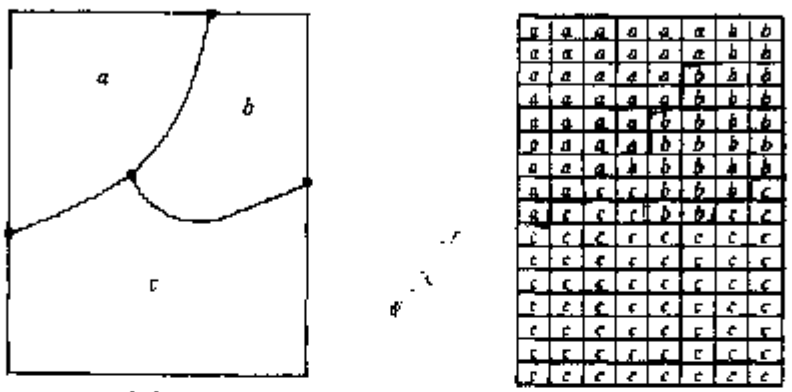


图 3-2 用矢量结构和栅格结构表示的例子

图 3-2 形象说明了两种数据结构表示方法。从直观上，我们可以看出：

(1) 矢量结构涉及的数据量少，所以占用的计算机存贮空间小；

(2) 对原物体的显示效果方面，矢量结构较栅格结构要好；

(3) 若要改变地块的形状或大小，矢量结构不仅要改变各矢量线段，而且还要改变各连接指针；而栅格结构则仅删添一些像元即可完成，所以栅格结构在数据更新、修改等方面优于矢量结构。

GIS 的空间数据结构在计算机系统中的实现通常是通过空间数据编码的方法进行的。由于 GIS 数据量极大，一般需要采用压缩数据的编码方式以减少数据冗余。一些常用的编码方式如图 3-3 所示。

矢量数据结构编码方式	栅格数据结构编码方式
—坐标序列编码	— 链码
—— 层次索引编码	—— 直接栅格编码
—拓扑结构编码	—— 游程长度编码
	—— 块码
	— 四叉树码

图 3-3 GIS 的常用编码方式

§3-2 矢量数据结构及其编码方式

一、矢量数据结构

矢量数据结构是通过记录坐标的方式，用点、线、面等基本要素尽可能精确地来表示各种地理实体。矢量数据表示的坐标空间是连续的，因此可以精确定义地理实体的任意位置、长度、面积等。如前所述，其显示精度较栅格结构高。事实上，它主要受数字化设备的精度和数值记录字长的限制。

在矢量结构中：对于点实体只是记录其在某特定坐标系下的坐标和属性代码。可以将其空间数据和属性数据结合在一起，将点的坐标直接作为点实体的两个属性值对待。在对线实体进行数字化时，就是用一系列足够短的直线段首尾相接表示一条曲线，当曲线被分割成多而短的线段后，这些小线段可以近似地看成直线段，而这条曲线也可以足够精确地由这些小直线段序列表示。在矢量结构中，只记录这些小线段的端点坐标，将曲线表示为一个坐标序列，坐标之间认为是以直线段相接，在某一精度范围内可以逼真地表示出各种形状的线状地物。对于面实体而言，在 GIS 中常用所谓“多边形”的概念来表述一个任意形状，并且边界完全闭合的空间区域。该闭合区域的边界线同前面介绍的线实体一样，也可以被看作是由一系列多而短的直线段组成，每个小线段作为这个区域的一条边，因此该区域也可看成是由这些边组成的多边形了。

矢量数据结构直接以几何空间坐标为基础，记录取样点坐标，可以将目标表示得精确无误。对于一个数字制图系统而言，按照这种简单的记录方式，再适当增加目标的注记名称、输出的线型和符号等，在矢量绘图仪上就可以得到精美的地图。另外，该结构还可以对复杂数据以最小的数据冗余进行存贮，相对于栅格结构来说，它还具有数据精度高，存贮空间小等特点，是一种高效的图形数据结构。

因为矢量结构的空间定位是根据坐标直接存贮的，而属性数据现在通常是建立表结构文件，用关系数据库管理。因此，矢量数据结构具有定位明显、属性隐含的特点。由于这一特点使其在某些方面有便利和独到之处。例如在计算长度、面积、形状和图形编辑、几何变换操作中，矢量结构有很高的效率和精度。但是同样由于这一特点，也使其图形运算的算法要较栅格结构复杂；在涉及叠加运算、邻域搜索等操作方面也不及栅格结构简单。

二、点和线实体的矢量编码

1. 点实体的矢量编码

点是空间上不能再分的地理实体，可以是抽象的也可以是具体的点。它的矢量编码简单直接，但是应将点实体的空间数据和属性数据记录完全，也就是说，除记录点坐标 (x, y) 以外，还需记录点的类型注记特征等属性。图 3-4 为点实体的矢量编码基本内容。

2. 线实体的矢量编码

线实体主要用来表示线状地物，如道路、河流、地形线等符号线和多边形边界。通常也称为“弧”。它的矢量编码相对来说也是比较简单的，其中包括如下内容：

(1) 唯一标识码：用来建立系统的排列序号。

(2) 线标识码：用来确定该线的类型。

(3) 起，终点：可以用点号或坐标表示。

(4) 坐标对序列：确定线的形状，在一定距离内，坐标对越多，则每个小线段越短，且与实体曲线越逼近。

(5) 显示信息：显示时采用的文本或符号，如线的虚实、粗细等。

(6) 其它非几何属性。

线与结点一起构成网络，从而产生线与线之间的连接问题，即拓扑关系中的连通性。因此还需要在线的数据结构中建立“指针”，指示其连接方向。除此以外，在结点上还应记录有交汇线的夹角，这样才能建立起正确的网络。在对城市管线网或道路交通网的研究中，上述工作是很重要的。

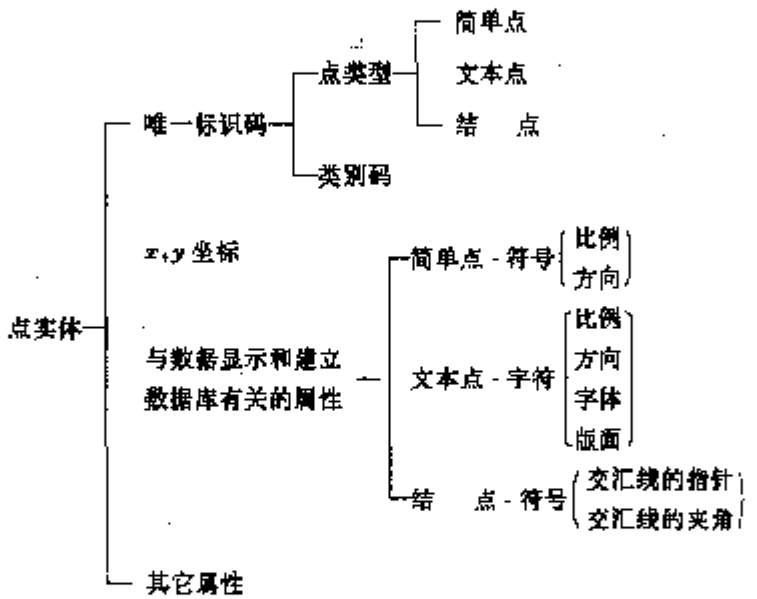


图 3-4　点实体的矢量结构表示

三、面实体的矢量编码

对于面实体，人们常采用多边形的概念。多边形数据是描述地理空间信息的最重要的一类数据。如行政区、土地类型、植被分布等具有名称属性和分类属性的地理实体均可用多边形来表示。多边形矢量编码不但要表示空间位置和属性，更为重要的是要能表达区域的拓扑性质，如相邻、连接、包含等关系。只有这样才能使其在进行专题性研究中得以方便地显示和操作。由于要表达的信息十分丰富，多边形的运算也较复杂，因此多边形矢量编码比起点和线实体的矢量编码要复杂得多，也更为重要。

多边形矢量编码除了为追求存贮效率要压缩数据以外，还有根据其自身特点提出的下列要求：

(1) 所表示的各多边形应有各自独立的形状、周长和面积等几何指标；

(2) 为便于进行空间分析，各多边形拓扑关系的记录方式应一致；

(3) 应能够明确表示区域层次，即多边形的包含关系，如“岛”在“湖”中而“湖”又在更大的“岛”上，形成岛—湖—岛的关系。

下面介绍三种多边形矢量编码方法：坐标序列法、层次索引法和拓扑结构法。

1. 坐标序列法

这是最简单的一种多边形矢量编码。因为多边形边界是由坐标集合及其说明信息组成，所以该法可看成是对线实体的一个简单扩展，即用定义多边形边界上的所有坐标来表示每个多边形。如图 3-5 所示多边形，其构成为下表 3-1。

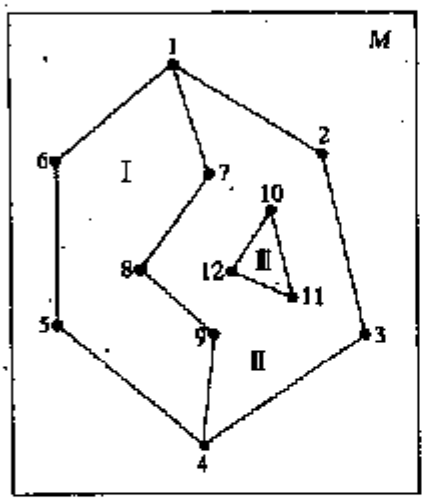


图 3-5 坐标序列法示例

表 3-1

多边形	坐标构成
I	(x1,y1),(x4,y4),(x5,y5),(x6,y6),(x7,y7),(x8,y8),(x9,y9)
II	(x1,y1),(y2,y2),(y3,y3),(x4,y4),(x7,y7),(x8,y8),(y9,y9),(x10,y10),(x11,y11),(x12,y12)
III	(x10,y10),x11,y11),(x12,y12)

该法文件结构简单，易于实现以多边形为单位进行的运算和显示。但是也有如下缺点：

(1) 多边形之间的公共边被数字化和存贮两次（如图 3-5 中的 7、8、9 三个点），由此产生数据冗余和边界不重合（由于数字化误差等因素造成）；

(2) 每个多边形自成体系而缺少相邻多边形信息，难以进行邻域处理。比如，合并同类时要消除公共边；

(3) “岛”只作为一个单个图形构造，没有与其外包多边形的联系；

(4) 不易对边界的拓扑关系进行检查。

如上所述，该法的应用有一定的局限性。

2. 层次索引法

该法是在坐标序列法的基础上进行了改进，它采用逐层索引的方法减少公共边的数据冗余以间接获得邻域信息。具体方法是，对多边

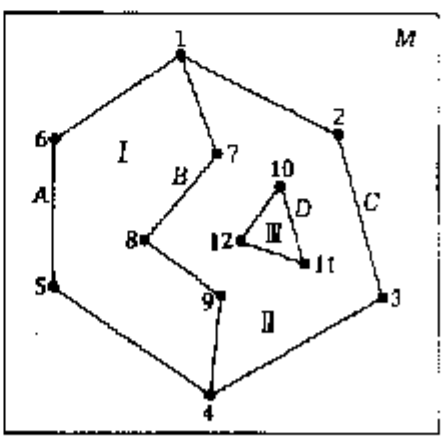


图 3-6 层次索引法示例

形所有边界点都数字化后，以顺序方式存贮各点坐标；各边界编号，由点索引与边界线号相联系，再用边线编号与多边形相联系，从而形成层次索引结构。仍以图 3-5 为例，可得到图 3-6 所示图形、表 3-2 所示的文件和图 3-7 所示的层次索引结构图。

该法消除了相邻多边形边界的数据冗余和不重合问题，可以直接对复杂边界线简化或合并相邻多边形。对有关邻域信息和岛多边形信息的获取可以通过对多边形文件的线索引处理得到，但是这样做较为繁琐，给邻域函数计算、消除无用边、处理岛多边形信息以及检查拓扑关系带来一定困难，而且工作量大时容易出错。

表 3-2

点文件		线文件		多边形文件	
点号	坐标（x，y）	线号	点号	多边形号	边界线号
1	$x1$，$y1$	A	1，4，5，6	I	A，B
2	$x2$，$y2$	B	1，4，7，8，9	II	B，C，D
…	……	C	1，2，3，4	III	D
12	$x12$，$y12$	D	10，11，12		

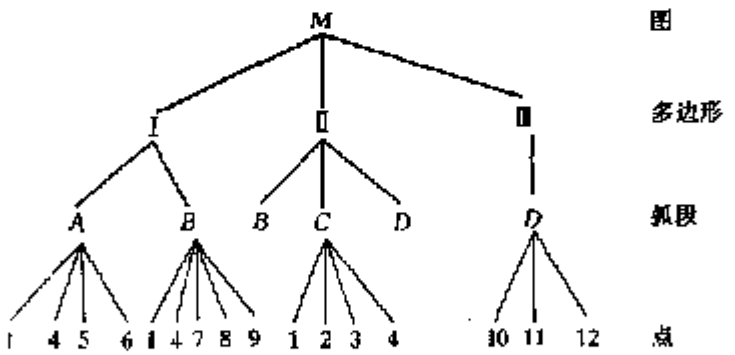


图 3-7 层次索引结构

3. 拓扑结构法

拓扑结构的建立可以通过两种途径达到：一是由用户在数据输入进行数字化时同时输入部分信息，如多边形边界的编号、边界结点的序号、边界左右多边形编号等标识信息；二是可以在全部数据经过数字化后，由专用软件自动搜索建立。属性可以通过每一多边形内建立编号与之联系。

所谓拓扑结构是指确定各地理实体空间关系的数学方法。为了准确描述空间目标的位置和空间关系，在几何上主要从两个方面进行。在几何形态方面，常采用解析几何方法来分析，这主要是以几何坐标为基础，涉及空间目标的角度、周长、方向、距离和面积等；在

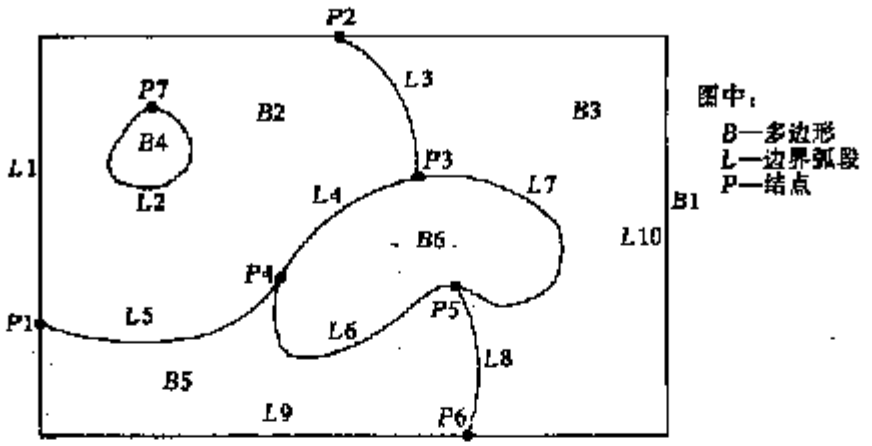


图 3-8 拓扑结构示例

空间关系方面，则采用拓扑几何方法来描述，主要涉及的是空间目标之间的“相邻”、“相连”、“包容”、“在里面”和“在外面”等关系。这样，就可能发生几何结构相差很大的图形，它们的拓扑结构却可能相同。所以说，若从几何形态的观点出发，主要研究的是图形的位置形状；而从拓扑观点看，则重视的是点、线和多边形之间的相互关系。

由图 3-8 所示图形，可以明确表达出“多边形—边界线—结点”的“从大到小”的拓扑关系。列出的关系表格为表 3-3 和表 3-4。

表 3-3 多边形—边界线的拓扑关系 $B_i=B\ (L_i)$

多边形编号	边界线编号
B_2	$L_2\ L_3\ L_4\ L_5\ L_6$
B_3	$L_3\ L_{10}\ L_8\ L_7$
B_4	L_1
B_5	$L_5\ L_6\ L_9\ L_8$
B_6	$L_4\ L_7\ L_8$

表 3-4 边界线—结点的拓扑关系 $L_i=L\ (P_i)$

边界线	起始结点	终止结点	边界点坐标
L_1	P_1	P_2	$(x_{p1}, y_{p1})\cdots(x_{pn}, y_{pn})$
L_2	P_7	P_7	…
L_3	P_2	P_3	…
L_4	P_4	P_3	…

续表

边界线	起始结点	终止结点	边界点坐标
L_5	P_3	P_4	…
L_6	P_4	P_5	…
L_7	P_3	P_5	…
L_8	P_5	P_6	…
L_9	P_7	P_6	…
L_{10}	P_2	P_6	…

还可以写出由“结点—边界线—多边形”的“从小到大”的拓扑关系表（表 3-5 和表 3-6）。

表 3-5 结点—边界线的拓扑关系 $P_i=F(L_i)$

结　点	边　界
P_1	L_1, L_8, L_9
P_2	L_1, L_2, L_{10}
P_3	L_3, L_4, L_7
P_4	L_4, L_5, L_6
P_5	L_5, L_6, L_7
P_6	L_8, L_9, L_{10}
P_7	L_2

表 3-6 边界线—多边形的拓扑关系 $L_i=L(B_i)$

边界线	左多边形	右多边形
L_1	B_1	B_2
L_2	B_2	B_4
L_3	B_3	B_2
L_4	B_3	B_4
L_5	B_3	B_5
L_6	B_4	B_5
L_7	B_4	B_5
L_8	B_1	B_5
L_9	B_5	B_1
L_{10}	B_1	B_3

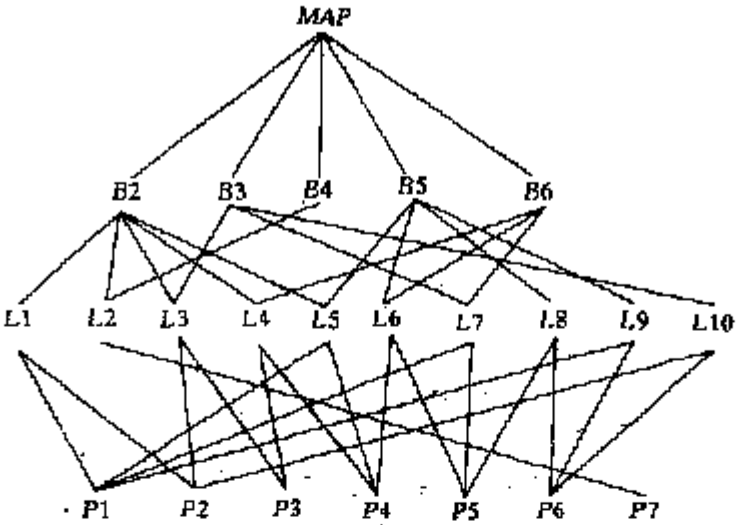


图 3-9　结点、边界和多边形之间的拓扑关系

此外，图 3-8 所示的图形还可以用图 3-9 所示的结构来表示。

如同前面讨论点、线实体的矢量编码一样，在形成多边形拓扑关系结构时，其编码内容应包括：①唯一标识码，②多边形标识码，③外多边形指针，④相邻多边形指针，⑤全部边界的记录，⑥范围（即最小和最大的坐标值）等几个方面。在数字化输入完成以后建立矢量拓扑结构的过程可以分为五个阶段来完成。首先，在所涉及的范围内对所有边界线进行分类，通过自相交或两两相交使各边界线都分割成两结点之间的具有单一拓扑性质的弧段，建立连接指针，构成图形外边界；然后，检查该外边界构成的多边形是否闭合，即应保证连接指针的正确，只有在单弧段组成的“孤岛”情况下，弧段指针才可能指向该弧本身；第三步从外边界连接各弧段构成各个内多边形，并给以顺序编号；第四步按梯形规则计算各多边形面积，并将其作为附加属性记录；最后，向建立的多边形连接其属性数据，建立与属性库的联系。

自动建立拓扑结构编码可以简化数据录入流程，减少工作量，生成的拓扑关系的可靠性和一致性较好。此外，该法还具有如下优点：

(1) 可将多边形生成完好的网络，减少数据冗余和公共边界不重合的矛盾；

(2) 因为生成的多边形、边界及其属性都存在着互相联系，从而可以实现邻域分析；

(3) 多边形与生成“岛”的层次关系不受限制；

(4) 其精度仅受数字化仪本身精度和计算机字长的限制。

总之，矢量结构自身的特点决定了矢量编码具有信息的完整性和运算的灵活性。但是从目前来说，尚无统一的最佳矢量结构编码方法。所以在具体工作中应根据任务的性质、要求和数据的特点灵活地应用和设计。

§3-3 栅格数据结构及其编码方式

一. 栅格数据结构

1. 栅格数据

栅格数据是最简单、最直观的一种空间数据结构，它是将地面划分为均匀的网格，每个网格作为一个像元，像元的位置由所在行、列号确定，像元所含有的代码表示其属性类型或仅是与其属性记录相联系的指针。在地理信息系统中，扫描数字化数据、遥感数据、数字地面高程数据（DTM）以及矢量—栅格转换数据等都属于栅格数据。在栅格结构中，如图 3-10 所示，点用一个像元来表示；线状地物用沿线走向的一组相邻像元来表示。由线的定义可知，每个线上的像元最多只有两个相邻像元在线上。面状地物或区域则用具有相同属性的相邻像元集合表示。由面的定义可知，每个像元可以有两个以上的相邻像元属于同一区域。来自遥感、数字摄影测量的数据属于典型的栅格结构。如遥感影像像元的数值表示该影像的灰度等级。由于栅格结构中的行列阵列的形式很容易为计算机存贮、操作和显示，特别是适合于 FORTRAN 等高级计算机语言建立文件或矩阵处理，所以这种结构容易在计算机上实现，而且算法简单，易于扩充和修改，特别是易于同遥感影像、数字摄影测量数据的结合处理，给地理空间数据处理带来极大的方便，受到普遍欢迎。因此，这也是栅格结构易于为多数地理信息系统设计人员接受的原因之一。图 3-10 为用栅格像元表示点、线、面实体的示意图。

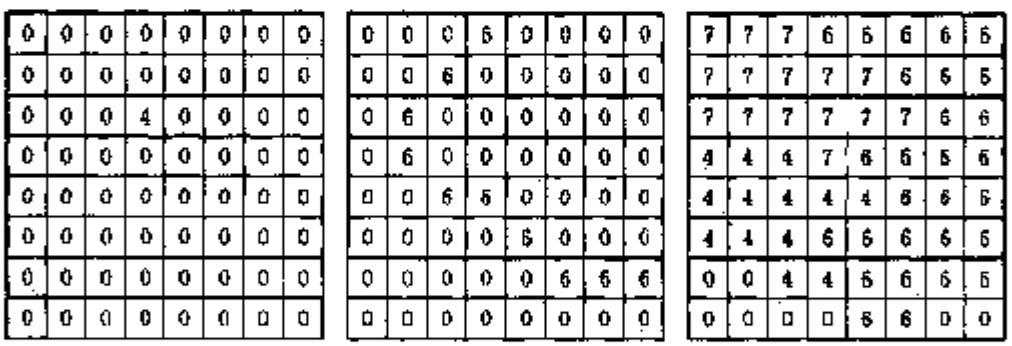


图 3-10 用栅格像元表示点、线、面实体

因为在栅格结构中，是将地面分划成相互邻接的大小均匀格网方块，然后使每一地块与一个栅格像元相对应。因此，由栅格结构所表示的地表数据是不连续的，是经过量化的离散值，而每一个像元大小与它所代表的实地地块大小之比就是栅格数据的比例尺。这样，在使用栅格数据计算面积、长度、距离、形状等空间指标时，若像元所代表的实地地块尺寸较大，则会造成较大的计算误差，如图 3-11 所示。

在图 3-11 (*a*) 中，*ac* 边的距离为 5，三角形面积为 6。而在图 3-11 (*b*) 中，*ac* 边的距离为 7，三角形面积为 7。这种误差是由于选择像元尺寸的大小引起的空间实体在形态上的

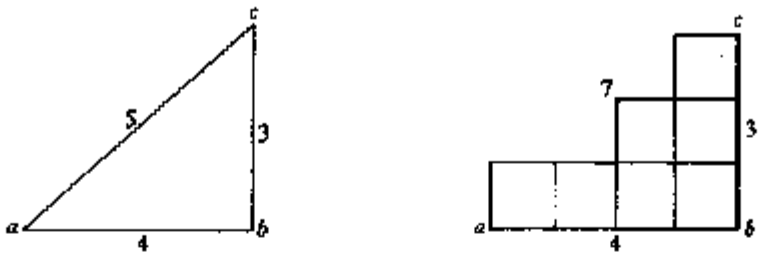


图 3-11　用栅格数据量测空间指标

畸变。从直观上可以看出，若进一步地使像元尺寸变小则这种形态畸变就可进一步地减小。此外，在许多栅格数据处理时，如数字影像处理，为解决由于像元尺寸大小引起的精度损失，常假设栅格表示的量化表面是连续的，以便可以利用某些具有连续求导的数学函数作为分析处理栅格数据的工具。当然，提高栅格数据的分辨率可以提高其数据精度，但是也带来数据量大的问题。通常，栅格数据的分辨率与其数据量之间呈平方指数率的函数关系。例如，一幅 1m×1m 的地图，若采用 1mm 的栅格像元，就有 10^6 个栅格数据，而一幅陆地卫星 MSS 图像有 3240×2340＝7.58×10^6 个栅格像元。

在栅格结构中，每个像元是处于二维数组中的一个元素，通常该元素仅有一个代码，表示该位置上的属性。但是由于在一个像元表示的地表范围内，可能存在多于一种的地物，即在同一空间位置上可能存在多重属性。这类似于遥感影像中的混合像元问题，如 Landsat 卫星影像中，每个像元对应的实地地块为 79 m×79 m 的矩阵区域，而影像上记录的光谱数据是每个像元所对应的地表区域内所有地物类型的光谱辐射的总和效果，因而这里包含有属性方面的偏差。

2. 栅格数据的获取

栅格数据通常可以由下述几种途径得到，即

(1) 格网法：在输入图上均匀划分网格，逐个网格地决定其属性代码，形成栅格数字地图文件，如图 3-2 所示。

(2) 由矢量结构数据转化为栅格数据。

(3) 扫描法：经过扫描对数据重采样和再编码得到栅格数据文件。

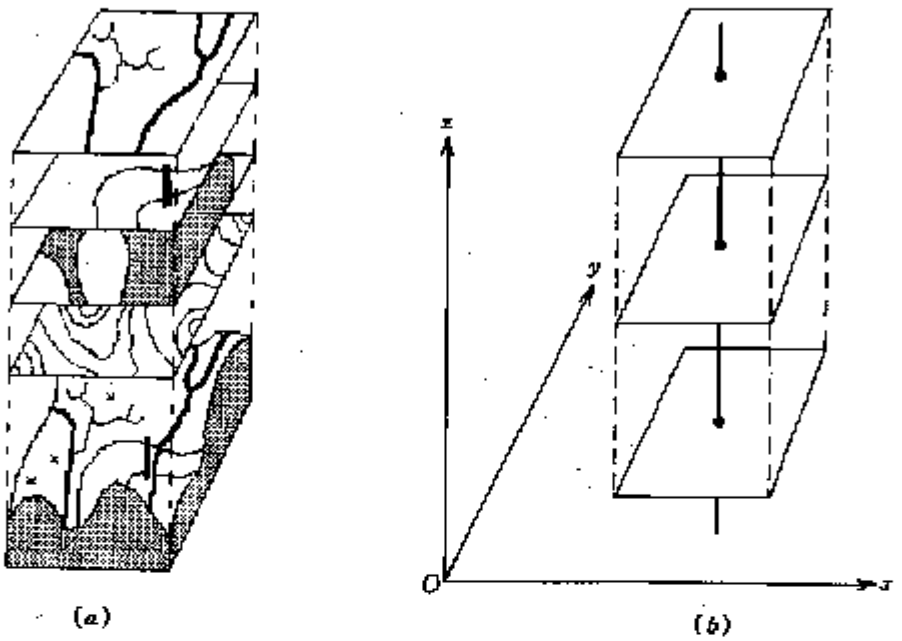
(4) 遥感影像数据。

无论采用上述何种方法获取栅格数据，都要求尽量保持原始数据的精度，要求尽量地减小由于像元尺寸的大小而产生的反映实体在形态上的畸变和属性上的偏差。

显而易见，缩小单个栅格像元的尺寸，即增加栅格像元的数目，可以使每个像元代表更为精细的地表矩形地块，从而使实体形态得到更为准确地描述，同时属性代码亦可较为正确客观地反映实际属性，也可表示更细小的地物类型，使模型更接近于真实状态。但是，细分格网由于增加了栅格像元的数目，在提高数据精度的同时，也带来了数据量大幅度增长，数据冗余严重的问题。

从栅格数据的属性精度考虑，解决多重属性的问题通常是采用数据层叠加的方法来处

理。在栅格结构中，对某空间位置上的几个地理属性，可以分别采用 n 个二维数据层叠加在一起后，构成与实地相符的地理模型。这种叠加技术对于绝大多数栅格数据处理是十分重要的。图 3-12 从几何上说明了数据层叠加技术的应用。



(a)　　(b)

图 3-12　层的概念与层的叠置

确定栅格像元的属性代码时应尽量保持与实地的一致性，保证最大的信息量。图 3-13 所示为某像元代表的一块矩形地块，内含 A，B，C 三种地物类型，O 点为中心点，在决定该像元的属性代码时，除了叠加方法之外，通常还可考虑如下一些确定方法。

(1) 中心点法，用处于栅格中心处的地物类型或现象特性决定该像元的代码。在图 3-13 中，中心点 O 落入属性代码为 C 的地物范围内，按中心点法规则，该地块相对应的栅格像元的属性代码为 C。中心点法常用于具有连续分布特性的地理要素，如降雨量分布等。

(2) 面积占优法：以占矩形区域面积最大的地物类型或现象特性决定像元的代码。图 3-13 中，B 类地物所占面积最大，故依面积占优法，相应的栅格代码为 B。该法常用于分类较细、地物类别斑块较小的情况。

(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像元代码。假设图 3-13 中 A 类地物是最重要的地物类型，则该区域相应像元的属性代码即可取 A。该法常用于具有特殊意义而面积又较小的地理要素，尤其是点、线状地理要素，如城镇、交通枢纽、交通线、河流水系等。在属性代码中应尽量表示这类重要地物。

3. 栅格数据文件的构成

在栅格结构中，叠加技术较常采用的是使每一数据上栅格像元与其属性代码的一一对应关系。因此而产生三种栅格数据文件的构成方法，如图 3-14 所示。

图 3-14(a)所示为在每个栅格像元上直接表示多重属性的方法；图 3-14(b)为在每个数

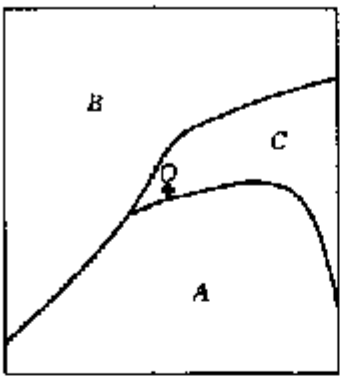


图 3-13　栅格属性代码的确定

据层上直接表示多重属性的方法，而图 3-14(c)则与前二法不同，它是根据在每个数据层上具有相同属性的像元来划分出不同区域，然后构成各个数据层，最后按层次结构组成地图文件。因为在这种方法中，每个数据层上不同区域的属性与构成区域的像元集合之间是一对多的对应关系，所以这种方法很适合应用压缩编码的方法来存贮数据，通常多采用链码、行程编码、块码和四叉树码等方法。应用栅格数据压缩编码方法的目的就是用尽可能少的数据量记录尽可能多的信息，减少数据冗余，节省存贮空间，提高操作效率。

二、栅格数据编码方法

1. 栅格矩阵法

栅格矩阵法是最简单、直观而又非常重要的一种栅格结构编码方法。通常就称按这种

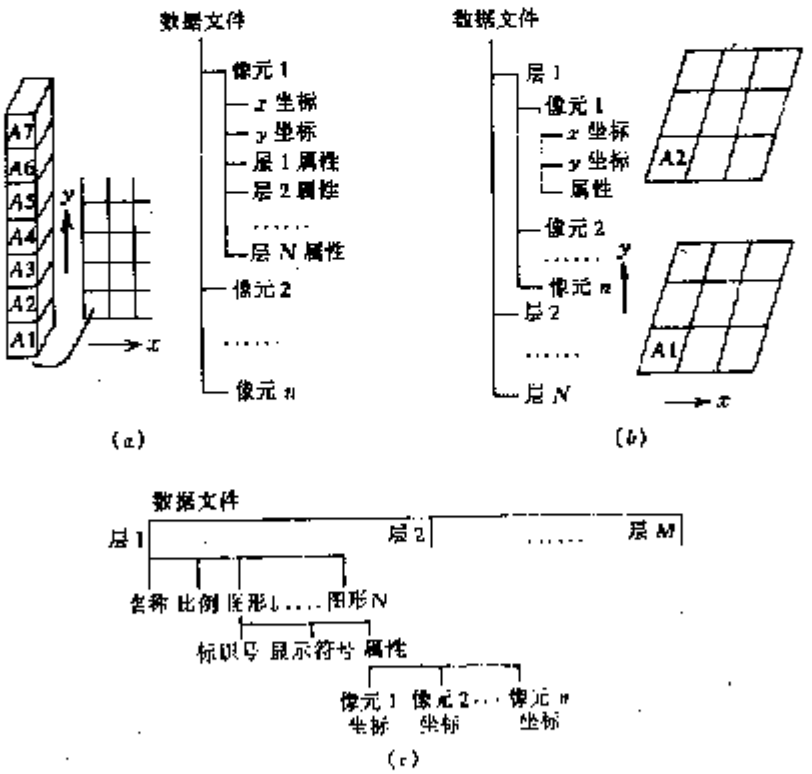


图 3-14　三种栅格数据结构形成的地图文件

编码产生的图像文件为栅格文件。实际上，栅格数据不论采用何种压缩编码方法，其逻辑原型都是采用栅格矩阵法产生的栅格文件，栅格数据是在二维表面上地理数据的离散量化值。栅格像元组成栅格矩阵，用像元所在的行列号来表示其位置。通常是以矩阵左上角开

始逐行逐列存贮，记录代码。可以每行都从左到右逐像元记录，也可奇数行从左到右而偶数行由右到左来记录。图 3-10 (c) 所示的栅格结构就可以记录成：7，7，7，6，6，5，6，6；7，7，7，7，7，6，6，6；7，7，7，7，7，7，6，6；4，4，4，7，6，6，6，6；4，4，4，4，4，6，6，6；4，4，4，6，6，6，6，6；0，0，4，4，6，6，6，6；0，0，0，0，6，6，0，0。在 FORTRAN 或 BASIC 程序中，也可采用顺序文件或随机文件记录这些数据，可以将每一行作为一个记录，记录长度＝单字长×列数，在每个记录结尾还可加标识符标记。一般是一个文件存贮一层信息，如果是多重属性，就采用多个文件存贮。当然，也可以在一个文件中存贮多层信息，记录每个像元的行列号以及属性代码。

2. 链码

这种方法又称为边界链码，它用一个起点和一系列在基本方向上的单位矢量描述出线状地物或区域边界。所采用的基本方向可以事先定义。例如，我们可以定义基本方向为：东＝0，北＝1，西＝2，南＝3，如图 3-10 (b) 所示的属性代码为 6 的线状地物，其位置可以表示为：1，4，3，2，3，2，3^2，0^2，3，0，3，0^3，其中前二个数字 1 和 4 表示起点为第 1 行第 4 列，从第 3 个数字起每个数字表示单位矢量的方向。数字的上标表示单位矢量的前进量。这里将单位矢量的长度默认为一个栅格像元。对基本方向的选择还可按下图来定义。

8 个基本方向上的单位矢量用 0～7 来记录，则图 3-10 (b) 又可记录为：1，4，5，5，6，7，0，7，7，0，0。其中，前 2 位数字仍为起点的位置。从第 3 位数字起，记录单位矢量的方向。

链码可以有效地压缩栅格数据，尤其是对于计算面积或长度或转折方向的凸凹度等运算较为方便，比较适合存贮图形数据。缺点是在进行叠加操作时，对边界的合并和插入等修改编辑工作比较困难。对局部的修改要涉及改变整体结构。因此效率较低。另外，由于链码是以每个区域为单位存贮边界，所以相邻边界将被存贮两次而产生数据冗余，这些问题使链码的应用受到一定限制。

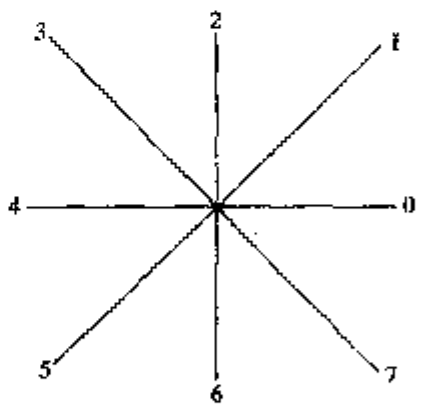


图 3-15　链码记录的基本方向

3. 行程编码

从前面讨论的图 3-14 (c) 知道，在每个数据层上不同区域的属性与构成该区域的像元集合之间存在一对多的对应关系，也就是说，在同一区域内，相邻像元的属性代码值是相同的，这样就非常适合于用压缩的编码方式，将相邻等值的像元合并，记录下行程长度和代码值，这种方法称为行程编码。

通常地理数据都具有较强的相关性，如上所述的一对多的对应关系，所以行程编码是栅格数据压缩的重要编码方法。对于在行（或列）方向上相邻的若干点具有相同属性代码的栅格图像，可以采取压缩那些重复的记录内容。通常有两种编码方案可实现对数据的压缩。一种编码方案是，在进行行列扫描过程中只在各行（或列）数据的代码发生变化时依

次记录该代码以相同代码出现的个数，从而实现数据的压缩。例如对图 3-10 (c) 所示的栅格数据，可进行如下行程编码：(7，3)，(6，5)；(7，5)，(6，3)；(7，6)，(6，2)；(4，3)，(7，1)，(6，4)；(4，5)，(6，3)；(4，3)，(6，5)；(0，2)，(4，2)，(6，4)；(0，4)，(6，2)，(0，2)。这些只用了 38 个整数就可以表示图 3-10 (c) 所示的栅格数据，而前述的栅格矩阵法中直接编码记录需要 64 个整数表示，可见行程编码压缩数据是十分有效且简便的。实际上，栅格图像的复杂程度直接影响到行程编码的压缩比，图形变化少，则行程数就少，压缩效率就高，反之压缩比就低。

另一种编码方案是在逐行（或列）记录属性代码时，仅记录下发生变化的位置和相应代码，仍以图 3-10 (c) 为例，可得到另一种行程编码为：(1，7)，(4，6)；(1，7)，(6，6)；(1，7)，(7，6)；(1，4)，(4，7)，(5，6)；(1，4)，(6，6)；(1，4)，(6，6)；(1，4)，(4，6)；(1，0)，(3，4)，(5，6)；(1，0)，(5，6)，(7，0)。

行程编码具有压缩效率高、易于检索、运算简单等优点，区域越大、数据相关性越强则该法节省存贮空间越多。这适用于计算机存贮容量小，且数据需大量压缩，而又要避免复杂的编码解码运算增加处理和操作时间的情况。此外，当需要提高栅格数据精度而采取加密栅格像元的情况下，由于数据的相关性，应用行程编码时，数据量没有明显的增加。

4. 块码（二维行程编码）。

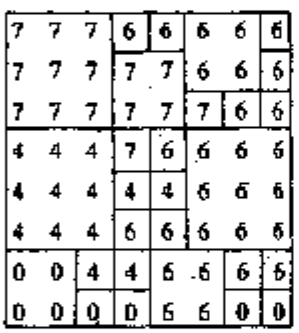


图 3-16 块码记录单元

在前面讨论行程编码时是在一维情况下(按行或列)记录像元的属性及其位置。现若采用方形区域作为记录单元，则可以将行程编码扩展为二维情况下的编码方式，即所谓块码。在块码方法中，每个记录单元应包括相邻的具有同样代码的若干栅格像元。它的数据结构可以由起始位置(记录单元之中左上角像元的行，列号)、记录方形单元的边长范围和该单元的属性代码三部分组成。以图 3-10 (c) 所示图像为例，按块码记录单元划分出方形区域如图 3-16 所示，则可得到块码编码为：

(1，1，3，7)，(1，4，1，6)，(1，5，1，6)，(1，6，2，6)，(1，8，1，6)；

(2，4，2，7)，(2，8，1，6)；

(3，6，1，7)，(3，7，1，6)，(3，8，1，6)；

(4，1，3，4)，(4，4，1，7)，(4，5，1，6)，(4，6，3，6)；

(5，4，1，4)，(5，5，1，4)；

(6，4，1，6)，(6，5，1，6)；

(7，1，2，0)，(7，3，1，4)，(7，4，1，4)，(7，5，2，6)，(7，7，1，6)，(7，8，1，6)；

(8，3，1，0)，(8，4，1，0)，(8，7，1，0)，(8，8，1，0)。

该例中，块码用了 112 个整数来表示栅格数据，较前述栅格矩阵法编码还多。这是因本例中为描述方便，所分栅格像元数目较少。在实际应用中，栅格划分很细，且涉及区域亦较大，则数据冗余要多得多，这样方能显示出块码压缩的效果。此外，块码记录本身还

可作一些技术处理，如行号可以通过行间标记而省去记录，行号和范围等也不必用双字节整数记录，这样可进一步减少数据冗余。

因为块码方法是行程编码在二维情况下的扩展，所以，块码方法与行程编码一样，地理数据的相关性越强，则其压缩效率越高。但随栅格图像复杂程度的提高而降低其效率。所表示的具有相同代码的图块越大，块码编码的压缩比就越高，反之则低。此外，块码在图像合并、插入、检查延伸性、面积计算等操作时有明显的优越性。但是，在某些情况下，按行程编码或块码编码的栅格数据还须通过解码使其返回到栅格矩阵编码的基本形式。

5. 四叉树编码

四叉树编码是以栅格数据二维空间分布的特点，将整个图像区域按照四个象限进行递归分割（$2^n \times 2^n$）最后逐步分解为一系列被单一类型区域内含的方形区域，所分解的最小方形区域为一个栅格像元。按照像限递归分割的原则，所分图像区域的栅格阵列应为 $2^n \times 2^n$（n 为分割的层次数 $n \geqslant 1$）的形式，无论分割到哪一层的象限，只要子象限上仅含一种属性代码或符合既定要求的少数几种属性时，则停止继续分割，否则就一直分割到单个像元为止。这样最后可得到一棵四分叉的倒向树。四叉树编码正是通过这种树状结构来记录和压缩栅格数据，并通过这种结构实现查询、修改和量算等操作。因此，四叉树编码法不仅能压缩数据，还可提高图形操作的效率。图 3-10（c）所示的栅格数据，经过四叉树编码得到的四叉树如图 3-17 所示。

在图 3-17（a）中各子象限的大小不同，它们是由组成该子象限的具有相同代码的栅格像元构成的子块而决定。图 3-17（b）中最上面的结点称作根结点。它对应于整个图形区域。在这个例子中，总共划分出四层结点。每层结点对应于不同尺寸的子象限。如第二层的 4 个结点对应于整个图形区域的四个象限，分别记作：

西北（NW）象限 =2

东北（NE）象限 =3

西南（SW）象限 =0

东南（SE）象限 =1

若某子象限仅含一种属性代码时，就不再继续划分，即被记为叶结点或终止结点。叶结点可以落在任意层上，它表示的区域仅具有单一类型的地物或符合既定要求的少数几种地物。图 3-17（b）所示的四叉树结构中，共有 43 个叶结点，这说明原栅格数据图被划分为 43 个大小不等的子象限。每个叶结点下的数字表示该子象限代表区域的属性代码。每个结点上方的数字表示该子象限的地址编号。图 3-17 中，共分割层次数 $n=3$，所得到四叉树的最大层数为 $n+1=4$。也就是说，若象限递归分割层次数为 n，则可得到结点的最大层次数为 $n+1$。

从对上述图形的四叉树编码过程中可见，位于结点层次较高的子象限尺寸也较大，说明其分解深度小，也即分割次数少；而低层次上的象限尺寸就较小，反映其分解深度大即分割次数多。这样编码后，可反映出整个图形区域的空间地物分布情况，在某些位置上单一地物分布较广，则采用较少的分割次数；在地物较复杂，变化较大的区域，则用加深分解深度，增加分割次数的方式编码。所以，采用四叉树编码方法能够自动地依照实际图形

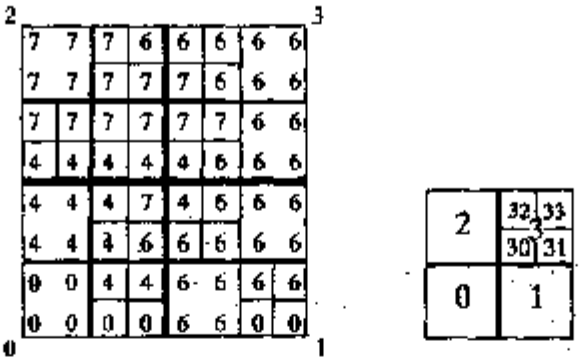


图 3-17 (a)　四叉树编码结构象限图

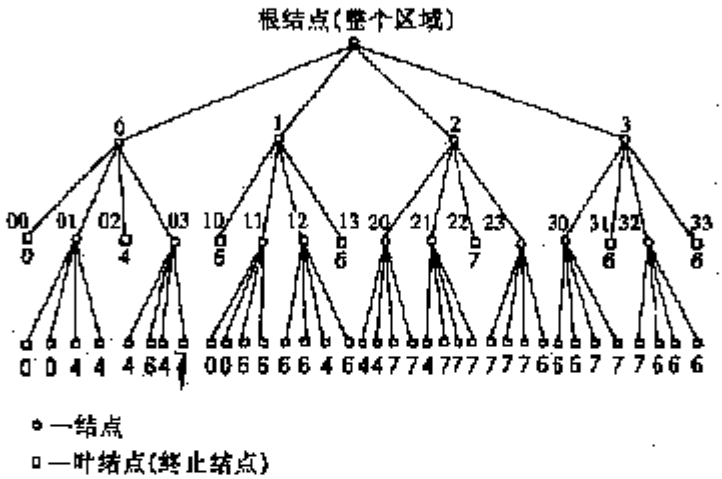


图 3-17 (b)　四叉树结构

区域的地物分布情况而调整象限分割的尺寸和疏密，从而产生较高的数据压缩效率。

一种四叉树编码的方式被称为指针四叉树（或常规四叉树），即通过在子结点和父结点之间设立指针的方式建立起整个结构。按这种方式，四叉树的每个结点通常存储 6 个量，即四个子结点指针、一个父结点指针和该结点的属性代码。这种方法除了要记录叶结点外，还要记录中间结点，一般要占用较大存储空间。但是另一方面，指针四叉树虽然增加了一些存储量，却换取了处理上的简便和灵活，例如，要将相邻的四块四叉树合并成一棵新树，对指针四叉树来说就十分容易，只要将这四块四叉树作为新树的子树，使它们的根结点指针指向一个共同的根结点就行了。因此，指针四叉树对于数据检索、多要素叠置分析和求实体间的空间关系等操作都很方便。在 GIS 中，还经常采用一种能以较小数据冗余存贮图形图像，而且又能方便地完成各种图形图像操作的编码方式，即所谓线性四叉树方法。该方法记录从根结点至叶结点的路径以及叶结点的属性代码，即仅记录每个结点的地址、深度和属性代码三个量。这种方法较指针四叉树节省存储量，并且由于记录了结点地址，就能够直接找到其在四叉树中的走向路径，也可以换算出栅格区域中的行列位置。

通常定义从根结点到叶结点的路径可以按照象限递归分割的顺序编号进行。无论分割到哪一层，总是用 0、1、2、3 分别表示 SW、SE、NW、NE 四个象限的编号，只是每个子

象限子结点编号的前缀必须为其父象限(父结点)的编号，如图 3-18 所示。

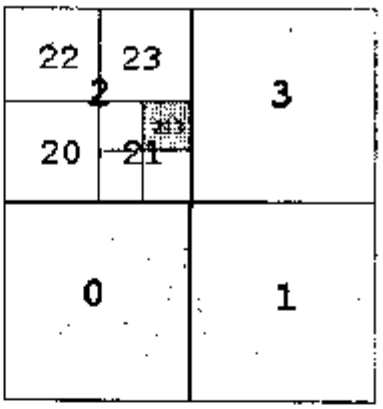


图 3-18 四叉树地址编码

上图中数字均表示该子象限的地址编号。图 3-17 (b) 给出按该法编号后的树状结构。每个子象限地址编号的位数等于该象限所处的分层深度 n。如图 3-18 中所示地址编号为 213 的子象限，就表示它为一个经过 3 次分割象限后，在第 3+1 层象限中的一个子象限。"213" 的意思为，经过第 1 次分割象限，该子象限位于 NW 象限，记作 2，经过第 2 次分割，位于 SE 象限记作 1，再经过第 3 次分割，落于 NE 象限记作 3，结合图 3-17 (b) 可知，该叶结点的属性代码为 7。

通常每个叶结点的地址编号在计算机中是用二进制数来表示的，在每一层上的象限位置 (0、1、2、3) 均可用两位二进制数写出。例如，0 记作二进制数 00，1 记作二进制数 01，2 和 3 分别记作二进制数 10 和 11。以图 3-18 为例说明，编号为 213 的子象限 (叶结点) 的地址可用二进制表示为：

编号 213→100111

这样，记录了各个叶结点的地址，再记上各自相应的属性代码值就记录了整个图像。并可在此基础上进行多种图像操作。

实际上，为了得到四叉树叶结点的地址码可以采用一种被称为 Morton 码的方法直接得到。Morton 码是一种自然数码以图像左下角为原点 (0 行，0 列开始起算) 的行列坐标系为基准，它的值与二维栅格阵列的位置相对应。如图 3-19 所示。

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

图 3-19 Morton 码与行列号的关系

由上图 3-19 可知，Morton 码的扩展顺序如同前述对子象限的编号顺序是相似的。根据 Morton 码，每个栅格像元对应着一个 Morton 码（十进制数），而每个 Morton 码又对应着二维栅格阵列的行列号，只要将行列号转化为二进制数，然后将它们按位交错排列放入 Morton 码变量中，即可得到四叉树各叶结点的二进制数地址码。如前例图 3-18 中编号为 213 的叶结点。它所对应的 Morton 码为 39，在以左下角为原点的行列坐标系中该叶结点的行列号为（5，3），将其分别化为二进制数为 101 和 011，然后将两数按位交叉排列，即可得到该叶结点的二进制数地址码为 100111。如图 3-20 所示。

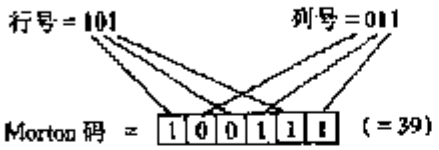


图 3-20 MORTON 码变量

应用 Morton 码可以将栅格数据的二维数组形式（阵列）转化为以 Morton 码为下标的一维数组。通过对一维数组的扫描，根据四叉树按四划分的特点，依次检查每四个相邻像元叶结点的属性代码是否相同，若完全相同，则将其合并成一个父结点象限，记录下地址和代码值，否则不作合并。完成第一轮检索合并后，又依次检索新形成的父结点层中每四个大块象限的代码值，如果不完全相同，则不作合并，否则就合并成更大的象限——更高一层的父结点。如此递归循环，直到全部像元检索完毕。图 3-21 所示为一个用 Morton 码为序的四叉树编码过程。为描述简便，这里给出一个简单例子。

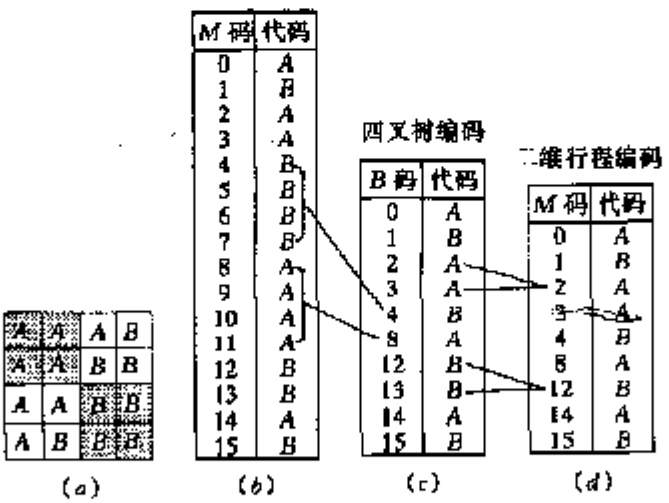


图 3-21 四叉树编码过程

图 3-21 (a) 给出一个含两种地物类型的二维栅格数据阵列，根据 Morton 码，记录下各像元的位置及其属性代码如图 3-21 (b)。通过对以 Morton 码为下标的一维数组扫描，以每 4 个相邻像元为单位，得到四叉树压缩编码形式如图 3-21 (c) 所示。此外，采用 Morton 码记录栅格像元的地址码还可以用在二维行程编码的过程中。若不考虑四叉树要求的按四分割的原则，仅按 Morton 码的扩展顺序进行扫描，当碰到前后的栅格像元代码值不相同时，记录下该像元的地址码和代码，就可得到对整个二维栅格阵列统一的编码，即二维行

程编码，如图 3-21 (*d*) 所示。要说明的是图 3-21 (*c*) 和 3-21 (*b*) 的地址码虽然相同，但二者的性质是不一样的，前者是按四叉树编码而后者实质上属块码范畴。

四叉树编码具有可变的分辨率。并且有区域的性质，压缩数据灵活，许多运算可以在编码数据上直接实现，大大提高了运算效率，是优秀的栅格数据压缩编码之一。其不足之处在于该方法的树状表示变换缺乏不变性，有时，相同形状和大小的两个区域可能表示为截然不同的结构。因此，在形态分析和模式识别方面较困难，难于设计出统一的算法。

前面讨论了 5 种栅格数据编码方法。一般说来，对数据的压缩是以增加运算时间为代价的。在这里时间与空间是一对矛盾。为了更有效地利用空间资源，减少数据冗余，有时不得不多费一些机器运算时间进行编码以及进行较为复杂的图形运算。一个好的压缩编码方法就是要在尽可能减少运算时间的基础上达到最大的数据压缩效率，并且使算法适应性强，易于实现。栅格矩阵法简单明了可直观地反映栅格图像数据，但是数据冗余太大；链码的压缩效率较高，已接近矢量结构，对边界的运算比较方便，但是不具备区域的性质，区域运算较困难；行程编码既可在很大程度上压缩数据，又可较大限度地保留原始栅格结构，而且编码解码也较为容易；块码和四叉树编码都具有区域性质，而且有可变的分辨率，压缩效率比较高。四叉树编码可以直接进行大量的图形图像运算，效率高，使用也日益广泛。当前，在四叉树编码的基础上，对三维栅格数据已开始发展所谓八叉树编码方法。

§3-4 矢量数据与栅格数据的转换

一、两种数据结构的比较

在前两节的讨论中已知矢量结构与栅格结构各有其特点，矢量数据结构精度高，易于表达拓扑关系，存贮量小；而栅格数据结构则结构简单，易于进行空间叠加和空间分析，并且具有操作速度快等特点。从表面上看来，两种数据结构似乎是两种截然不同的空间数据结构。矢量结构表示空间位置明显，但表达属性隐含；而栅格数据则相反，表达属性明显，表示位置隐含。然而，从实质上说来，两种结构是从两个不同的方面对空间数据进行的描述，按通常意义上的理解，栅格结构只是矢量结构在某种精度上的一种近似。在 GIS 中，无论采用何种数据结构，所涉及到的数据量与数据精度总是构成矛盾的双方，要提高数据精度，矢量结构需要记录更多的线段中间点和结点，而栅格结构则需要更多的栅格单元。因此，在表示空间数据方面，两种数据结构可以是同样有效的。

矢量结构在表达线状地物方面具有直观性，尤其是在反映网络信息方面如管线、交通具有其独特的优势；此外，矢量结构还适合用在建立拓扑关系和进行拓扑关系的搜索，比如计算多边形形状，搜索邻域层次信息等方面。矢量结构的另一重要特点是它的计算精度与数据存贮量方面的优势。一般来说，要使栅格结构描述的图形取得与矢量结构相同的表达精度，在数据处理量上矢量结构要大大少于栅格数据量。但是，由于矢量结构对面状地物的描述是通过对其边界的描述来表达的，若从空间操作和空间分析方面来考虑，例如，地图的叠加操作，若采用矢量结构的数据操作就较为复杂。

归纳起来，矢量结构的优点为：

(1) 显示图形的质量高，可视性好；

(2) 存贮数据量较小；

(3) 应用网络分析可建立完整的拓扑关系；

(4) 显示数据精度高；

(5) 可对图形及其属性进行检索、更新和概括。

其不足之处为：

(1) 数据结构较为复杂；

(2) 多图叠加操作较困难；

(3) 由于拓扑关系的不同，使得模拟操作困难；

(4) 数据输出的费用较高；

(5) 难于进行某些空间分析运算。

栅格数据操作总的来说比较容易实现，比如在对空间位置坐标的检索上，采用栅格结构极为方便；在对给定区域内的统计指标运算时，用栅格结构可很快得到多边形面积、线密度、周长等统计量。此外，在给定范围内进行开窗、缩放等操作也较矢量数据优越得多。栅格结构除了在空间分析方面的优势以外，还有个重要特点是作为GIS的主要数据源之一的遥感影像和数字摄影测量的扫描数据均为栅格结构。因此，栅格数据结构也是一种有发展潜力的方法。它的数据冗余较大的缺点，目前正在通过改进压缩编码技术使数据冗余降低，同时也提高某些图形操作的效率。另一方面，随计算机硬件技术的飞速发展，大容量高速记录介质如可读写光盘的出现，大大减轻了数据冗余对计算机系统的压力，从而使人们更加关注操作的效率。

对栅格结构的优点可归纳为：

(1) 数据结构简单；

(2) 易于与遥感影像和数字摄影测量等数据结合；

(3) 适于进行各种空间操作和空间分析；

(4) 易于进行模拟操作；

(5) 有发展潜力。

其缺点为：

(1) 图形数据量大；

(2) 减少数据量要以牺牲精度为前提；

(3) 图形显示的可视性不如矢量结构；

(4) 网络分析较难；

(5) 投影变换较为耗时。

从以上讨论中已知，两种结构各有其优缺点，但在表示空间数据上又是同样有效的方法。对于一个与遥感数据相结合、要进行各种空间操作的GIS来说，栅格结构是必不可少的；而对地图数字化、拓扑检测和矢量绘图而言，矢量结构数据也是不能缺少的。所以在GIS中较为理想的方案是采用两种数据结构，即含有栅格结构又保持矢量结构，以形成一种混合数据结构，通过计算机程序实现两种结构的高效转换，数据根据需要按矢量结构或栅

格结构存贮，最大限度地减少数据冗余，提高数据精度；根据数据的提取和分析、输出，由程序自动地根据操作的需要选取合适的数据结构，以获取最强的分析能力和时间效率。矢量和栅格两种结构的并存、结合及转换，对于提高GIS的空间分辨率，数据压缩和增强系统分析，输入输出的灵活性是十分重要的。还要说明的一点是，这两种数据结构事实上也是与当前一定的输入输出设备相联系的。目前流行的线跟踪式数字化仪对地图数字化的结果首先是形成矢量结构，而高精度的增量式绘图仪输出数字地图也要求是矢量结构的；栅格结构可由人工读点或扫描式数字化仪得到，对于屏幕显示和打印机或喷墨绘图仪输出来说，栅格数据又是必须的。

二、两种数据结构的相互转换

对地图数字化主要是采用两种方式，即手扶跟踪数字化产生矢量结构数据和图像扫描产生栅格结构数据。在GIS中，这两种数据结构是最基本的。因此矢量数据与栅格数据的相互转换的方法也就必然要成为在不同GIS之间进行数据传输、信息共享所不可缺少的手段；即使在同一GIS系统中，在实现其空间分析和数据操作的过程中，为使系统运行的总体效率和功能得以圆满完成，这种数据格式的相互转换也是不可避免的，并且也是GIS的重要功能之一。

一般地，GIS中的数据格式转换基本上可分为矢量数据向栅格数据转换和栅格数据向矢量数据转换两种方式。这种互相转换是通过各种算法编制的转换程序得以实现的。近年来已发展了许多高效的转换算法以适应不同的环境需要。

1. 矢量数据向栅格数据的转换

在实际工作中，由于扫描输入设备较昂贵，所以目前仍多采用手扶跟踪数字化方法进行地图数字化。而许多图形操作和输出表示都需要栅格数据，因此，研究矢量数据向栅格数据的转换是很重要的。

从前述讨论中已知，空间图形可抽象地以点、线、面（多边形）三种要素来描述。对于以矢量结构表述的点状实体而言，每个实体仅由一个坐标对表示其空间位置，而在栅格结构中的点实体（像元）的位置，则是由该点所处的行列位置所确定。因此对于点实体的两种结构转换基本上只是一个坐标精度变换的问题，在技术上并不难解决。用矢量结构表示的线实体是由一系列点的坐标对来表述其位置的，在变换为栅格结构时，按解析几何中的两点式直线方程，根据栅格精度要求，在每两坐标点之间插入一系列栅格像元，每个坐标点变换为栅格结构的行列坐标，可见在线实体上实现这种结构转换也是不难实现的。因此，通常对转换算法的研究主要放在对面实体或称多边形的数据结构转换上。为便于叙述，这里对所谓多边形及其矢量格式和栅格格式给予具体定义。

多边形：指由折线段组成的封闭曲线围成的多边形区域，具有可唯一标识的多边形编号。该编号可作为多边形属性代码也可是指向属性数据记录地址的指针。

矢量格式：用在某坐标系下的坐标对序列表示多边形边界的几何位置。整个多边形边界由该多边形与其邻接多边形的公共边界弧段组成，每个边界弧段又由多个短折线相连组成，按弧段的前进方向，记有弧段左右侧多边形编号记录。

栅格格式：表示为一个数据阵列文件或数组，其中每个数据在阵列中的行列位置表示

一个栅格像元的相对几何位置，其数值则表示该栅格像元所在位置的多边形编号（属性）。

所谓矢量格式向栅格格式的数据转换，就是在矢量格式的多边形内部给所有栅格像元上赋以相应的多边形编号，故又可称为多边形填充。现将几种主要的转换算法描述如下：

（1）内部点扩散算法

该算法由每个多边形的一个内部点（种子点）开始，向周围8个方向的邻点扩散，然后判断各个新加入点是否在多边形边界上。若新点在边界上，则该点不作为种子点，停止扩散；如果该点不在多边形边界上，则该点作为新的种子点与原有的种子点一起继续向外扩散，并将新的种子点赋以该多边形编号。重复上述过程直到所有的种子点填满该多边形并遇到多边形边界停止为止。

扩散算法程序设计比较复杂，需要在栅格阵列中进行搜索，占用内存较大。当计算机内存容量受限制时，由于搜索过程中读写磁盘文件需花费大量的时间，故该法难于采用。另外，由于栅格精度的限制，在一定的栅格精度下，如果复杂图形中某个多边形两侧的边界落入同一个或相邻的两个栅格像元之内，就会造成该多边形不连通，如图3-22（e）所示。此时，仅用一个种子点就不能完成整个多边形的填充。

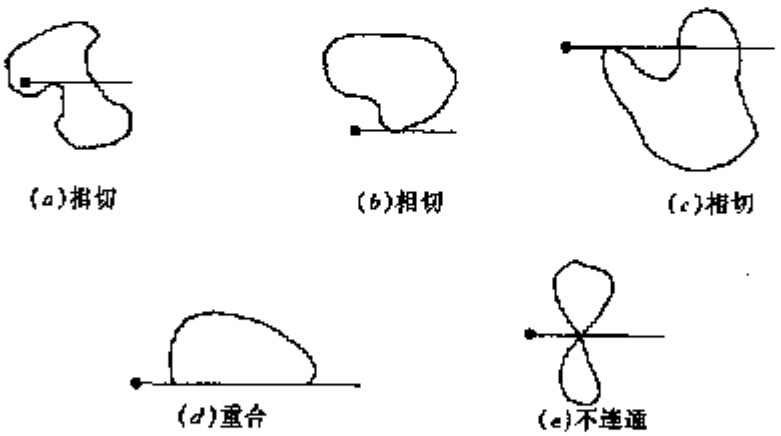


图3-22　射线与多边形边界相交的几种特殊情况

（2）射线算法

射线算法是逐点判断各栅格点是位于多边形内或在多边形之外。通过从待判点向外引射线，然后判断该射线与多边形的边界相交的次数。如果相交偶数次，则该点在多边形外部；若为奇数次，则该点为多边形内部点。如图3-23所示，当n＝奇数时，可给该点赋值该多边形编号。

射线算法要计算射线与多边形边界的交点，计算量较大。此外，判断射线与多边形边界相交时，有一些特殊情况会影响判断的正确性，如图3-22所示的几种特例情况，必须予以排除。而这又会造成算法的不完善并增加编程的复杂性。

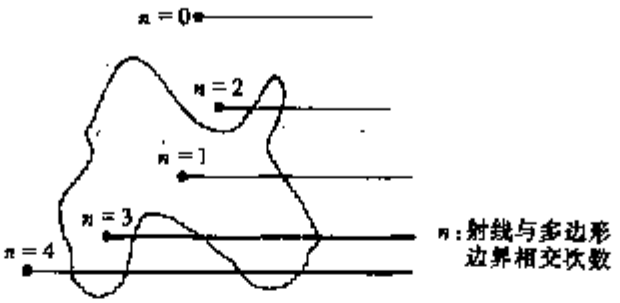


图 3-23　射线算法示意图

(3) 扫描算法

扫描算法是射线算法的改进。通常情况下，沿栅格阵列的行（或列）方向进行逐行(列) 扫描，对扫描线每两次遇到多边形边界的交点之间的栅格，判别为属于多边形内部点给予赋值该多边形编号。扫描算法较射线算法大大提高了计算效率。但是，射线算法中的一些特殊情况，如图 3-22 所示，在扫描算法中仍然存在，仍需给予特殊处理。

(4) 边界代数算法

边界代数法是一种基于积分思想的矢量格式向栅格格式转换的方法。在平面直角坐标系下，沿多边形边界对坐标轴积分，可以得到闭合多边形的面积。模仿积分求多边形面积的过程，首先以单个多边形的简单情况为例，说明边界代数法的转换原理，然后再推广到由多个多边形组成的区域转换方法。

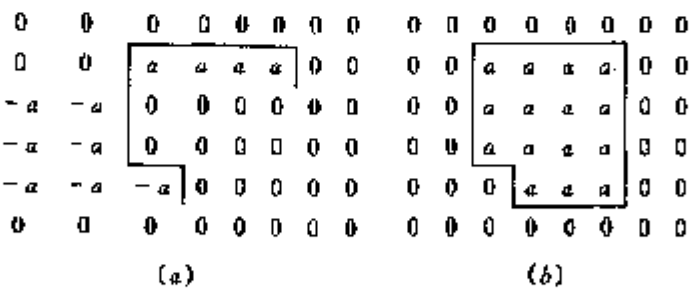
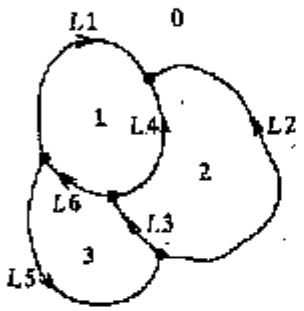


图 3-24　单个多边形边界代数法转换示意图

图 3-24 表示转换单个多边形的情况。设该多边形编号为 a，对初始化的栅格阵列各像元赋值为零，欲完成转换，就是要将该多边形区域内的所有栅格像元均赋值为 a，而多边形区域以外的像元仍保持原值零。因此，转换时以栅格行列为参考坐标轴，从多边形边界上某点开始顺时针搜索边界线。当边界上行时，如图 3-24 (a)，位于该边界线段左侧的具有相同行号的所有栅格像元被减去 a 值；当边界下行时，如图 3-24 (b)，该边界左侧（或为边界行进方向的右侧）的所有栅格像元加上一个值 a，当沿边界搜索一周重新回到起点后，所有多边形内部像元点都已被赋值 a，而多边形外部的栅格像元值不变，从而完成多边形填

充的转换。

以上讨论了单个多边形转换的方法，但在实际工作中，每幅地图都是由多个多边形区域组成的。因此还必须研究多个多边形区域的转换问题。在本章§3-2节中讨论了矢量编码的拓扑结构法，利用多边形边界弧段的拓扑信息可以使这种转换效率大大提高。根据拓扑结构编码的矢量数据，图上每条边界弧段都与两个不同编号的多边形相邻，并且，按边界弧段的前进方向记录下了每个弧段的左右多边形编号。下面以三个多边形组成的区域为例，讨论多个多边形的矢量向栅格转换问题，如图3-25所示。



(a)

边界弧段号 Li	左多边形	右多边形
$L1$	0	1
$L2$	2	0
$L3$	3	2
$L4$	1	2
$L5$	3	0
$L6$	3	1

(b)

图3-25　多个多边形的情况

根据前面对单个多边形转换的运算原则，这里要分别对三个多边形填充其代码值为1、2、3，而外多边形仍保持其原值零（所有加减运算均对该边界左侧栅格像元而言）。因此，可以得到：

多边形1：$L1$ 上行 -1，下行$+1$；
　　$L4$ 上行 $+1$；
　　$L6$ 下行 $+1$，上行-1；
多边形2：$L2$ 上行 $+2$，下行-2；
　　$L4$ 上行 -2；
　　$L3$ 上行 -2；下行$+2$
多边形3：$L3$ 上行 $+3$，下行-3；
　　$L6$ 上行 $+3$，下行-3；
　　$L5$ 上行 $+3$，下行-3；

至此，三个多边形填充完毕，转换结束。下面按边界弧段号排列，顾及公共边的综合赋值情况以及每个弧段上下行时对该弧段左侧栅格像元进行赋值的情况给以综合整理归纳，得出运算规律为：

弧段 $L1$　上行：左多边形编号0－右多边形编号$1=-1$
　　下行：右多边形编号1－左多边形编号$0=1$

弧段 $L2$ 上行：左多边形编号 2—右多边形编号 0=2
下行：右多边形编号 0—左多边形编号 2=—2

弧段 $L3$ 上行：左多边形编号 3—右多边形编号 2=1
下行：右多边形编号 2—左多边形编号 3=—1

弧段 $L4$ 上行：左多边形编号 1—右多边形编号 2=—1
下行：右多边形编号 2—左多边形编号 1=1

弧段 $L5$ 上行：左多边形编号 3—右多边形编号 0=3
下行：右多边形编号 0—左多边形编号 3=—3

弧段 $L6$ 上行：左多边形编号 3 — 右多边形编号 1 = 2
下行：右多边形编号 1 — 左多边形编号 3 = —2

由此得到按边界代数法进行矢量向栅格转换的基本思想：在转换过程中，沿每幅图中的全部具有左右多边形编号的边界弧段按其前进方向进行逐个搜索。当边界线上行时，将该边界线左侧的全部栅格像元赋值 = 左多边形编号—右多边形编号；当该边界线下行时，则对该线左侧的栅格赋值 = 右多边形编号—左多边形编号。搜索过程中不必顾及各边界弧段的排列顺序。图 3-26 给出两个多边形转换过程的示例。

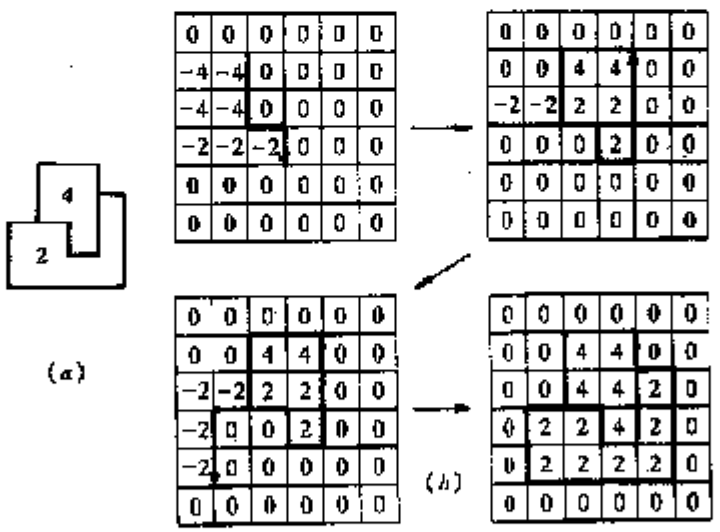


图 3-26 两个多边形转换示例

边界代数法是根据组成多边形的各边界弧段的拓扑信息，通过简单的加减代数运算将拓扑信息动态地赋给区域内的各个栅格像元。整个转换过程对每个栅格来说是"全息"的，搜索完全图的多边形边界后，也就完成了所有多边形的转换。边界代数法与前述的几种方法不同，它不是逐点搜寻判断边界，也不需要考虑边界与搜索轨迹的关系，因此具有算法简单、可靠性好的特点。此外，由于该法仅采用加减代数运算，每条边界弧段仅计算一次，省去了对公用边的重复计算，又无需考虑边界存放的顺序，因此还具有运算速度快、受计

算机内存限制小的优点。

2. 栅格数据向矢量数据的转换

栅格数据向矢量数据的转换实际上就是将具有相同属性代码的栅格像元集合表示为以边界弧段以及边界的拓扑信息所确定的多边形区域。而每个边界弧段又是由一系列小直线所组成的矢量格式边界线。通常这种转换包括对多边形边界的提取、多边形边界的跟踪、拓扑关系的生成和去除冗余点及曲线光滑等步骤。上述步骤是为了使栅格数据中包含的空间实体之间的拓扑关系和固定的属性代码在转换过程中仍保持其原有关系和原代码，保证数据转换的真实性和一致性。所谓多边形边界提取实际上是通过确定边界点和结点来实现的；边界线跟踪则是根据已经提取的结点或边界点，判断跟踪搜索方向后，逐个边界弧段地进行跟踪；拓扑关系生成则是将原栅格数据含有的边界拓扑关系转换为矢量拓扑数据结构并建立与属性数据的联系；去除冗余点以及曲线光滑是为减少数据冗余，将因逐点搜索边界点造成的多余点去掉，并采用一定的插补算法对由于栅格精度限制造成的边界曲线不圆滑进行处理。

下面介绍一种双边界直接搜索算法。其基本思想是通过边界提取将边界弧段左右多边形的拓扑信息保存在边界点或结点上，在对边界跟踪搜索时采用 2×2 栅格阵列作为窗口，顺序沿行（或列）方向对整个栅格阵列进行全图搜索，根据当前窗口内的四个栅格代码值的结构模式可以确定下一个窗口的搜索方向以及与被搜索边界的拓扑关系。具体步骤可分为三个阶段：

(1) 边界点和结点的提取

对于一个 $m\times n$ 的栅格图像阵列，采用 2×2 栅格阵列作为窗口顺序沿行、列方向对全图进行扫描。若当前窗口内的四个栅格像元的代码值相同，则表示它们属于同一区域，不是边界点；若当前窗口内的四个栅格值有且仅有两个不同的值，则该窗口内的四个栅格可确定出边界点。这说明该点位于以这两个值为编号的多边形边界上。为此可以将这四个栅格作为确定边界点的标识，并保留各栅格的原属性代码值；若窗口内四个栅格中有三个以上不同编码，则可以确定出结点，说明该点是三条以上不同边界弧段的交汇点，为此可用这四个栅格作为确定该结点的标识，并保留各栅格原代码值。若窗口中四个栅格出现对角线上两两相同的情况，说明该处多边形不连通，此时仍将这四个栅格确定的点当作结点处理。图 3-27、图 3-28 所示分别为结点和边界点的几种结构模式。

(2) 边界线搜索与拓扑信息的生成

边界点和结点提取后，即可在此基础上进行边界线搜索。边界线搜索是逐个弧段进行的，对每个弧段由一个结点开始，选定与之相邻的任意一个边界点或结点进行搜索。首先记录边界点的两个多边形编号作为被搜索边界的左右多边形，而搜索方向则由进入当前点的方向和当前点下一步要走的方向来确定。因此每个边界点只能有两个走向：一个是前点的进入方向，另一个是要搜索的后续点方向。如图 3-28 (3) 所示的边界点只能有两个走向，即向下方和右方。若该边界点是由下方搜索得到的，也就是说，若其前点位于它的下方，则该点的搜索方向只能是右方，该边界弧段的左右多边形编号应分别为 a 和 b；反之，如果该点是被其右方的点搜索到，即右点是该点的前点，则后续搜索方向应确定为下方，此时，该

边界弧段的左右多边形编号应分别为 b 和 a。其它情况依此类推。由此可见，这种结构可以唯一地确定搜索方向，从而大大地减少搜索时间，同时形成的矢量结构带有左右多边形编号的拓扑信息，容易建立拓扑结构和与属性数据的联系，利于提高转换效率。

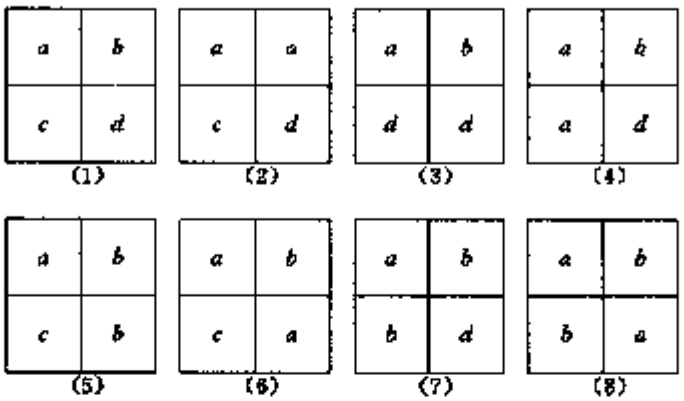


图 3-27　结点的八种结构

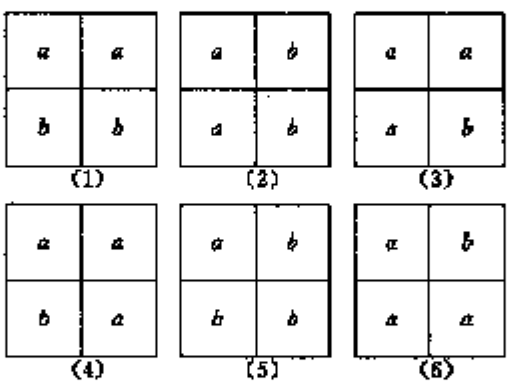


图 3-28　边界点的六种结构

(3) 去除冗余点和曲线光滑

在进行边界搜索时，由于是沿边界逐点搜索，当遇到边界弧段是直线的情况时，就会产生多余点，从而造成数据冗余。为此，需要除去这些多余点记录。去除多余点的基本思想是根据解析几何中的直线方程来确定需要去除的点。在一个边界弧段上的连续三个点 (x_1, y_1)，(x_2, y_2) 和 (x_3, y_3)，如果在一定的精度范围内可以认为它们是处于一条直线上，即满足直线方程，则三个中间的一点可以被认为是多余，应予以去除。有直线方程：

$$(x_1 - x_2)/(y_1 - y_2) = (x_1 - x_3)/(y_1 - y_3)$$

或

$$(x_1 - x_3)/(y_1 - y_3) = (x_2 - x_3)/(y_2 - y_3)$$

由于在算法中要尽量避免出现除零的情形，上式可转化为：

$$(x_1 - x_2)(y_1 - y_3) = (x_1 - x_3)(y_1 - y_2)$$

或

$$(x_1 - x_3)(y_2 - y_3) = (x_2 - x_3)(y_1 - y_3)$$

所以，只要上式成立，则 (x_2, y_2) 点为多余点，可予以去除。

通过边界搜索产生的边界线由于栅格精度的限制，可能不够光滑，通常可以采用一些插值算法进行光滑处理，常用的算法有：

①线性迭代法；②分段三次多项式插值法；③正轴抛物线平均加权法；④斜轴抛物线平均加权法；⑤样条函数插值法。

§3-5 GIS 数据模型

GIS 涉及的数据量很大，而且数据之间的联系又是错综复杂，因此，GIS 数据库的数据组织是一个核心问题。数据组织的好坏将影响到 GIS 效率的高低以及用户对 GIS 的使用。通常，数据库技术在处理数据库组织时，是从全局出发，对数据内部的联系和用户要求进行综合平衡来考虑。在设计数据库全局逻辑结构时，现有的数据库管理系统通常使用三种方法对数据构成三种不同形式的数据库数据结构，即

(1) 层次方法：将数据库数据按层次结构的形式进行组织；

(2) 网状方法：将数据库数据按“有向图”的结构形式组织；

(3) 关系方法：将数据库数据按关系形式，即表格结构的形式进行组织。

对于不同形式的数据库数据结构方法，对应有不同类型的数据库数据的操作集合，以及为了保证数据库数据的完整性规定的完整性规则集合。一般情况下，数据库数据模型就是由上述的数据库数据结构、数据库操作集合和完整性规则集合三者组成。所谓模型是人们对客观世界的认识和理解，是对客观现实系统的近似描述。所以数据模型又称为数据库结构，它反映了数据库中数据的整体逻辑组织，反映了实体之间的逻辑联系。

GIS 数据模型是描述 GIS 的数据内容和数据之间的联系工具，它是衡量 GIS 数据库能力强弱的一个主要标志，也是数据库系统的基础。GIS 数据库主要涉及对图形数据和属性数据的管理和组织，它是以一定的组织方式将互相关联的数据集合存贮在一起，并且能够以最佳方式和最小的数据冗余为多种目的服务。另一方面，由于计算机处理的数据既可以存放在数据文件里，也可以存放在数据库中，例如目前很多 GIS 系统就是用文件系统管理空间数据（图形数据），用数据库系统管理属性数据，所以也可以说数据库就是一些相互关联的存贮文件的集合。

目前在数据库领域中，流行的数据模型主要有三种，即层次模型、网状模型和关系模型，除此之外还有最近兴起的面向对象模型。需要说明的是，所谓关系数据库管理系统，就是说数据库管理系统所支持的数据模型是关系的，或者说该数据库管理系统采用的是关系方法。类似地，还有网状数据库管理系统和层次数据库管理系统。从目前数据库技术的状况和发展趋势来看，关系数据库将占主导地位。另外，面向对象技术用在数据库设计领域后，受到人们的重视，有趋势表明可能将发展成下一代数据库的主流。

一、层次模型

用树状结构来表示实体之间联系的模型称为层次模型。它是以结点来表示数据库中的

记录类型的有向树。如图 3-29 所示为一个简单的多边形地图，按层次方法可将图 3-29 构造成树结构的层次模型，如图 3-30 所示。

由图 3-30 可看出，层次模型揭示的是实体（记录）之间的一对多（1：n）的联系。通常把表示 1 的实体放在上方，称为“父结点”；而将表示 n 的实体放在下方，称为“子结点”；最上层只能有一个结点，称为根结点。为符合 1：n 的联系，除根结点外，其它的结点都有且仅有一个“父结点”，但是每个父结点可以对应于多个子结点。最下层的末端结点也称为叶结点。

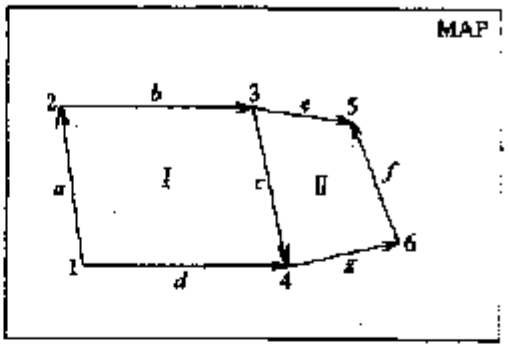


图 3-29　地图 MAP 及其多边形实体 I 和 I

在层次模型中，检索是向从根开始的某条路径提出询问，依照从左到右的顺序检查路径。图 3-30 是用层次模型表示的简单多边形地图数据库结构，该结构按多边形、弧段和结点分为不同层次，如查询结点 6，则从根结点开始，可循多边形 I —弧段 f—结点 6 的路径搜索，层次关系非常清晰。

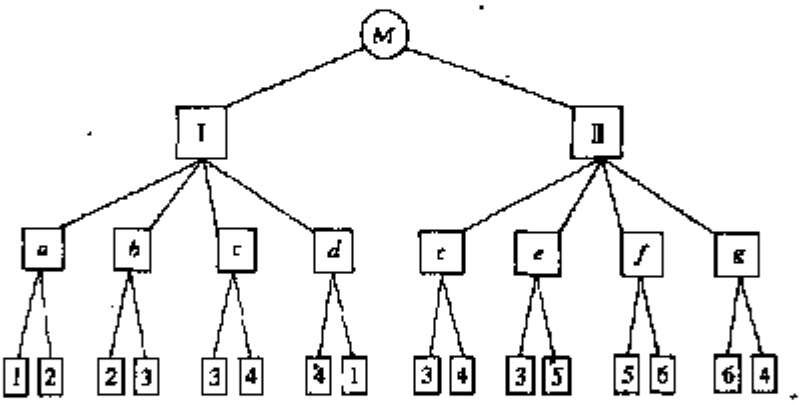


图 3-30　层次模型

如上所述，可得出层次模型的两个限制条件：

(1) 有且仅有一个结点无父结点，即根结点；

(2) 除根结点之外，所有结点有且仅有一个父结点。

这就使层次模型不能直接表示实体之间多对多（m：n）的联系。若需要用层次模型表示 m：n 联系，就必须设法先将该联系分解为 1：n 联系后，才能用其表示。这是令人遗憾的缺陷。在 GIS 中，若采用这种层次模型将难以顾及公共点、线数据共享和实体元素间的拓扑关系，导致数据冗余度增加，也给拓扑查询带来困难。归纳起来，层次模型的优点是容易理解，单码查找速度快，易于更新和扩充；但是多码查找比较困难，一般需要较大的索引文件，所以产生数据冗余。

二、网状模型

如果取消层次模型中的两个限制，即允许有多于一个以上的结点没有父结点，且允许

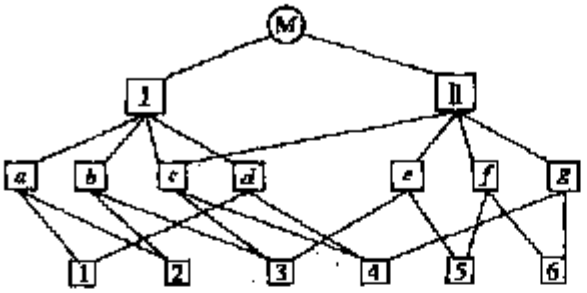


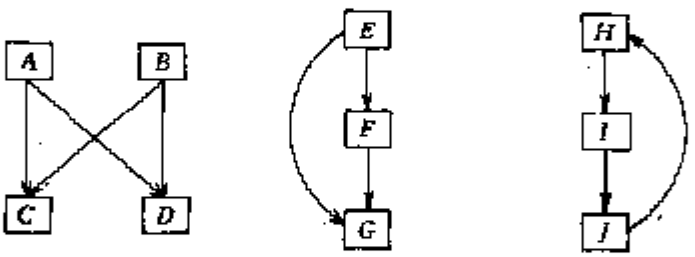
图 3-31 网状模型

每个结点可以有多个父结点，则树形结构便可形成网络形式。因此，网状模型就是用网络结构来表示实体之间联系的模型。理论上，由于网状模型没有层次模型的那两点限制，所以可以直接表示 $m:n$ 联系。但是在实际应用上，一些已经实现的支持网状模型的数据库管理系统上，对直接处理实体之间的 $m:n$ 联系作了限制。

网状模型和层次模型在本质上是一样的。从逻辑上看它们都是用连线表示实体之间的联系，用结点表示实体集。如图 3-31 所示为图 3-29 的网状模型，从物理上看，层次模型和网状模型都是用指针来实现两个文件之间的联系。区别仅在于网状模型中的连线更加复杂，更加纵横交错，从而使数据结构更加复杂。

由此可知，网状模型是以有向图表示的网络结构，每个结点仍然表示数据库中的一个记录类型（实体）。网状模型的有向图与层次模型的有向树比较，其区别在于：

(1) 可以有零个或多个结点无父结点（图 3-32 (a)）；



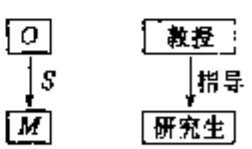
(a) (b) (c)

图 3-32 网状模型示例

(2) 至少有一个结点有多于一个父结点（图 3-32 (a, b)）；

(3) 允许两个结点之间有两种或多种联系（图 3-32 (c)）。

网状模型中，结点之间的联系是任意的，为区分这些结点之间不同的联系，对这些联系要进行命名，通常将其称为“系型”(Set Type)。系型表示各结点之间，即记录类型之间的联系。例如，记录型 O 和记录型 M 之间的联系命名 S，则系型可表示为图 3-33 (a) 所示；若将该关系用在教授与研究生之间的指导关系，则可表示为图 3-33 (b) 所示。



(a) (b)

图 3-33 系型示例

在基于矢量结构的 GIS 中，图形数据通常采用拓扑数据模型，这种模型非常类似于网状模型，只是在拓扑模型中一般采用目标标识来代替网状结构中的指针。

图 3-34 是用网状模型表示自然要素、社会经济要素和地理位置要素之间联系的示例。

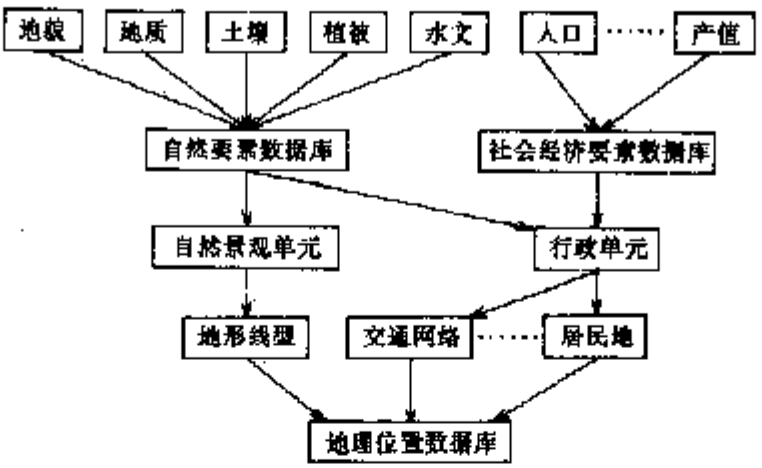


图 3-34　网状模型示例

总之，网状模型较层次模型扩充了对实体之间联系的限制，可以较灵活地表示实体之间的多种关系，对确定的数据表示效率较高，数据冗余也较小，适合于表示关系较复杂的地理数据和具有网络状特征的地理实体，但网状模型的数据指针比较复杂，数据更新也较为繁琐。

三、关系模型

表格是人们熟悉的数据表示方法。用表格数据表示实体和实体之间联系的模型称为关系模型。

在层次模型和网状模型中，文件中存放的是数据，各文件之间的联系是通过指针（连线）来实现的。而在关系模型中，文件中存放两类数据：一是实体本身的数据，二是实体间的联系，这里的联系是通过联接字段实现的。

关系模型源于数学，它把数据看成是二维表格中的元素，而这个二维表格就是所谓关系。表中的每一行代表一个记录或一个取样；每一列称为关系的一个属性集，列可以命名，称为属性名，或数据项类型。显然，关系是记录（行）的集合。如果表格中有 n 列，则称该关系是 n 元关系。

一个实体可由若干关系组成，而关系表的集合就构成关系模型。对这种数学化的模型，每个关系应满足下列条件：

(1) 表中的每一列属性都是不能再分的基本字段；

(2) 各列被指定一个相异的名字；

(3) 各行（记录）相异，不允许重复；

(4) 行、列次序无关。

综合以上四点可知，一个关系是一个概念文件，该文件中的每个记录是唯一的，所有记录具有相同个数和类型的字段，也就是说，所有的记录有同样的固定长度和格式。图 3-

29 所示的多边形地图，可用下列关系表示出多边形实体与其边界弧段实体和结点实体之间的关系。如图 3-35 所示各实体之间的联系可以通过连接字段来实现。例如关系 1 中边界弧段号 c 可与关系 2 中相同字段边界弧段号 c 进行连接，同样道理，关系 2 中结点号 3 也可与关系 3 中结点号 3 进行连接。

关系模型的最大特色是对实体描述的一致性，上述示例中用连接字段实现各实体之间联系正说明这一点。关系模型正是利用数据本身通过公共值隐含地表达它们之间的联系，并且可采用关系代数和关系运算来操作数据。在关系模型下，实体及其实体之间的联系均采用关系来描述，通过关系之间的连接运算还可建立新的关系，对关系数据库的查询和统计操作也都可以采用关系运算得以实现。此外，关系模型还具有结构简单、灵活、数据修改和更新方便、容易维护等特点。所以它是当前数据库中较常用的数据模型。目前，很多 GIS 中的属性数据都是采用关系数据模型，甚至某些关系还采用关系数据库管理系统来管理空间数据。

关系 1：多边形边界关系表

多边形编号	边界弧段号	边长
I	a	30
I	b	32
I	c	28
I	d	36
II	c	28
II	e	20
II	f	30
II	g	25

关系 2：边界数据结构表

边界弧段号	起点号	终点号	左多边形号	右多边形号
a	1	2	0	I
b	2	3	0	I
c	3	4	II	I
d	1	4	I	0
e	3	5	0	II
f	6	5	II	0
g	4	6	II	0

关系 3：结点坐标关系表

结点号	x 坐标	y 坐标
1	$x1$	$y1$
2	$x2$	$y2$
3	$x3$	$y3$
4	$x4$	$y4$
5	$x5$	$y5$
6	$x6$	$y6$

图 3-35　关系模型示例

但是GIS中的关系模型在效率、数据语义、模型扩充和目标标识等方面还存在一些问题，特别是在处理空间数据库所涉及的复杂目标方面，传统的关系模型显得难以适应，此外，关系数据库系统的管理也较为复杂，查找速度与网状和层次模型相比也要慢一些。

综上所述，数据库理论是GIS的理论基础，尽管目前流行的通用数据库管理系统在对空间实体的定义、描述和检索运算上仍存在着很大的不足，进行GIS的技术设计时，完全采用通用数据库管理系统远非理想；但是，由于通用数据库管理系统在数据定义、结构扩充、数据更新和数据运算上效率高、通用性较强，因此目前很多GIS仍倾向于采用通用数据库管理系统作为其数据管理的支持系统。而与此同时，更适合于GIS需要的具有最小冗余、最大灵活性以及最快速的空间型数据库管理系统也正在发展之中。

第四章　GIS的基本功能

GIS技术在全球范围迅速发展及其具有的强大生命力是与其对地学信息的处理功能强大而紧密相联的。简言之，GIS就是通过对地学信息和有关社会经济信息进行采集、编辑、数据管理、查询、分析和输出等工作，来实现对地学信息的计算机化管理过程。GIS作为一个空间信息系统要求至少具备五项基本功能，即数据输入、图形与文本编辑、数据存贮与管理、空间查询与空间分析、数据输出与表达。

§4-1　数据输入功能

建立GIS的地理数据库是一项最重要最复杂的核心任务。这项工作是将地面上的实体图形数据及其描述性属性数据输入到数据库中。所以，数据输入就是对数据进行必要编码和写入数据库的操作过程，此过程又称为数据采集。此外，由于GIS可以有多种数据源，所以，对不同的数据格式的数据要进行格式转换，以保证数据格式的一致性。

通常，GIS主要考虑以下三个方面的数据输入问题：

(1) 统一的地理基础；

(2) 位置数据（空间数据）：定义地面实体特征相对于某一坐标系统所处的空间位置；

(3) 属性数据（非空间数据）：定义地面实体特征所表示的内容。

一、统一的地理基础

地理基础是地理信息数据表示格式与规范的重要组成部分，它主要包括统一的地图投影系统、统一的地理坐标系统以及统一的地理编码系统。各种来源的地理信息和数据在共同的地理基础上反映出它们的地理位置和地理关系特征。

地理信息系统之所以区别于一般的信息系统，就在于它所存储记录、管理分析、显示应用的都是地理信息。而这些地理信息都是具有三维空间分布特征且发生在二维地理平面上的，因而它们需要有一个空间定位框架，即共同的地理坐标和平面坐标系统。所以说统一的坐标系统是地理信息系统建立的基础。

通常，地理信息系统中的投影坐标系统配置具有以下一般性特征：

(1) 各个国家的地理信息系统采用的投影系统与该国的基本地图系列所用投影系统一致；

(2) 各地区的地理信息系统中的投影系统与其所在区域适用的投影系统一致。

因此，对一般地理信息系统中地图投影配置的选择原则为：

(1) 投影系统应与相应比例尺的国家基本图（基本比例尺地形图、基本省区图或国家大地图集）投影系统一致；

(2) 一般地只考虑至多采用两种投影系统，一般服务于大比例尺的数据处理，另一种服

务于中小比例尺，如我国基本比例尺地形图（1∶100万、1∶50万、1∶25万、1∶10万、1∶5万、1∶2.5万、1∶1万、1∶5000）中，除1∶100万图采用兰勃特投影外，其余均采用高斯-克吕格投影为地理基础。

二、空间数据的输入

对空间数据的采集是GIS数据采集的主要功能。这里的空间数据主要指图形实体数据。通常在GIS中用到的图形数据类型包括：各种地图与地形图、航测照片、遥感数据、点采样数据等。因此，空间数据的输入主要是对图形的数字化处理过程。而输入方法可以采用数字化仪、扫描仪、摄影测量仪以及测量全站型速测仪、GPS接收机等能以数字形式自动记录测量数据的测量仪器。至于选择哪种输入方法需根据应用图形数据的方式、图形数据的类型、现有设备状况、现有人力资源状况和经济状况等因素综合考虑，可以选用单一方法或几种方法结合起来输入所需要的图形数据。下面简单介绍采用数字化仪和扫描仪进行数据输入的工作过程。

1. 用数字化仪进行数据输入

采用数字化仪对构成需要图形的点、线、面或栅格单元的坐标进行编码和输入可以大大减少用手工敲击键盘输入的大量劳动。目前，市场上数字化仪的规格按其可处理的图幅面积来划分，有A0、A1、A3等幅面。典型的用于制图的数字化仪是A0规格，其幅面为1.0 m×1.5 m。对于一般应用而言，A1幅面的数字化仪也可以满足对0.5 m×0.5 m常规地形图的数字化工作。在试验或培训等数字化工作中也可采用A3或更小幅面的数字化仪，但在对一幅地形图数字化过程中，可能需要定向几次，操作比较麻烦。

数字化仪上点的坐标是用一个被称作“游标”（或称“鼠标”）的定位装置送入计算机的。游标上有一个精确固定在窗口内的十字丝。需要数字化点的坐标时，只须将十字丝精确对准该点后，按动游标上相应的按钮即可。一般的数字化仪的游标上都设有4个也有12个或16个或更多的键。这些键可用于附加的程序控制，便于数字化操作员不必离开数字化仪去用计算机键盘输入，就可选择数字化命令和数字化内容。同时，这些键还可用来对所进行数字化的对象加入标号以便以后与有关的非空间属性数据相连接。

在进行数字化时，首先将待数字化的地图或航片等固定在数字化仪上，输入其比例尺，并用游标定出图幅上的左下角点和右上角点，确定出数字化范围，然后数字化员小心地用手扶动游标，使游标十字丝尽可能地保持在待数字化线段上，再按动游标键钮进行输入。

在一般的数字化工作中常见的有两种基本输入方式，即流方式和点方式。

所谓流方式数字化是指作业员只需仔细保持使游标十字丝沿待数字化的线段连续移动，而由计算机自动地按等时间隔或等距间隔来控制点位的数据录入。在这种方式下，由于作业员只需保持线迹跟踪而不必考虑点位选择和录入问题，因此这种方式易于操作，且数字化速度较快，工作效率也较高。缺点是若按等时录入，则采样点的疏密程度与作业员移动游标十字丝的速度有关；而若按等距录入，则对曲线弯曲复杂的部分就不能正确地反应出弯曲顶部的形状。

点方式数字化则是由数字化作业员自行选择采样点和确定采样密度，逐点地对目标进

行数字化的方法。在该方式下，对每个采样点，作业员都是仔细地用游标十字丝对好目标后，再由作业员按键取样录入数据。这种方式可以让作业员来选择最有利于表现曲线特征的点位进行数字化，因此数字化精度较高；缺点是由于对每个点都要独立地进行目标重合和单独录入，而不是像流方式那样边进行线迹跟踪移动边进行数据录入，所以数字化速度太慢，工作效率不高。

实际上，在数字化作业过程中对于大多数有经验的数字化作业员来说，采用的数字化方式既不是纯粹的流方式也不是纯粹的点方式，而是根据这两种方法的特点结合起来的一种所谓基于流方式的点方式数字化采样方法。按这种方法，作业员在沿线段连续地跟踪移动游标十字丝的同时（流方式的特点），还要根据目标特征，按作业员的主观判断随时按键录入数据（点方式的特点）。这种方法兼有流方式和点方式的优点，同时又避免了二者的缺点，因而具有工作效率高，录入速度快，而且能人为控制采样数量及采样位置，使概括综合误差的影响降低。缺点是这种数字化方式需要作业员具有一定的技术水平和经验，对初学者而言不易掌握。

数字化数据的精度受数字化仪本身的分辨率大小、数字化方式和操作员的经验和技术水平等多因素影响。由于分辨率高的数字化仪售价也较高，所以选择数字化仪自身分辨率大小时既要考虑其是否满足应用精度要求又要顾及其价格高低。例如，待数字化的原图是用 0.4 mm 绘图笔所绘制的，若选用分辨率 0.02 mm 的数字化仪就很不合算。另外，在购置数字化仪时应注意，数字化仪的实际分辨率一般比其标定分辨率要低。例如，标定分辨率为 0.025 mm 的数字化仪，经实践证明，测试时的最大偏差可以达到 0.07 mm 甚至 0.15 mm。

通过对人工手扶跟踪数字化方法的分析，概括地讲，这种方法的数字化误差可以大致归纳为如下两个方面：

(1) 线迹跟踪误差：数字化过程中，使游标十字丝在被跟踪线左右产生偏移造成的误差；

(2) 概括综合误差：由于对某线段选点的位置和数量进行选择时产生的误差。

2. 用扫描仪进行数据输入

扫描仪是 GIS 图形及影像数据输入的一种重要工具。从当前数字化工作发展趋势看，有可能采用扫描仪来把现有的纸质地图和图像、影像转换成数字地图而取代繁重枯燥的手扶跟踪数字化仪的工作。

扫描仪按其扫描方式可分成以栅格数据形式扫描的栅格扫描仪和直接沿线划扫描的矢量扫描仪。

栅格扫描仪的工作原理并不复杂。设地图上任意部分都是由黑白两色组成。将装有电荷耦合器件 (CCD) 的高分辨率镜头的摄像机与扫描装置以及低功率激光源组合在一起，对数字化对象进行扫描后，CCD 将摄取的影像分解为离散的像元，然后将其记录成“黑”或“白”的像元。当然，这样得到的数据还需进行大量的处理工作后才能变成有用的数字地图要素，如等高线、道路等。

在扫描工作中，选取像元尺寸大小不仅要考虑需满足扫描数据的精度，同时还要考虑扫描数据处理量的大小。通常像元的大小可在 20 μm 至 100 μm 之间变化。即使这样，目前

采用扫描仪进行数字化工作的数据量仍是十分巨大的。例如，处理一幅中等复杂程度的常规地形图，采用 VAX-780 小型机对等高线进行扫描也需要近 20 MB 的存储空间。

除了上述的数据处理量大的问题以外，栅格扫描数据还存在噪声和中间色调像元的处理问题。噪声是指那些不属于地图内容的斑点污渍和其它模糊不清的东西形成的像元灰度值。噪声灰度的范围较广，很难用一种简单有效的方法加以清除。中间色调（灰度）是由于在扫描过程中，光孔或一个 CCD 元件没有被地图上的线划完全覆盖住，从而产生的介于“黑”和“白”之间的灰度值，这种现象常常使数字化的线划变成一条很粗的带状。

矢量扫描仪是直接跟踪被扫描图上的线划并直接产生矢量数据的扫描仪器。目前用得较多的是激光扫描。这种方法是通过把扫描仪和绘图仪结合在一起，利用光电转换的方法对透明膜片上的线划进行数字化扫描或在膜片上绘出线划和文字。

在矢量扫描过程中，操作员将地图的透明膜片复制品投影到屏幕上，然后由操作员用光标引导激光束进行扫描。例如，当要数字化某曲线时，将激光束引导到该曲线的起点后，激光束能够自动沿线移动并记录其坐标，碰到连结点或扫描闭合曲线时碰到起始点就自动停止移动，等待操作员进行下一步的引导。一旦一条线扫描完成后，就由另一激光束在屏幕上绘出该线，然后操作员就可在该线上加入一个识别符供以后连接属性用。

该法的优点是能够即时完成对扫描对象的扫描工作，并且得到的数据直接变成符合比例尺要求的矢量数据。其不足之处在于，在扫描过程中操作员必须进行大量的引导工作，此外，还要事先制作数字化材料的透明膜片，而且要保持清洁。

二、属性数据的输入及其与空间数据的连接

1. 属性数据的输入

属性数据是用来描述空间数据特征性质的。例如，一栋房子除了需要记录它的位置坐标等空间数据以外，还需要存贮它的属性信息，如房主、房屋面积、建筑日期等。因此，这种非空间的属性数据也可以被称为空间实体的特征编码。很显然，属性数据是与空间实体相关的。通常，可以采用公共识别符的方法建立属性数据与空间数据的有效联系，从而有效地存贮和处理这些属性数据。

某些信息系统在进行属性数据输入时与空间数据一起用键盘进行输入，把属性数据与空间数据组织在数据文件的同一个记录中。这种输入方法和存贮方式，虽然可以在一个记录中同时反映出空间位置及其特征信息，但是当数据量很大时这种记录方式在数据管理过程中就显得很不灵活，同时又会造成很大的数据冗余，从而使数据处理时间增加、降低系统的效率。

还有一些系统把属性数据以单独的数据文件方式与空间数据文件并存于文件系统中。这种管理方式，对于某些具体应用可能是简单实用的，但局限性很大，结构不灵活，难以实现数据共享。

因此，一般而言，GIS 要求其数据结构应满足通用性的要求，具有使用方便、灵活性好、冗余度小、管理程度高、逻辑操作性能强等特点。从地理空间数据的实际考虑，由于涉及到的空间数据往往关联诸多的属性数据项目，大量的属性数据通常就采用关系型数据库与空间数据分别存贮的方法，也可通过公共识别符或者建立一个程序将空间数据与属性数据

连接起来。尤其是当数据量较大时，一般都是将空间数据的输入与存贮分别进行的。采用的方法是首先将属性数据输入一个顺序文件，经编辑、检查无误后再转存数据库的相应文件或表格中。

2. 与空间数据的连接

如上所述，在输入空间数据时虽然可以直接在图形实体上附加特征编码，但是当数据量较大时，这种交互式输入的效率就太低了。因此，可以用特定的程序把属性数据与已数字化的点、线、面空间实体连接起来。这样只要求空间实体带有唯一性的识别符即可，除了用程序自动生成以外，识别符也可以用手工来输入。

通常，属于一个空间实体的属性项目可能有很多，所以也可以将其放入同一个记录中，而该记录的顺序号或者是某一特征数据项可作为该记录的识别符。该识别符与所对应的空间数据的识别符一起构成了它们之间相互检索的联系纽带。

§4-2 图形与文本编辑功能

在GIS的数据输入过程中，通过各种输入设备采集到的数据难免会产生或引入一些差错，比如，使用数字化仪进行手扶跟踪数字化过程中，数字化操作员的经验和技术水平的不同及在线迹跟踪过程中受到的心理性和生理性因素的影响和局限性都可能会产生数字化误差。即使是扫描得到的数据，也可能会有一些噪声斑块或线条出现。所以，一般要求对GIS中的空间数据和属性数据应具备图形与文本的编辑功能，以修正所出现的错误。

通常，大多数GIS的数据编辑都是比较耗时的交互式处理过程。在编辑过程中，除了要逐一修改所能发现的数据错误之外，还要进行对图形的合并或分割，数据的更新等工作。而这些编辑工作可以通过把数据显示在屏幕上，然后利用键盘或数字化图板来控制数据编辑的各项操作活动。

一、对空间数据（图形数据）的编辑

对图形数据进行编辑，一般应要求系统具备图幅定向、文件管理、图形编辑、生成拓扑关系、图形修饰与几何计算、图幅拼接、数据更新等功能以及设计一个用户界面友好的人机对话窗口。

1. 图幅定向

数字化过程中一项重要内容是对待数字化材料进行图幅定向工作，将图幅坐标归化为地理坐标。因此，要求图形编辑系统应具有自动完成数字化仪坐标→地图坐标→屏幕坐标的转换功能，并且可以设置定向允许限差值，能够修改或删除定向点，输入点坐标，对定向点进行数字化以及可以进行图幅定向平差并显示平差结果。

2. 文件管理

图形编辑中的文件管理是指对图形文件的读写功能。通常，在这一功能菜单中应包括：创建新文件、打开已有文件、添加文件、存贮文件与更改文件名、输出图形文件等基本项目。

3. 图形编辑

这一功能要求对图形数据不仅可以进行逐点或逐线段地增、删、改操作，还可以对图形进行开窗、缩放、移动、旋转、裁剪、粘贴和拷贝等操作。同时，在图形编辑中还应能对图形进行分层显示；能针对所开窗口的任一位置，输出其地理坐标或相对于某参考系的坐标值；并且还应具备一种情除功能，以便能将各种不必要的改动清除掉。除上所述，由于对 GIS 中矢量数据的处理主要是以点和线为基本对象（面要素是以线要素为基础进行处理的），所以在编辑中还应具有对线段的特定处理功能。比如，修改或删除某线段的一部分或全部，线段的连接与断开，产生平行线，使曲线光滑以及镜面反射等功能。

4. 生成拓扑关系

能够自动使矢量数据生成拓扑关系是 GIS 与一般数字测图系统的主要区别之一。通常建立拓扑关系可以根据相应的结点和弧段经过编码由计算机自动组织成 GIS 中的线状或面状地物。例如，可由相应多边形的内部点组成的文件得到某多边形边界的左右多边形信息，从而建立线与面的扑拓关系。在图形编辑中建立拓扑关系则是使用鼠标人工装配地物或修改已建立的拓扑关系。尤其是在处理复杂地物的情况下，对图形数据的编辑中应能够让用户自行定义复杂地物，分解复杂地物，删除或显示复杂地物以及作多边形填充的能力。

5. 图形修饰与几何计算

编辑地图时需要根据不同的地物类型，设置不同的线型、颜色和符号，还应具有注记的功能，包括设置字体大小、方向和注记位置。此外，用户有时还要根据需要自己设计一些特殊的符号存入符号库中，以便在地图整饰时调用。所以系统应具备线型设置、颜色设置、注记设置和符号选择等功能，同时还应可以进行创建、编辑和存贮符号的操作。

除上述功能外，系统还应能通过几何坐标计算多边形的面积、周长、结点间距离、线段长度、支距极坐标等。

6. 图幅拼接

在数字化过程中，由于人为的或（和）仪器的误差影响致使两个相邻图幅的数据库在接合处产生不一致，即所谓“裂隙”。这种“裂隙”有两种情况。一种是所谓“几何裂隙”，另一种是“逻辑裂隙”。“几何裂隙”是指数据库中的边界数据所分开的一个实体的两部分不能精确地彼此衔接。比如，数据不完整、重复、位置不正确、变形、比例尺问题等。“逻辑裂隙”则是指某个实体在一个数据库中有性质 A，而在另一个数据库中却具有性质 B；或者同一实体在两个数据库中具有不同的附加信息，例如，同一条公路具有不同的宽度，空间与属性数据连接有误，属性数据不完整等问题。因此，在图形编辑系统中应定义相邻图幅的接边范围，开窗检索接边范围内的图形及属性信息，将跨接图幅的目标，根据接边方向按横坐标或纵坐标进行排序，对逻辑相同的实体在几何上自然连接，以消除接边误差。对于存在的逻辑上不一致的实体，则应先用交互编辑方法达到逻辑上的一致性。

7. 数据更新

为保证数据库的现势性必须进行数据更新，以防止数据的老化和不完美。这要求在对数据编辑中，能确定出数据库中那些正发生变化的数据以及变化的程度。

为了使上述 7 项编辑基本功能充分发挥出性能，系统还要求设计一个用户友好的界面

以保证使人机交互操作顺利而简便地得以实现。因此，这样的用户界面应至少满足以下3点要求：

(1) 图形编辑过程应采用下拉弹出式菜单驱动方式。为充分利用有限的屏幕空间，菜单的弹出和取消应不破坏原有画面。此外，一个开放性系统，还应具有用户自定义菜单的功能，在子菜单内直接扩充应用模块；

(2) 界面应能向用户提供实时在线帮助功能（HELP功能），使用户可以随时通过HELP功能得到更详细的操作说明，以帮助用户理解和处理各种可能出现的问题；

(3) 界面应能对用户的任何操作（包括错误的操作），从显示屏幕上都能作出相应的响应；应该对用户执行的任何操作都有简单明确的提示说明；并且还有允许用户反悔的UNDO功能。

二、对属性数据（非图形数据）的编辑

图形和属性数据分别管理可以提高操作效率和数据管理的灵活性。属性数据库模式可根据任务的性质给予任意定义和修改。因此为建立属性数据与空间数据的联系，需要在图形编辑系统中，设计属性数据的编辑功能，将实体的属性数据与相应的空间数据（如点、线、面）进行连接。

一般而言，属性数据较为规范，适应于采用表格形式表示，所以许多GIS都采用关系数据库管理系统管理属性数据。通常的关系数据库管理系统（RDBMS）都为用户提供了一套功能很强的数据编辑和数据库查询语言，即通常所说的SQL（结构化查询语言）。系统设计人员可以适当组织SQL语言，建立友好用户界面，以方便用户对属性数据的输入、编辑与查询。表4-1为一个属性数据库管理模块的基本功能。

表4-1 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	计算平均值
添加文件	拷贝结构	删除数据	显示几行属性	统计最大值
存贮文件	删除结构	修改数据	显示几列属性	统计最小值
打印文件	合并结构	合并数据库	条件查询	报表设计
Dos Shell	查询结构信息	移动数据	多表连接查询	图示分析
退出系统		行拷贝数据		打印报表
		列拷贝数据		

除文件管理外，属性数据库管理模块的主要功能之一是用户定义地物种类的属性数据结构。由于GIS中种类地物的属性特征不同，描述它们的属性项及值域亦不相同，所以系统还应提供用户自定义数据结构的功能。比如表4-2、表4-3所示为用关系表格结构表示的

属性数据结构，使用户易于理解和掌握，也便于制表输出。

表 4-2 属性数据关系结构表

类别码	属性 1	属性 2	…	属性 n
N_1	A_{11}	A_{12}	…	A_{1n}
N_2	A_{21}	A_{22}	…	A_{2n}
…	…	…	…	…
N_m	A_{m1}	A_{m2}	…	A_{mn}

表 4-3 属性数据文件结构表

文件类型	字段名	类型	长度	小数位
多边形	类别码	字符型	5	
	面积	数字型	10	6
弧 段	类别码	字符型	5	
	长度	数字型	10	6
点	类别码	字符型	5	

通常系统还应该具备修改结构的功能，以便对数据结构设计不合理之处进行修改。

数据结构设计完成后，用户既可以在图形编辑系统中输入属性数据，也可以在属性数据库管理系统中输入属性。这两种途径输入的数据应该在同一数据库中，只是在不同数据处理模块和界面中进行。如表 4-1 所示，一般来说，在属性数据编辑模块中编辑属性数据比在图形编辑系统中的功能更强。因为属性库管理模块不仅可以作数据输入、删除、修改等基本编辑，还具有对数据进行移动、行拷贝、列拷贝以及合并等操作。

此外，采用属性数据库管理模块不但具备像图形编辑系统中根据图形目标查询属性的功能，而且还能够借助 SQL 语言提供的丰富的查询语言进行多种灵活的数据库查询。用户可以浏览属性库的内容，或单独显示某几行属性数据，或某几列属性，另外还可以操纵多个表进行跨表连接查询。

属性库管理模块还提供了统计计算和分析报表的功能，不仅可以计算平均值、最大值、最小值等，还可以按一定要求建立报表，提供用户进行分析之用。

§4-3 数据存贮与管理功能

在 GIS 中对数据的存贮管理主要是通过数据库管理系统来完成的。GIS 数据库不同于一般的数据库，它具有数据量特别大，空间数据与属性数据之间不可分割的联系，数据应

用面广等特点。为此，GIS 的数据库应做到数据集中管理，数据冗余度小，数据与应用程序相互独立，选择合适的数据模型和有数据保护措施。

一、数据存贮

地理对象通过数据输入和编辑以后，送入到计算机的存贮装置上。如前所述，GIS 是数据密集型系统，如一张中等密度地形图一般就需要 1 MB 以上的存贮容量；一幅 1024×1024 的遥感图像，一个波段就需要 1MB 的容量。因此，如果没有高密度存贮介质和较快传输速度，GIS 很难做到高效率运转。

磁带、磁盘和应用激光技术的光盘是目前计算机的主要存贮装置。磁盘是一般计算机必备之存贮设备。磁盘分硬盘和软盘两种，硬盘的存贮容量和存取速度比软盘大得多。目前微机用的硬盘容量已达到 1GB 以上，并且在实际工作中还经常使用大容量的外接硬盘。现今的 5 英寸软盘容量达 1.2 MB，3.5 英寸软盘可达 2 MB。但是考虑到用磁盘作数据交换处理的风险性以及永久放置的磁盘存贮系统的费用较高，因此也常用磁带来存贮那些需长期存贮的数据。

磁带存贮也是数据存贮的一种重要方式。目前使用的磁带一般为 0.5 英寸宽的多磁道磁带，记录密度为 1000 比特/英寸至 6250 比特/英寸。外观上看，目前计算机磁带有两种类型，一种是通常见的圆盘磁带，另一种是形似于录像带的盒式数据流磁带。一盒数据流磁带具有 150 MB 以上的存贮容量。需要说明的是，尽管磁带存贮方式在存贮容量和存贮成本以及存贮风险上要比磁盘存贮方式为优，但是在数据的交互特性方面确不如磁盘存贮方式。

激光技术的应用使计算机的存贮容量有了较大突破。目前市场上销售的只读光盘（CD-ROM）的存贮容量可达 600 MB 以上，且制造成本仅几元人民币。CD-ROM 驱动器的发展也很快，已从最初的单速（数据传输率 150 K 字节/秒），发展到双倍速（数据传输率 300 K 字节/秒）、3 倍速（数据传输率 450 K 字节/秒）、4 倍速（数据传输率 600 K 字节/秒）等系列产品。目前计算机市场上的 4 倍速 CD-ROM 驱动器也已成为低档产品，数据传输率达 900 K 字节/秒的 6 倍速和数据传输率达 1.2 M 字节/秒的 8 倍速 CD-ROM 驱动器已批量生产，正逐步成为市场主流。据报道，美国等一些主要 CD-ROM 驱动器生产厂家已在研制 10 倍速和 12 倍速的光盘驱动器，可望于近年内推出。

最理想的真正有价值的光盘驱动器应是可擦写的。目前市面上可以见到的一次写入型 WORM（Write Once，Read Many）和可擦写型 EDRAW（Erasable Direct Read After Write）光盘驱动器多属相变光盘 PCD（Phase Change Disk）驱动器。PCD 采用光学技术来读写光盘，利用激光加热记录介质，改变相变记录介质的晶体状态，用结晶状态和非结晶状态来区分“0”和“1”信息，达到读写的目的。目前，这种光盘驱动器的价格较贵，所以其用户不如 CD-ROM 驱动器用户普及，但是由于 PCD 技术的价值所在，国外许多公司，如日本松下电器公司、北美飞利浦电子公司以及 IBM、HP、SONY 等都在积极开发这种相变媒体技术。随着计算机硬件更新速度的加快，可以预见在近几年内，光盘驱动器会朝着大容量、高速度、可擦写、低价格的方向迅速发展，前景十分乐观。

总之，无论采用哪种数据存贮介质，还是几种介质混合使用，GIS 用户应该综合考虑其

数据存贮容量和存贮费用的问题。选择作为永久存贮的数据应是那些经过挑选又必不可少的数据。为使这些数据作长期保存并为多用户使用，应有良好的具有安全保护的环境要求：温度控制在 10～32℃，湿度控制在 40%～60%，磁场强度＜50 奥斯特。

二、数据库管理系统

GIS 的海量地理数据需要采用数据库管理系统进行管理。这好比是对图书馆的图书进行编目、分类存放、设立索引以便于管理人员或读者快速查找所需要的图书。图 4-1 所示为一个 GIS 数据库管理系统的基本结构图。

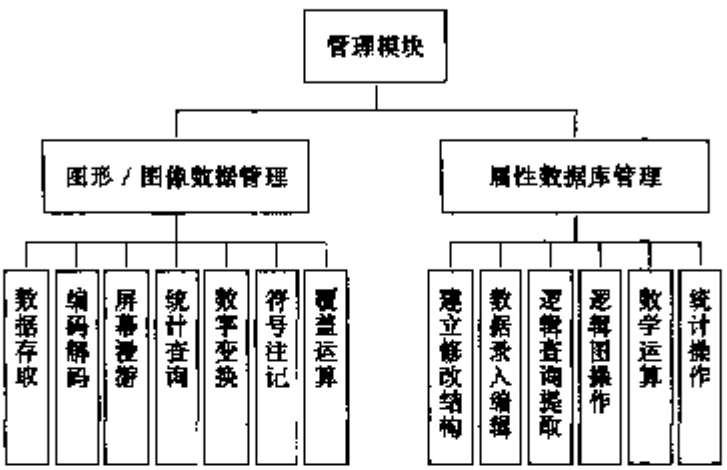


图 4-1　某 GIS 数据库管理系统基本结构图

数据库管理系统的主要功能应包括：数据库定义、数据库管理、数据库维护和数据库通讯等几个方面。

1. 数据库定义

数据库定义是通过数据库提供的数据描述语言实现的。描述语言用来定义数据库的逻辑结构，数据库的结构框架，定义数据级，建立记录类型，定义记录间的联系，指定安全性控制要求，指定数据完整性控制设备等。而这其中又以定义数据库的逻辑结构最为重要。因为它确定了整个数据库设计的基础。

数据库的逻辑结构规划出了数据库的整个框架，并确定了该数据库能完成的任务和形成数据库的总体数据模型。一般而言，数据库的逻辑结构应满足下列要求：

(1) 在综合考虑处理速度、数据吞吐量、操作响应时间、可维护性和存贮要求等因素的基础上，使数据资源共享，降低数据输入量以及数据维护的事务处理方面都达到最大效率。

(2) 保证数据质量，使数据内容与格式的一致性，减少数据冗余，减少或限制数据误差及其不确定因素的影响。

(3) 为最大限度地发挥数据库系统的性能，应具有灵活简便的数据检索、分析功能，能支持各种用户视图并有利于用户进行应用开发。

(4) 数据模型应独立于计算机的硬件和软件并且是面向应用的，易于为用户理解的。维护数据的独立性也是共享数据资源的前提条件。

2. 数据库管理

这一功能应包括对整个数据库的运行控制、数据存贮、更新管理、数据完整性及有效性控制，以及数据共享时的并发控制等。数据库管理系统接受用户发出的操作指令是通过作为用户与数据库之间的接口语言——数据操作语言来完成的。通常它应具有以下功能：

(1) 从数据库检索出满足条件的数据（数据项、记录等）；

(2) 向数据库中插入数据；

(3) 修改数据库中的数据；

(4) 控制操作。

除了执行数据操作语言提供的基本操作外，数据库管理系统还应具有修改索引的功能、存贮块的功能以及缓冲区的管理功能，以致进出数据库数据能够快速有效地装入或调出存贮区。

3. 数据库维护

对数据库的维护功能主要是指系统应具备对数据库重新定义、数据重新组织、性能监督和分析以及对数据库整理和发生故障时恢复运行的能力。

当将原始数据按数据库要求的格式和结构输入数据库时，应同时进行数据的检验工作以避免粗差或无效数据入库，确保数据的正确性。为保证数据库安全并改善数据库效率，数据库维护工作应包括以下几方面的内容：

(1) 改善系统的使用性能。及时掌握数据库性能变化情况，当系统性能下降到一定程度时，进行必要的干预，如对数据进行整理或重新组织，消除降低性能的因素；

(2) 数据库受损后的复原。数据库的安全是极为重要的，对数据库的维护，一方面要采取有效措施防止各种损害数据库的活动；另一方面，必须具备系统受损后的复原手段；

(3) 用户应用管理。数据库是许多用户共享的，为了避免由于局部的使用错误引起整个数据库的彻底破坏，必须对用户实行统一管理，分配数据库子模式的使用权限，并防止应用程序非法使用数据库；

(4) 拓宽数据库用户的需求。根据用户的要求，修改数据模式，重新书写和编译，重新描述数据库的有关内容，并根据新模式组织数据。

4. 数据库通讯

通讯功能是指数据库系统具备与操作系统的接口处理，与各种语言的接口以及与远程操作的接口处理等。如前所述，由于数据库数据的共享性，除本系统的专业用户能够方便地应用数据库中的数据外，还应该能与其它数据库系统进行通讯，发送或接收到其它数据库系统的数据。因此，系统需要有通讯软件和数据格式或转换程序作为这些接口处理。

§4-4 空间查询与空间分析功能

GIS 用户可能提出的并要求系统解答的实际问题可能是各种各样的，比如说以下几例就是经常出现的空间查询问题：

① 物体 A 在何处？

② 物体 A 处于物体 B 的什么地方?

③ 具有属性 A 的地块的面积是多少?

④ B 地物有多大（包含的面积、周长)?

⑤ 各种空间数据的交、并运算结果是什么?

⑥ 各种空间数据叠加运算后结果是什么?

⑦ 具有一定属性组合的物体与什么物体相邻?

⑧ 按某属性组合如何将物体重新分类?

系统要回答这些问题，就需要采用适当的数据分析方法，需要用数据检索的组合方法和选择适当的变换方法。就一般空间查询而言，用户可以就某个地物或区域本身的直接信息进行双向查询，即根据图形查询相应的属性信息，如问题④；按照属性特点，查找相对应的地理目标，如问题③。除此之外，经过适当的选择变换方法，还可以从 GIS 目标之间的空间关系中获取派生的信息和新的知识，来回答有关空间关系的查询和空间分析，如问题⑤、⑥、⑦和⑧等。

GIS 之所以在处理空间信息的性能上强于其它的信息系统，就在于它具有很强的空间查询和空间分析能力。而这种能力主要是由其分析、变换功能所决定的，它们可归纳为对空间数据的拓扑和空间状况的运算、属性数据运算和空间数据与属性数据的联合运算等，如图 4-2 所示。

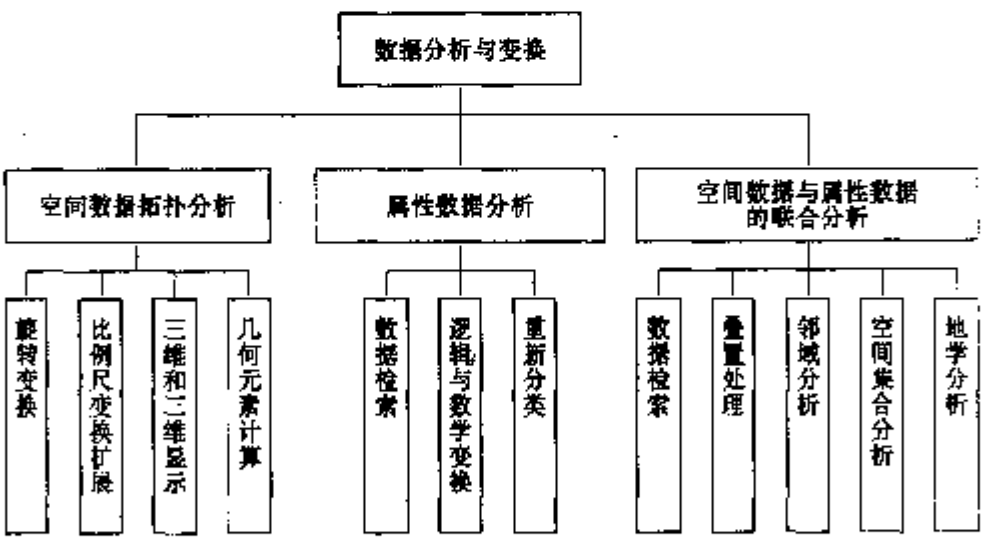


图 4-2 GIS 空间查询和空间分析

GIS 中采用图 4-2 所示的方法是通过一些基本的空间分析操作，如叠置分析、缓冲区分析、拓扑空间查询、空间集合分析、地学分析等得以实现的。下面对其进行讨论。

一、叠置分析

一般情况下，为便于管理和应用开发地理信息（空间信息和属性信息)，在建库时是按分层进行处理的。也就是说，根据数据的性质分类，性质相同的或相近的归并到一起，形成一个数据层。例如，对于一个地形图数据库来说，可以将所有建筑物作为一个数据层，所

有道路作为一个数据层，水系作为另一个数据层等。每个数据层一般都由一种或几种公共属性来定义，比如建筑物数据层的类型就是多边形实体；道路或水系决定的数据层类型为弧段或网络结构；由地下管线井，如雨水井、电信井等构成的数据层为点实体类型等。

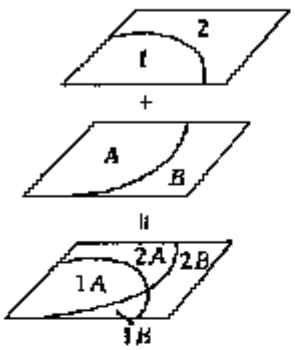
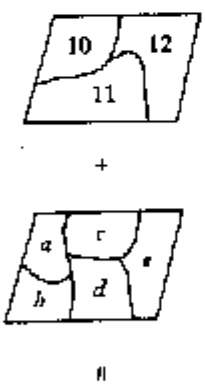


图 4-3　多边形叠置的例子

为确定空间实体之间的空间关系，可以将不同数据层的特征进行叠加，从而产生具有新特征的数据层。也就是说，将同比例尺、同区域的两个或多个数据层叠置在一起，根据它们的边界交点可以建立具有多重属性的图形，即多边形叠加，如图 4-3 所示。或者根据图形范围的属性、特征进行多个属性数据的统计分析，即统计叠加，如图 4-4 所示。在直观概念上，叠加操作就是将两个或两个以上的地图重叠放在一起，产生新的多边形和新多边形范围内的属性。因此，叠加后输出图（或输出数据层）的属性（记作 Y）就是各个参与叠加的地图（数据层）上相应属性（记作 X_i）的函数。可用下述关系式表达：

$$Y=f\ (X_1,\ X_2\cdots,\ X_n) \tag{4-1}$$

式中，f 函数的形式取决于各输入图层上属性与用户需要之间的关系，n 为参与叠加的图层个数。



区域	类型	面　积				
		a	b	c	d	e
10	5					
11	3					
12	4					

图 4-4　属性数据统计叠加的例子

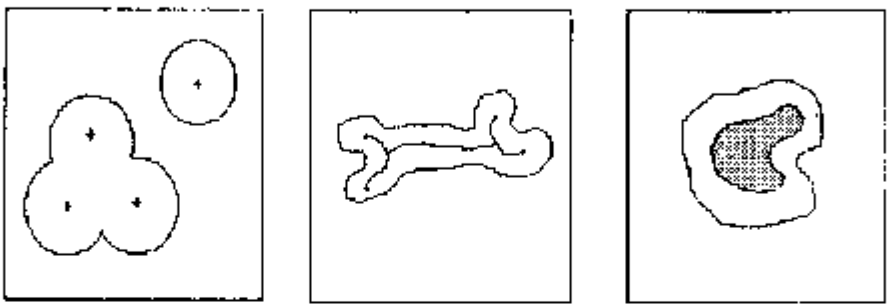
在图 4-3 所示中，叠加操作的输出图上产生了许多新的多边形，每个新多边形也被赋予了新的多重属性值。所以，通过这种区域多重属性的叠加操作，可以寻找和确定同时具有几种地理属性的分布区域。例如，为显示出某城市工业区中具有不稳定土壤结构的所有地区，就可以将土壤结构分布图与土地利用分区图进行多边形叠加操作，产生其合成图。

在图 4-4 所示的统计叠加操作中，所得到的是新的属性数据，其目的是统计计算某种要素在另一种要素中的分布特征。得到的新属性值可能与原叠加数据层上的属性意义不同，可以是分类值、名称等，也可以是基本元素分析或集合分析的结果。例如，进行土地评价所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、土壤结构以及地形地貌等各个数据层的信息，如果直接对这些数据层上的属性值进行数学运算，得到的结果可能是毫无意义的，必须将其变成另一基本元素（如用数值量化的土地适用性）后才能进行这种多因素分析的数学运算，其运算结果可以是统计报表。

二、缓冲区分析

在GIS的空间操作中，涉及到确定不同地理特征的空间接近度或邻近性的操作就是建立缓冲区。例如在一个城市中，要对某个地区做一些改变，就需要通知该地区及其周边地区一定距离（如500 m）范围内的所有单位或居民；在林业方面，要求距河流两岸一定范围内规定出禁止砍伐树木的地带，以防止水土流失；在建设某项目进行选址时，可利用建立缓冲区来查找沿某公路两侧10 km以内尚未被利用的土地分布情况，或在街区改造中，沿某条街两侧200 m以内三层楼以下的建筑物分布情况等。

因此，缓冲区分析就是研究根据数据库的点、线、面实体自动建立其周围一定宽度范围的缓冲区多边形。如图4-5所示为对点、线、面实体建立的缓冲区示意图。



点的缓冲区　　线的缓冲区　　面的缓冲区

图4-5　缓冲区分析

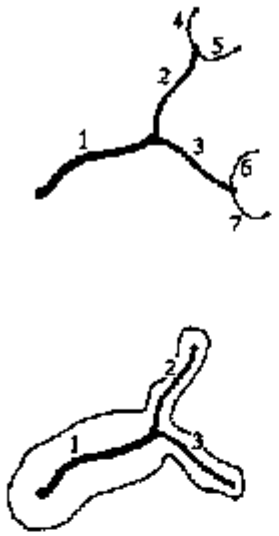
应该说明，以应用缓冲区操作生成的缓冲区多边形将构成新的数据层，该数据层的数据并非是由数据输入时生成的。另外，根据地理实体的性质和属性，对其规定不同的缓冲区距离，通常也是十分重要的。例如，沿河流两岸绘出的禁止砍伐树木带的宽度应根据河流的类型以及河流两岸土质而定；在规划研究中，距交通线、居民点、中心商业区等线状或点状地理实体的距离是进行土地评价和空间布局规划的重要指标，结点的影响范围和相互引力等也需要通过扩散距离来决定，所以在进行缓冲区分析时应允许设置可变距离值的缓冲区。当对某地理实体的可变距离项设置为零时，将不对该实体建立缓冲区。图4-6所示为一有选择地对某一水系建立缓冲区的例子。

三、拓扑空间查询

根据拓扑学原理，一幅图的诸元素可大致分为点、线、面三种基本形式，空间实体的拓扑特征就体现为这三种基本元素的拓扑关系。这种拓扑关系可以分为两大类，一种是所谓关联关系，即描述点、线、面基本元素之间的关系，用以表达几何元素间的拓扑数据结构；另一种是描述空间实体之间的空间位置关系和相关关系。

我们知道，在拓扑结构中，一幅图是由若干多边形（面元素）或线段或点组成的，其中，线段是由若干点组成，多边形是由若干线段组成。所以对拓扑空间的查询，就是对点、线、面三种基本元素相互之间的关系进行分析处理后提取的拓扑特征。

通常，叠置分析是进行这种拓扑空间查询的有力工具，而所要进行分析处理的拓扑关



河流编码	河流类型	缓冲区距离/m
1	3	1000
2	2	600
3	2	600
4	1	0
5	1	0
6	1	0
7	1	0

图 4-6　有选择地建立缓冲区

系一般包括以下几个方面：

1. 点、线、面之间的关系。其中又可分为：

(1) 点与点之间的关系。

查询与某特定点有一定距离的其它点的信息。例如，查询距火车站 3 km 以内的所有旅馆；

(2) 线与线之间的关系。

例如查询与某国道关联的等级公路情况；

(3) 面与面之间的关系。

查询某地区与其周边地区的地理分布。

2. 点与线的关系。

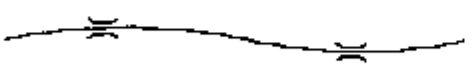


图 4-7　点与线的拓扑关系示例

点与线进行叠置操作的结果是使点与线都具有了新的属性特征。例如，检索某河流上的桥梁，如图 4-7 所示。

3. 点与面的关系。

将一个数据层上的点特征叠置到另一数据层的多边形上，可确定每个点落入哪个多边形。例如，查询某城市内所有邮局的分布，如图 4-8 所示。

4. 线与面的关系。

将一个数据层中的多边形特征叠置到另一个层的线段上，可以确定多边形内所隐含的弧段。例如，对存贮道路的数据层与存贮县界多边形的数据层进行叠置运算，可以得到每个县内所包括的不同类型道路的长度。如图 4-9 所示。

四、空间集合分析

GIS 能够为人们做出的最主要贡献之一就是它具有丰富的空间分析功能。这种功能是

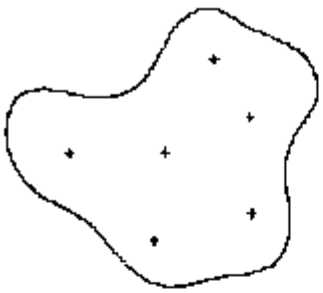


图 4-8 点与面的拓扑关系示例

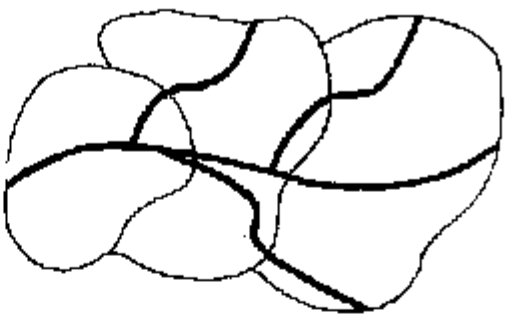


图 4-9 线与面的拓扑关系示例

基于空间实体的位置和形态特征的空间数据分析技术，其目的在于提取和传输空间信息。

空间集合分析是以叠置分析运算和布尔逻辑运算为基础，按用户给定的空间数据组合条件来检索，查询其它的属性项目或图形数据。实际上就是在叠置分析的基础上进行的一个逻辑选择过程。通常是按照逻辑子集给定的条件进行逻辑交运算（AND）、逻辑并运算（OR）和逻辑差运算（NOT）。

实际工作中，GIS 的用户经常需要从多个数据层或多幅地图中提取数据。例如，用户可能希望知道某地区种植小麦的土地是什么样的土壤类型以及与水利资源的分布关系等，而土地利用类型、土壤类型以及河流水系分布数据分别存贮在三幅不同的地图上（或数据层上）。当前，GIS 就是采用将数据层叠置后，根据各数据层指定的条件，运用逻辑运算方法进行逻辑选择提取出用户需要的结果。

1. 逻辑交运算

下面再用三个简单例子来说明逻辑选择的具体用法。假设用户需要查询在粘性土壤上种植小麦的土地 C，可以从土地利用图中得到小麦地的子集 A，再从土壤类型图上得到粘性土壤子集 B，然后对两个子集做逻辑交运算，得到其交集 $C=A\cap B$，如图 4-10 所示。

2. 逻辑并运算

如果用户需要知道的是全部粘性土壤的土地和全部小麦地的情况，就可以采用逻辑并运算。在获取了粘性土壤子集 A 和小麦地子集 B 后，对 A 和 B 两个子集做逻辑并运算，得

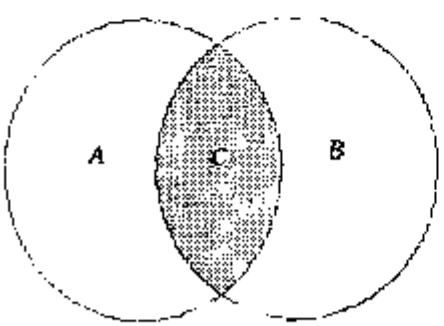


图 4-10 逻辑交运算（A AND B）

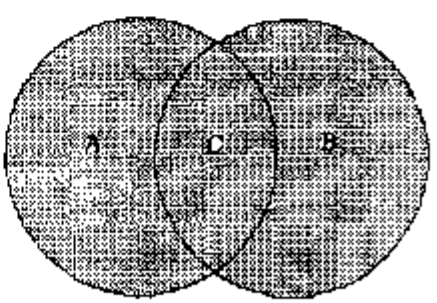


图 4-11 逻辑并运算（A OR B）

到其并集：$C=A\cup B$，如图 4-11 所示：

3. 逻辑差运算

假如用户要求查询的是粘性土壤上没有种植小麦的土地，则可以运用逻辑差运算作用在粘性土壤子集 A 和小麦地子集 B 之上，求出其补集 $C=A-B$，如图 4-12 所示：

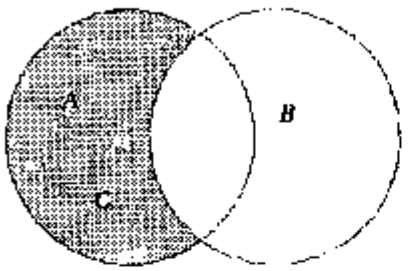


图 4-12 逻辑差运算（A NOT B）

由以上例子可以看出，这些逻辑运算都是按照特定组合条件去检索提取出相应的数据。需要说明的是空间集合分析既适用于基于矢量结构的数据，也适用于基于栅格结构的数据。但是，由于矢量结构数据在进行叠置操作时要比栅格结构数据复杂且速度慢，此外，对矢量数据层叠置还易产生细碎多边形误差和使边界搜索困难等原因，一般用户多愿意用栅格数据形式进行空间集合分析。

五、地学分析

地学分析是用来描述地理系统中各地学要素之间的相互关系和客观规律信息的方法。可以从数字高程模型分析、地形分析和地学专题分析三个方面来论述。

1. 数字高程模型（DEM）

数字高程模型（DEM）是对地理空间起伏连续变化的数字表示形式，用来描述地理空间中的第三维坐标——高程。尽管 DEM 是为了模拟地面起伏而发展起来的，但也可以用其模拟其他二维表面上连续变化的特征，比如说还可表示地面景观的属性等。此时的 DEM，也可称为是数字地形模型（DTM），

DEM 最普通的形式是高程矩阵，也就是说，将地表单元上高程的集合用矩阵来表示。数字高程模型的建立（高程矩阵的生成）除了可以用解析立体测图仪从立体航测像片对上直接获取外，主要可以由规则格网对离散数据点内插而得到。例如，采用人工格网输入时，将地形图蒙上格网，逐格输入格网中心点和角点的高程值，从而构成高程矩阵。由于在计算机中处理矩阵比较方便，所以以栅格数据为基础的 GIS 中，采用高程矩阵较普遍。如美国和英国都采用该法从全国 1：25 万地形图上产生了全国的高程矩阵。此外，高程矩阵还有利于计算等高线、坡度、坡向、自动描绘流域轮廓等。但是高程矩阵法得到的高程是原始采样点的派生值，在内插过程中将损失高程精度，并且采样工作量较大；同时还存在数据冗余与精度不足的矛盾。如地形简单地区数据冗余，而固定的格网尺寸又无法正确反映起伏复杂地区的情况。尽管目前有采用可变格网分辨率的方法采样，但是仍未彻底解决数据冗余问题。

另一种建立数字高程模型的方法是所谓不规则三角网法（TIN）。不规则三角网数字高程模型是由连续的相互联接的三角平面组成，三角形的形状和大小取决于不规则分布的高程数据点的密度和位置。其构成方法是对有限个离散点高程，每三个最邻近点联接成三角形，根据每个三角形的平面方程可计算各格网点高程，生成 DEM。TIN 与高程矩阵不同之处在于它能随地形起伏变化的复杂性而改变采样点的密度和决定采样点的位置。因而能够

克服地形起伏变化不大的地区产生数据冗余问题，同时还能按地形特征点如山脊线、山谷线、地形变换线和其它能按精度要求进行数字化的重要地形特征获得 DEM 数据。TIN 结构也可以用来产生坡度图、等高线图、晕渲图、三维立体块图等。但是，采用 TIN 法要求有足够的离散数据点的情况下，才能保证模型精度。此外，TIN 的数据存贮量一般较大，在进行大规模大区域的规范化管理以及与 GIS 的空间数据或遥感影像数据进行联合分析应用中尚存在一定困难。

当前一般的 GIS 都提供两种数字高程模型的软件工具，用户可以根据需要自行选择。

2. 地形分析

在 GIS 中应用数字高程模型为工具，可以获得多种派生产品，从而作为地形分析的有力手段。

(1) 等高线图

等高线图适用于对连续变化的地面进行模拟表示。在地形图上绘制等高线表示地面起伏是人们传统上观测地形的主要手段。人们可以在等高线图上定量地量取，获知地形起伏程度以及有关区域内各部分的高程信息等。在 GIS 中等高线数据既可以通过数字化原图得到，也可以利用 DEM 的高程矩阵或 TIN 数据得到。

从高程矩阵中获得等高线图的方法是把高程矩阵中各单元的高程分成适当的高程类别，然后用不同的颜色或灰度输出每一类别，两高程类别之间的分界线即可视为等高线。

从不规则三角网 TIN 数据中产生等高线的方法是用水平面与 TIN 相交割而产生等高线。TIN 中的山脊线、山谷线等数据可用来引导等高线的起始点。

(2) 三维立体图、透视图

等高线图虽然可以定量地描述地面起伏变化，但是在进行数据分析和模拟计算方面尚有不足，此外，在视觉感应上也有欠缺。从现在以及今后，GIS 用户将越来越需要能从直观上观察地形的概貌和进行数据分析与模拟计算。因此，GIS 应具备绘制三维立体图和透视图的功能。

三维立体图在显示多种土地景观信息中非常有用，利用 DEM 以数据形式表示地表数据变化，通过计算三维立体图的程序用线划描绘或是用栅格阴影显示来表示出立体图形。三维立体图计算要求用户指定一个观测点和在高程方向上放大的比例尺，并设定透视因子计算中包括了解决隐藏线或隐藏物的方法，从而可以使结果产生立体透视图的模型，使图形更加逼真。有的系统采用彩色三维格网显示，则视觉效果更佳。

(3) 坡度和坡向

由于 GIS 采用了数字高程模型，以数字形式用高程矩阵或 TIN 结构描述地面起伏状况，使得 GIS 用户可以对地貌的描述和比较进行定量分析。

坡度的定义是水平面与局部地面夹角的正切值。实际计算坡度时将其分成两个部分：斜度——高度变化的最大比率，即我们常说的坡度概念；坡向——变化比率最大值的方向。斜度和坡向两个因素基本上能满足地形分析的要求。通常斜度用百分比来量度，坡向按从北方向起算的角度量测。斜度和坡向的计算通常是在 DEM 的数据矩阵中采用开窗连续移动搜

索以完成整幅图的计算工作。输出值不仅可以是数字形式，而且还可以派生出坡度分布图和坡向图。

(4) 自动形成地形轮廓线

使用数字高程模型进行数量分析时，需要具备提取和描绘线、面特征的能力。比如，在DEM上叠置的卫星影像上需勾绘出江水范围线，可以使遥感数据与特殊地理景观联系在一起。目前，人工进行水系流域分析和水网分析需花费大量时间并且还难以避免数据误差或错误的出现。因此，在GIS中建立数字高程模型DEM，借助于自动描绘地形特征线程序是很有意义的。

系统中设计有专门的计算方法可以自动地探测山脊线和山谷线，并将其提取出来。此外，还可以对汇水范围或流域范围进行确定。这种计算不仅能确定范围边界线，而且还能将整个范围从整个数据库中分离出来，用户可以对其重新编码形成某一流域的分布图。这对水系的流域分析是十分必要的。

3. 地学专题分析

GIS除有以上所述功能外，还提供了许多专业性较强的应用计算程序。比如，利用DEM数据，可以很容易地建立断面图，从而为工程设计和工程量算提供依据；可以对工程土石方量和填挖范围进行计算；还能够根据DEM格网或TIN三角形的每个基本单元面积，累积求和算出整个区域地表面积等。从而有效地提高了工作效率和经济效益。

此外，GIS中还具有一些专用分析模块，例如网络分析模块能够用来进行最佳路径分析，以及追踪某一污染源流经的排水管道等；土地适应性分析是多因素分析的一种，可用来评价和分析各种开发活动，包括农业应用、城市建设、农作物布局、道路选线等用地，优选出最佳方案，为土地规划提供参考意见；发展预测分析可以根据GIS中存贮的丰富信息，运用科学的分析方法，预测某一事物如人口、资源、环境、粮食产量等，对今后的可能发展趋势给出评价和估计，以调节控制计划或行动。另外，利用地理信息系统还可以进行最佳位址的选择；新修公路的最佳路线的选择；辅助决策系统和地学模拟分析等等。用户可以根据需要自行选择应用分析模块。

六、应用实例

1. 道路拓宽改建过程中的拆迁指标计算

本例说明在计算一条道路拓宽改建过程中的拆迁指标时，如何应用GIS的建立缓冲区、拓扑叠加和特征提取等功能解决具体问题。

步骤一：明确分析的目的和标准

本例的目的是计算由于道路拓宽而需拆迁的建筑物面积和房产价值。道路拓宽改建的标准是：

——道路宽度从原来的20 m拓宽到60 m；

——拓宽后的道路应尽量保持直线；

——部分位于拆迁区内的10层以上的建筑物不拆除。

步骤二：准备进行分析的数据

本例需要二类信息：一类是现状道路图；另一类是分析区域内的建筑物分布图以及相

关信息。

步骤三：进行空间操作

首先选择拟拓宽的道路，根据拓宽半径，建立道路的缓冲区；然后，将此缓冲区与建筑物数据层进行拓扑叠加，产生一幅新图，此图包括所有部分或全部位于拓宽区内的建筑物信息。

步骤四：进行统计分析

对部分或全部位于拆迁区（拓宽区）内的建筑物进行选择，凡是落入拆迁区内且楼层高为10层以上的建筑物，将其从选择组中去掉，并且对道路的拓宽边界进行局部调整。然后对所有需拆迁的建筑物进行拆迁指标的计算。

步骤五：将分析结果以地图和表格的形式打印输出

2. 辅助建设项目选址

本例说明如何利用GIS空间操作和特征提取功能为某建设项目选择最佳建设位置的步骤。

步骤一：建立分析的目的和标准

分析目的是确定一些具体地块，作为一个有轻度污染的工厂的可能建设位置； 厂选址的标准为：

——地块建设用地面积不小于1万m^2；

——地块的地价不超过1万元/m^2；

——地块周围不能有幼儿园、学校等公共设施，以避免受到工厂生产的影响。

步骤二：从数据库中提取用于选址的数据

为达到选址的目的，需要准备两种数据：一种为包括全市所有地块信息的数据层；另一种为全市公共设施（包括幼儿园、学校等）分布图。

步骤三：进行特征提取和空间拓扑叠加

从地块图中选择所有满足上述标准中前两项的地块，并与公共设施数据层进行拓扑叠加。

步骤四：进行邻域分析

对叠加结果进行邻域分析和特征提取，选择出满足要求的地块。

步骤五：将选择的地块及相关信息以地图和表格形式打印输出。

§4-5 数据输出与表达功能

GIS的数据输出与表达是指借助一定的设备和介质，将GIS分析或查询检索结果表示为某种用户需要的可以理解的形式的过程；或者是将上述结果传送到其它计算机系统的过程。这种输出形式转化为人们能理解的共同形式就是地图、表格、图形和图像等；转化为计算机兼容的形式就是能读入其它计算机系统的磁盘、磁带或光盘记录形式等，以及某些通讯网络、电话网、无线电连接设施等电子传输形式。由此可知，输出就是将GIS的信息形式表达成适合用户需要的过程。由于在GIS中所有的图形图像信息都以数字形式存储，

而且它们都以有效的数据结构与各种专题信息有机地联系在一起，因此，GIS 在最终成果的表示与输出方面显示出极大的优势。这表现在如下一些方面：

——地图表现内容的灵活性：不仅可以为用户输出全要素图，也可根据用户需要，分层输出各种专题地图，如行政区划图、道路交通图、等高线图等；还可以通过空间分析得到一些特殊的地学分析用图，如坡度图、坡向图、剖面图等；另外，还可根据用户应用需要，随时叠置专题信息，形成专题地图。

——地图表现形式的可变性：通过图形编辑功能，设计图形符号、填充花纹和复杂线型，可方便地制作图例、拼图表、指北针、比例尺等各种图饰；可自由选择各种字体进行标注，并实现图形与文字排版；有的还具有编图功能，能处理图形的复杂压盖、避让等特殊图形问题。由于所有图形都已数字化，可以反复利用，随时修改以满足不同用户对地图形式变化的要求。

——产品类型的多样性：表现专题信息的空间分布特征和相互关系，可以输出专题地图或专题影像地图；表现统计信息、决策信息等可输出直方图、饼状图等统计图；若要获得属性数据库的查询、统计和运算结果，可以输出正规的统计表格；此外，还能产生一系列的数字化产品，如符合各种图形交换标准的地图数据、多媒体的电子地图等。

——图形设备的适应性：GIS 的图形输出一般与设备独立，它能输出各种标准格式的图形图像数据文件，从而可以在打印机、笔式绘图机、静电绘图机、喷墨绘图仪和激光照排机上输出。

数据的输出形式可以分为三种：硬拷贝、软拷贝和电子输出。硬拷贝输出是永久显示，信息被打印在纸上、摄影胶片或类似的材料上。地图和表格通常用这种格式输出。软拷贝是一种能在计算机屏幕上观察的输出格式，这种输出可以是黑白或彩色方式输出的文本或图形，它用于操作员与计算机之间的交互操作和正式硬拷贝之前对数据的浏览。软拷贝一般不用作最后输出，因为它的尺寸大小和当屏幕影像被照相或电子获取时有质量损失。电子输出是由计算机兼容文件组成的输出方法，通常用于计算机系统之间的数据传递、进一步分析或带往远地产生硬拷贝输出之用。

对于输出地图、表格和图形、图像等能被人们理解的数据的设备可以分为两大类：一种是在电子屏幕上显示出 GIS 的分析结果或输出内容，如图形终端等计算机屏幕；另一种是在纸张、聚酯薄膜或其它材料上产生永久性图形或文本数据的所有装置，例如，打印机、绘图机等。此外，对于计算机兼容输出形式的设备通常有磁带机、磁盘驱动器等以及必要的计算机网络通讯设备。图 4-13 所示，为 GIS 数据输出的一个简单示意图。

通常输出显示装置还分为矢量显示设备和栅格显示设备两类。这两类设备常与它们连接的地理数据库的数据结构（矢量结构和栅格结构）相一致。当然，在实际工作中，并非一定要使矢量输出设备连接到矢量数据库，栅格设备连接栅格数据库，很多实践已经证明数据结构与输出设备的类型不一定相同。但矢量数据在矢量绘图仪上输出，绘出的地图质量较高，当然它也有输出速度较慢、实现面状渲染图较难的不足。

下面分别对各种输出设备作简要说明。

一、屏幕显示器

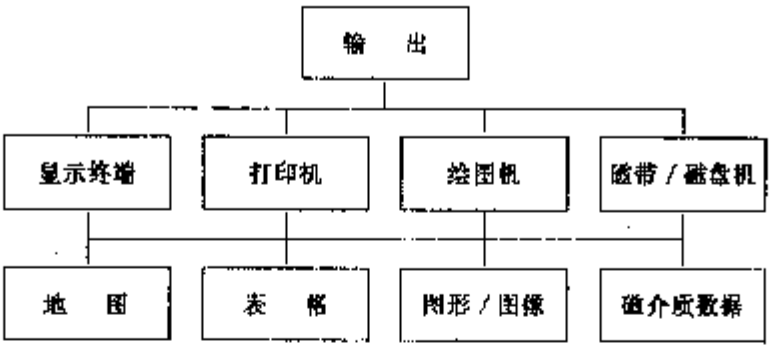


图 4-13　数据输出

采用电子显示屏幕作为人机交互的数据输出设备具有代价低、速度快、得到的图形、图像色彩鲜艳，且可以动态刷新等特点；但也有显示幅面小、显示精度不高、显示比例失真等不足之处。此外，屏幕显示不是永久性输出，关机后就无法保留其影像（对屏幕图像进行硬拷贝需用特殊装置进行），所以这种输出形式一般只是作为暂时性输出，供人机交互操作时参考之用。但是，在GIS运行中的人机交互操作量是很大的，屏幕显示从数据输入、编辑到空间操作等各个阶段都要用到，并且在很多情况下，用户并不要求图形输出的硬拷贝，而只是在屏幕上显示地图。所以屏幕显示输出是使用最频繁的一种方式。

在屏幕显示中，为保持图像在屏幕上显示，是通过让电子束重复扫描屏幕来实现的。为了不致使屏幕上的图像有闪烁和模糊现象，屏幕的刷新频率至少应有 60 次/秒（屏幕刷新频率又称之为扫描率）。通常屏幕图像的产生是使用栅格扫描的，处理的方法是将屏幕划分为离散栅格和像元阵列，每个像元的亮度是单独控制的；大幅图像在屏幕上分幅显示，栅格越小，所表达的图像的细节越详细，分辨率也就越高。一个典型的高分辨率彩色显示器能够提供约 0.3 mm 大小的屏幕像元尺寸。一般微机显示器屏幕大小主要有 14、15、17 和 19 英寸等规格。为了让屏幕能够显示文字或图形，还必须配备有相应的显示卡，目前基于微机的数字测图系统或GIS系统常选用VGA显示卡或Super VGA显示卡。通常拥有高分辨率显示卡的屏幕能够产生较佳的画面。分辨率的计算单位是以像元表示。一个像元可代表一个点。例如，在显示文本数据时，VGA 卡的分辨率可达 720×400；在显示图形时的分辨率为 640×480。这里分辨率 640×480 表示该屏幕每行拥有 640 个像元（点），每列拥有 480 个像元。一般的 Super VGA 卡的分辨率为 800×600。有些计算机厂商为突出自己的产品，在分辨率能力上都尽量地使其得到提高，如有的 Super VGA 拥有 1024×768 分辨率，或 1280×1024 分辨率。通常，高分辨率的屏幕显示器使用大约 1 百万个像元点表达一幅屏幕图像。

对于一个实用的GIS运行系统应该配备有高分辨率的图形终端。同时为提高显示效果，在经济能力允许的前提下，应尽量选择大屏幕显示器。通常对工作站硬件配置的图形终端而言，最好选择 19 英寸以上，分辨率 1280×1024 的彩色显示器为佳。除此之外，由于图形数据操作的计算量大，为加速图形计算，减轻主机的负荷，一般还应该配备图形卡或称图形芯片，比如 Sun 工作站配置的 GX 卡或 GS 卡。

屏幕画面是依靠电子束连续从屏幕上部到底部逐行连续扫描像元阵列产生的。在扫描过程中，每束光束的强度是随要显示的每个像元所规定的亮度而变化的。要重复显示高分辨率的图像，就需要使扫描过程重复连续地以 60 次/秒的较高数据传输率进行。当前为减少传输率和降低硬件成本，常采用所谓“隔行扫描”的方法，即每次扫描只需传输一半数据，1/60 秒的时间仅显示屏幕的奇数行或偶数行。这样做的代价是图像清晰度会有一定的损失，并且图形也可能会产生一定的变形。

二、打印机

打印输出一般是直接由栅格结构方式进行的。与一般微机配套的行式打印机就是一种最简单的栅格显示设备。行式打印机具有打印速度快、售价成本低的特点。但行式打印机的最大缺点是输出图形质量粗糙、精度低，并且输出图形的纵横比例不一，使输出结果失真。因此，目前多采用行式打印机输出文本数据或输出图形草图。

点阵打印机的出现有效地解决了行式打印机输出结果比例失调的问题。这种打印机可以根据用户需要进行设置，如正方形、长方形、点阵密度和图形符号等。用点阵打印机还可以打印输出各种线划图、灰度图和面积填充符号图等。如果设置点阵中的点打印，则分辨率可达到 0.1 mm 以下。若采用三色带点阵打印机还可打印比例正确的精美彩色地图。因此，点阵打印机具有打印速度快、精度较高、成本较低的特点，但由于输出幅面较小，大的输出图需进行拼接，因而限制了其应用范围，目前可在小型 GIS 系统中采用。此外，采用静电复印原理的高分辨率点阵打印机，如在市场上常用的激光打印机，可以产生高质量的输出图形和图像。其分辨率达到 600 dpi 甚至 1200 dpi，具有速度快、精度高和图形美观等特点。

三、绘图仪

1. 笔式绘图仪

笔式绘图仪分为平板式和滚筒式两种。它属于矢量显示设备，也是目前国内最主要的图形输出设备。所有输出信息都是由计算机控制绘图笔（或刻针），在图纸上或模片上绘出。

笔式绘图仪的灵活性和绘图速度在很大程度上取决于绘图软件的程序功能。简单的绘图仪必须具备落笔、抬笔和移动笔这些基本程序命令。因此一台简单的绘图仪也必须包括编制好的绘图程序、字体库和符号集等，只需简单的计算机命令就能调用这些功能。

笔式绘图仪的输出质量在很大程度上取决于笔控制马达的步进量。对制图而言，其步进量不应大于 0.05 mm。对一些高精度的地图绘制还可安装一些专用附加装置，以减少绘图误差。如在绘图笔的位置上装配激光束，可直接在印刷板膜层上绘制或直接绘在胶片上等。

对于实用 GIS 系统，选用的绘图仪幅面应不小于 A0 的幅度，实际分辨率应高于 0.05 mm，当对精度有较高要求时，分辨率可要求达到 0.01 mm～0.02 mm，笔数≥8。

笔式绘图仪是矢量绘图设备，因此在输出矢量结构数据时可得到较高的制图精度。但是这需要一套功能齐备的相应绘图软件的支持，以保证输出数据按用户理解和乐于接受的方式表达出来。因此应对矢量图形的整饰输出问题进行研究，至少需要考虑图面整饰、符号管理和绘图输出三部分内容。

(1) 图面整饰：由图面布置、符号注记、多边形填充、图例生成、文字注记等组成。无论是普通地图还是专题地图，均应包括矢量数据、图名、图框、图例、比例尺等内容。该

项工作的好坏将直接影响输出的效果。

(2) 符号管理：对不同类型的图形符号，如点状、线状和面状符号采用不同方式加以管理和使用。由于图形符号是用来表示地图上需要反映的地物的，因而应以尽可能形象地表现地物主要特征的原则来设计有关符号。如有可能还应建立相应的图形符号数据库及其管理系统。

(3) 绘图输出：包括向用户提供选择绘图机类型、设置仪器状态、确定绘图数量及图幅分割问题等功能。还应考虑有关绘图驱动软件以及有关硬件的选择等问题。

综上所述，笔式绘图仪作为一种矢量输出的设备在输出矢量数据时具有输出精度高、表现精美、输出图幅大等优点。但是，它也有设备成本较高、软件开发较复杂、对用户要求较高的不足之处。

2. 喷墨绘图仪

喷墨绘图仪是栅格显示输出设备的一种。它的分辨率，即喷到纸张上墨滴的大小，受到栅格数据的像元值的控制。每个栅格像元根据其不同的灰度对应于一定数量的静电电子，从而达到控制灰度值深浅的目的。静电电子数与每个栅格像元的位组合模式相匹配。

目前市场上的彩色喷墨绘图仪的分辨率可以达到 10～12 个点/mm，即 300 dpi 的分辨率，彩色绘制时能产生 400 多种颜色，并且幅面可以达到 A0 幅面规格。不足之处在于价格比较昂贵，成本太高。

采用喷墨绘图仪进行图形输出，适合于对栅格数据或经过格式转换后形成的栅格数据进行处理。输出时，每个栅格像元对应一个“墨点”，最后输出一幅比例准确、表现精美的彩色地图。

与矢量绘图输出相类似，进行栅格数据输出时，也应考虑其图形整饰、图形符号管理和绘图输出三个部分的内容。

四、数据表示

如上所述，GIS 用户可以接受的数据输出形式通常为地图、表格和图形、图像等。这就要求经 GIS 的查询检索和分析计算之后，将有关的空间数据和属性数据按满足某标准图式的地图形式、文件表格形式以及按指定格式输出数据内容。根据标准图形软件的要求，对矢量数据结构的基本描绘单元采用线单元、标记单元和文本单元等。线单元用来绘制弧段；标记单元用于控制符号位置；文本单元则用来确定文字说明和注记的位置和方向。栅格数据结构则是采用像元阵列和填充单元为基本单元。在像元阵列单元上不仅可以附加上属性数据，而且还可以作为简单实体加以定义和处理。

GIS 用户通常希望从系统得到的地图产品具有与传统地图产品相一致的制图标准。在输出地图上，也应注明地图图例和分类间隔、图名、定向标记、地图比例尺等信息。此外，对地图的颜色和灰度级别的选择应该以能清晰明确地传送信息为原则。

采用 GIS 进行输出数据能够实现以往用手工方法无法实现的一些数据表达形式。例如使用彩色喷墨绘图仪可以进行三维地貌形态的显示；在三维立体图上或在表示地面起伏的地貌晕渲图上叠加上专题信息等。这些方法能够产生形象生动、效果很强的图形。GIS 功能强大的空间分析能力和 GIS 的数据输出方法相结合，使 GIS 产生神奇的效果。

第五章　GIS 工程设计

地理信息系统工程是应用系统原理和方法，针对特定的实际应用目的和要求，统筹设计、优化、建设、评价和维护实用的 GIS 系统的全部步骤和过程。它的基本原理是系统工程原理，即从系统的观点出发，立足于整体，统筹全局，同时，又将系统分析和系统综合有机地结合起来，采用定量的或定性与定量相结合的方法，提供合理的 GIS 建设思路和模式。

地理信息系统工程设计的主要内容包括确定系统目标，进行系统分析和系统设计以及系统实施和维护、评价等工作。为了加强系统的实用性，降低系统开发和应用的成本，提高系统的生命周期，采用先进的系统设计方法，合理的设计路线是十分重要的。因此，系统设计人员在开发地理信息系统时，应该首先明确用户需要，并据此确定系统要做的工作，形成系统的逻辑模型；然后根据系统的功能，进行子系统划分，使各个子系统分别满足所规定的要求；通过制定统一的工作计划，开发设计有关的子系统，并对各个子系统进行一致性测试，以完成整个系统的功能。

§5-1　设计模式概述

一、结构化设计模式

地理信息系统设计方法基本上起源于比较成熟的信息系统技术，在应用中采用了结构化分析和设计原理。所谓结构化就是有组织有计划和有规律的一种安排。结构化系统分析方法，就是利用一般系统工程分析法和有关结构概念，把它们应用于地理信息系统的设计，采用自上向下，划分模块，逐步求精的一种系统分析方法。这种结构化分析和设计的基本思想包括如下的要点：

——在研制地理信息系统的各个阶段都要贯穿系统的观点。首先从总体出发，考虑全局的问题，在保证总体方案正确、接口问题解决的条件下，按照自上向下，一层层地完成系统的研制，这是结构化思想的核心。

——地理信息系统的开发是一个连续有序、循环往复不断提高的过程，每一个循环就是一个生命周期，要严格划分工作阶段，保证阶段任务的完成。例如，没有调查研究，掌握必要的数据，就不可能很好地进行系统分析；没有设计出合理的逻辑模型，就不可能有很好的物理设计等等。这是系统设计的基本原则。

——用结构化的方法构筑地理信息系统的逻辑和物理模型，包括在系统的逻辑设计中，分析信息流程，绘制数据流程图；根据数据的规范编制数据字典；根据概念结构的设计，确定数据文件的逻辑结构；选择系统执行的结构化语言，以及采用控制结构作为地理信息系统设计工具。这种用结构化方法构筑的地理信息系统，其组成清晰，层次分明，便于分工

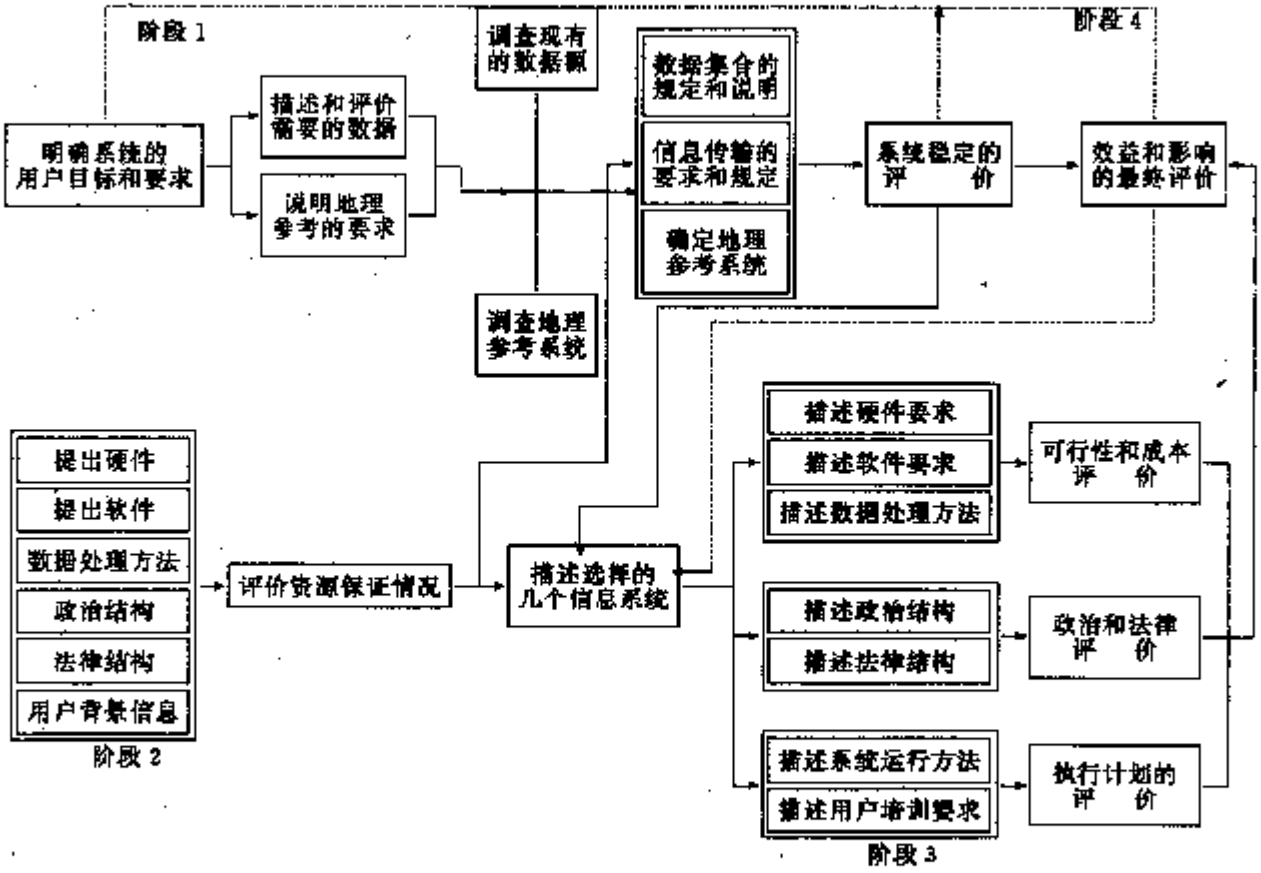


图 5-1　结构化的系统设计模式

协作，而且容易调试和修改，是系统研制较为理想的工具。

——结构化分析和设计的其它一些思想还包括：系统结构上的变化和功能的改变，以及面向用户的观点等，是衡量系统优劣的重要标准之一。

根据上述结构化系统设计思想，可以构造出结构化系统设计模式的如下四个阶段，如图 5-1 所示。

1. 通过访问用户，调查用户的需求和数据源，确定系统的目的、要求和规定，从而分清任务的执行顺序；

2. 描述和评价与系统设计过程有关的资源和限定因素，例如现有的硬件、软件和有关的政治和法律因素等，明确任务的执行条件和分支结构；

3. 说明和评价所拟定的不同系统，重复执行某项任务直到定义的条件满足，这些系统能够满足所规定的要求；

4. 对拟定的系统作最后的评价，从中选择一个运行的系统。

该模式的主要特点是强调对用户的调查和系统功能需求的分析。在系统设计的各个阶段都要写成有关的文件，以便进行评价，以及用户要参与系统的设计，以免系统设计的失误。

结构化设计中最重要也是最流行的方法是自顶向下逐步精化的顺序设计方法。该法将系统描述分为若干层次，最高层次描述系统的总体功能，其他层次则一层比一层更加精细更加具体地描述系统功能，直到分解为程序设计语言的语句。这种方法的特点是结构描述比较清晰，便于掌握系统全貌，也可逐步细化为程序语句，是一种很有效的设计方法。

二、现代原型化设计模式

自结构化系统设计模式应用于地理信息系统设计以来，其设计思想已经有了很大的发展。原来的设计模式是假定系统的大部分组成（除了硬件以外，包括所有的软件和数据库）都需要由系统设计人员来完成，有时甚至包括处理空间数据的某个专门的硬件。现在的情况不同了，不但有许多处理空间数据的重要软件，而且有现成的系统和空间数据库。因此需要对原来的地理信息系统设计模式进行修改。其主要的设计思想，是强调对现有的各个组成部分，包括硬件、软件和数据库，进行深入认真的评价，以研究其满足系统功能的程度，保证所设计系统的实用可靠以及有效地处理数据和使用周期长等要求。该模式采用了管理信息系统和软件工程的一些设计理论，其中包括：

1. 目的与任务。每个系统都要对目的和任务作详细的说明，指出该系统的目的，谁是主要的用户，以及如何使用该系统。关于任务，要说明所要完成的工作，以及总体评价所采用的方法。目的和任务的说明要非常详细，以便用户进行评论和评价，而且这种说明代表着用户和系统设计人员对话的开始，并且在系统设计的过程中还要继续进行这种对话。

2. 概念的定义。介绍系统的各个主要组成部分，分别按照输入、输出、主要过程和数据库来说明系统的基本结构，包括主要模块、系统开发的主要资源、主要的限制条件等。

3. 功能的要求。具体说明该系统要做什么，对每一种功能要求都要说明，包括功能的

技术特征、功能的目的、具体的标准和满足的条件等等。功能的要求一般分为以下三类：

(1) 该系统必须具有的功能；

(2) 如果时间和资源条件允许，希望具有的功能；

(3) 其它有意义的功能。

实际上只有第一类功能才是系统的真正要求，每一种功能要具体规定：

——输入（来源、数据、频率）；

——输出（格式、数据量、用户）；

——功能需要的处理步骤；

——功能成功地实现所需要的条件；

——功能生成的数据。

除了规定要完成的功能外，还要说明该系统期望的性能和特征、质量控制措施以及该系统与其它部分的接口等。显然，在系统开发的过程中，对系统的功能要求可能发生变化，因此要制定专门的计划进行处理。

4. 性能测定。在系统设计过程中，要对各个组成部分分别进行测试，对综合以后的整个系统要进行最后的测试。具体测试的内容包括：硬件设备、软件模块、数据库的质量控制等。测试根据所说明的功能要求和规定的标准进行，测试时一般应考虑以下几个条件：

(1) 系统的正常操作条件；

(2) 重点测试，包括最坏情况和极端操作条件；

(3) 逻辑测试，指检查各种可能的逻辑条件；

(4) 线路测试。

从认识论的角度来看，地理信息系统设计思想也有很大发展。早期的系统设计方法大多采用线性模型，即将系统设计与系统实施处理成无回归的单向发展过程，如图 5-2 所示。

随着地理信息系统技术的发展，这种线性模型的设计的方法已不能适应 GIS 的系统设计。这是因为 GIS 系统的服务对象（即用户）以及用户需求不是固定不变的，而是逐渐发展的。因此，初期拟定的目标和系统数据规范等很难保持一成不变。与此相反，不断地要求修改和完善是必然的。此外，现代技术的发展变化也很快，要使地理信息系统的应用跟上新技术的发展常常也要求改变原有的设计。

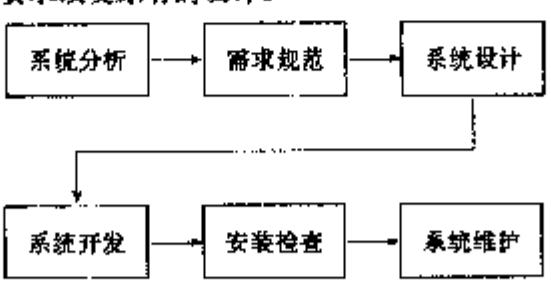


图 5-2 系统设计的线性模型

为适应这种发展和变化的要求，地理信息系统的设计人员对传统设计方法中的原型法 (Prototype Method) 进行了研究和发展，使之更能适合于 GIS 设计工作。这种设计思想

要求在系统建设的早期阶段，生成一个实实在在的系统原型；然后将该原型提供给用户使用，听取用户的批评意见，再根据反馈信息修正系统原型，补充新的数据、数据结构和应用模型，再提交给用户使用。这样，在系统设计过程中，包含有一个再设计阶段，形成了迭代的过程而不像多数传统设计方法那样仅是线性过程，如图 5-3 所示。这种方法的特点是不需一开始即清楚地描述一切，而是在明确任务后，在设计过程中逐步对系统进行定义和改造，直至系统完成。这是一种动态的设计技术。

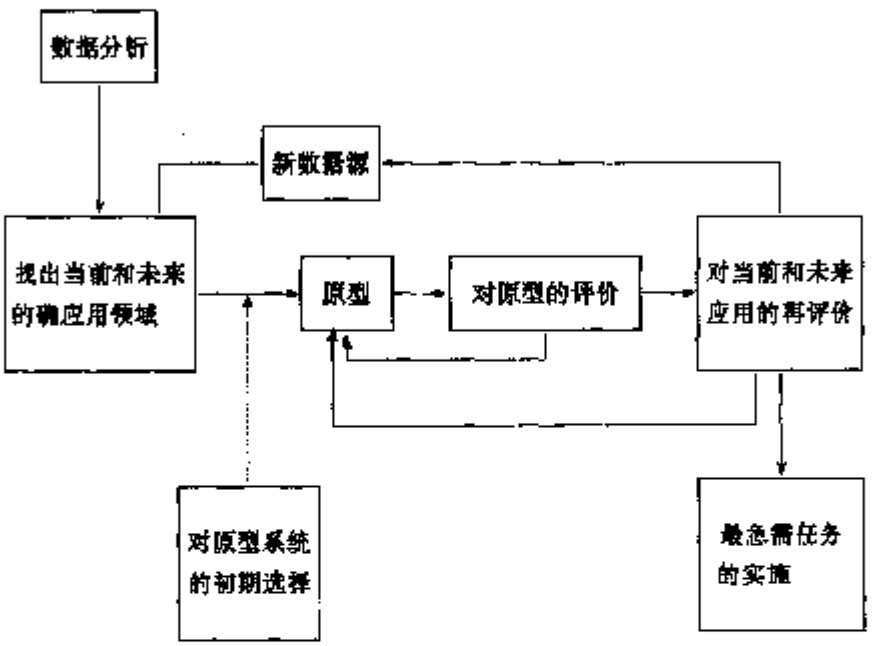


图 5-3　原型设计方法的基本模型

上述设计思想能够比较好地适应地理信息系统的建设特点。它的好处是能创立一个看得见摸得着的样板系统，便于用户试用和提出意见，这样也就更有利于吸引用户介入系统设计工作。同时这种设计思想对于较复杂和具有不确定性的系统目标有较强的适应性，可以使设计与实施达到更为紧密的结合。应该说明，系统原型是一个正式运行的信息系统的早期版本，它能够显示出该系统的基本特征，在重新设计或局部重新设计阶段，还应该顾及以下三点：

(1) 有一个基本的系统可以运行。这个系统应能进一步扩展，并能容纳因存入新的数据而增添的新的应用；

(2) 基本的系统应能有效地试验应用项目，以便使他们能有效地集成起来；

(3) 使负责系统的管理人员能对一些具有紧迫性的应用项目及时作出有效的响应。

三、面向对象 (Object-Oriented) 的设计模式

面向对象的设计方法是近年来发展起来的一种新的设计技术，其基本思想是：将系统所面对的问题，应用封装机制，按其自然属性进行分隔，按人们通常的思维方式进行描述，建立每个对象的领域模型和联系，既模拟信息实体的内在结构又模拟动作机制（如路径选

择和图像解释就是矢量数据与栅格数据两类应用的典型范例），使设计出的软件尽可能直接地表现出问题求解的过程。整个系统只由对象组成，对象之间联系通过消息（messages）进行。由于采用了将数据和操作行为封装在一起的模块化结构，从而使系统很容易重组，而其他系统就必须重写。这对于结构复杂的系统是难以承受的。因此面向对象设计方法的优点就是能保护现有资源，同时也很容易扩充和重组。

所谓面向对象的定义是指无论怎样复杂的事物都可以准确地由一个对象表示。例如，地图上多边形的一个结点或一条弧段可定义为对象；一条河流，或一个省，也可定义为一个对象。下面是面向对象技术的一些有关概念：

（1）对象：是事物的抽象单位，具有特征的内部状态、性质、知识和处理能力，通过消息传递与其他对象相联系，是构成系统的元素或说是封装了数据和操作集的实体。

（2）消息（message）：是请求对象执行某一操作或回答某些信息的要求，用以统一数据层和操作控制，将对象联系起来。

（3）分类（classification）：是关于同类对象的集合。具有相同属性和操作的对象组合在一起形成类。属于同一类的所有对象共享相同的属性项和操作方法，但每个对象可能有不同的属性值。以一个城市的GIS为例，它包含了建筑物、街道、公园、给排水管道、电力设施等类型，而中山路51号是建筑物类中的一个实体，即对象；建筑物类中可能还有诸如地址、房主、用途、建筑日期等其他属性，并可能需要显示对象、更新属性数据等操作。

（4）概括（generalization）：在定义类型时，将几种类型中某些具有公共特征的属性和操作抽象出来，形成一种更一般的所谓超类，称为概括（或父类）。例如，饭店、商店、学校、医院等都涉及到建筑物，所以可以将建筑物抽象出来，形成一种超类，建立饭店、商店、学校、医院等子类的公共属性项和操作。子类还可以进一步分类，如饭店类可以进一步分为餐馆、旅店、涉外宾馆、招待所等类型。所以一个类可能是某个或几个超类的子类，同时又可能是几个子类的超类。

（5）联合（association）：在定义对象时，将同一类对象中的几个具有相同属性值的对象组合起来，为了避免重复，设立一个更高层次的对象来表示那些相同的属性值。例如，某农户拥有两块农田，使用同样的耕种方法，种植同样的庄稼，这里农田主、耕种方法和庄稼三个属性相同，因而可把这两个对象（农田）组合成一个新的对象，而新对象中包含有这三个属性。

（6）聚集（aggregation）：聚集有点类似于联合，但聚集是将几个不同特征的对象组合成一个更高层次的对象。每个不同特征的对象是这个聚集的一部分。它们有自己的属性描述数据和操作，这些是不能为聚集所公用的，但聚集可以从它们那里派生得到一些信息。例如，房子从某种意义上说是一个聚集，因为它是由墙、门、窗、房顶等组成的。

（7）传播（propagation）：传播是作用于联合和聚集的工具，它通过一种强制性的手段将子对象的属性信息传播给高层次的组合对象。就是说，高层次的组合对象，联合和聚集的某些属性值并不单独存于数据库中，而是从它的子对象中提取和派生。例如，一个多边形的位置坐标数据并不直接存于多边形文件中，而是存于弧段和结点文件中。多边形文件仅提供一种组合对象的功能和机制，即借助于传播工具可以得到多边形位置信息。

面向对象技术所提供的机理有以下三个主要的特征：

——应用封装机制（encapsulation）：这种机制使所考虑对象内不仅包含有数据，而且还能提供对其数据进行外部干涉的能力；从而可以支持 GIS 成为多数据类型混合操作平台，并支持模型操作。图 5-4 所示为封装机制示意图。

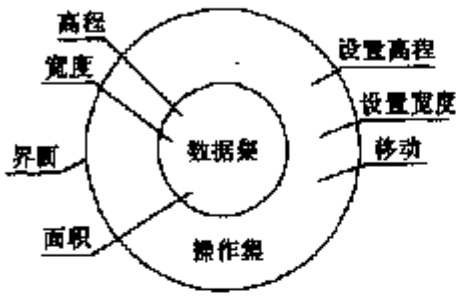


图 5-4　封装机制示意图

——应用继承机制（inheritance）：继承是作用于概括的机制，子类继承其超类（概括）的全部描述，即子类的某些属性和操作来源于它的超类。因此，在同一超类下的子类拥有某些公共属性，并且在继承过程中，还可以将超类的操作和属性遗传给子类的子类。继承是有力的建模工具，有助于进行共享说明和实现几何数据与属性数据一体化，兼容多种数据类型。

——应用多性机制（polymorphism），这种机制使不同类中的对象，对系统发出的同一信息都具有反应能力。例如在绘图时，向不同对象，如方形、圆形或其它形状等，发出绘制（draw）命令时，它们都有反应能力。

目前，国际上已出现了一些面向对象的 GIS 系统，如 Intergraph 的 TIGRIS 系统利用面向对象的抽象数据类型，使几何数据与属性数据一体化，消除了两套管理系统的界面，提高了效率；PRIME 公司的 SYSTEM 9 在现有关系数据库面向记录的数据类型的顶上加一层面向对象的数据模型，应用继承机制对属性数据建立子类与超类等等。但是应该说明，当前面向对象技术在 GIS 中完全实现尚存在一些具体问题。如：

——大对象的操作问题。目前对象的大小仍然受到硬件条件的限制（一般系统生成的对象不超过缓冲区的尺寸）。

——矢量数据中的各个多边形实体，由于受控于同一的操作方法以及公共边的存在，难于在物理上选为具有独立意义的对象并进行独立的存储和封装。

——面向对象的运作机制是消息的传递，而不是面向推理的。如何与推理机制相结合还有待于进一步开发。

§5-2　系统目标

地理信息系统是一个信息交换频率高、实施面和受益面广、应用性强的空间信息系统。要实现一个结构完整、功能齐全、技术先进、适合行业管理特点、实用性好的信息系统，必须经过较长时间的努力。因此，科学合理地确定系统的建设目标是非常必要的。

一、确定目标的原则

系统目标是概括全局，决定全面的东西，只有在充分掌握了各种有关的信息，并进行综合分析比较后，才能正确地确定系统的目标。对地理信息系统而言，无论从信息存贮量还是从功能上划分都包含着广大的范围。从数据量看，其范围可以大到一个现代化大城市

或一个地区甚至国家的综合信息系统，小到某个专业部门的管理与维护信息系统。从功能上看，也有一个延伸到两个极端的连续范围，即从完全不具备辅助决策功能到非常强调辅助决策功能之间的广阔范围。

因此，在这样一种广泛的可能性中，要确定比较适宜的系统目标，就需要首先确定择标原则。当前，在确定地理信息系统目标时，通常都遵循如下这样一些原则：

1. 针对性

以提高信息管理的效率，提高信息质量，为决策者提供及时、准确、有效的信息，向社会提供所需信息为出发点。对具体的专业应用要有具体的设计目标，如对城市环境信息系统，就应考虑除完成日常城市环境规划、管理、决策等工作的数据处理之外，还应能对污染源及环境质量的现状进行评价和预测等。此外，还应能利用航天和航空遥感数据等。

2. 阶段性

系统建设要自上向下，从总体到局部地对系统进行全面规划和整体设计，然后再自下向上，由分到总地分期实施。应该做到既有总体结构的描述，又有子系统的划分。比如，在逻辑结构上可以分为三个基本层次：

(1) 直观目录。用尽可能扼要的方式说明系统的所有功能和主要联系，是解释系统的索引；

(2) 概要图。简要地表示主要功能的输入、输出和处理内容，可以用符号和文字表示每个功能中处理活动之间的关系；

(3) 详细图。详细地用接近编制程序的结构描述每个功能，使用必要的图表和文字说明，再向下则可进入程序框图。

3. 实用性

根据我国现行地理信息系统发展状况，大多数单位（或城市、地区）都难以在短期内建成一个完善的系统，为充分发挥系统的经济效益和社会效益，应注重实用性。初期建设重点在数据建库、处理与查询等工作上。所谓实用系统，不仅要考虑诸如算法设计、软件开发、模型建立等方面的方法和手段，而且还要考虑大量数据的存贮、维护与更新的方法。系统的生命周期应该包括系统的运行与维护阶段，应是一个相当长的时期，而不是仅到系统建成之日为止的相对短的时期。此外，应尽可能考虑建立一个与行政管理体系相适应的系统，因为 GIS 的最终目标是为社会的政治经济服务，它的成功运行与行政管理体制有密切的关系。反之，一个新系统投入运行后，又必然会对现行系统及行政管理体制产生巨大影响。

4. 预见性

要充分考虑国家对有关行业管理的政策、方针和立法以及当今信息技术的快速发展，在系统功能设置时应留有发展余地和良好的接口。系统的功能、系统管理的数据、系统的应用领域以及硬软件均应可扩展，尽量建成一个可扩展的系统。

5. 先进性

要考虑计算机及外设、基础软件的新版本，新的操作系统等先进设备、先进技术的应用。

二、确定目标的依据

在确定系统目标时主要考虑以下因素：用户需求、经费、系统建设时间的要求、技术条件以及数据情况等作为确定依据。

1. 用户需求

系统目标必须围绕用户需求来确定。对用户的需求应在广泛调查的基础上，进行综合分析，权衡利弊得失，从而确定适当的目标。围绕这个问题，要进行以下的工作：

(1) 传统需求的调查与分析；

(2) 潜在需求的探索；

(3) 对需求进行分类；

(4) 基于多数用户的主要需求及满足需求的可能性来确立系统目标。

2. 经费

经费是制约系统目标的主要因素之一。建设一个地理信息系统需要有大量的投入。在我国当前情况下，争取足够的经费是相当困难的，因此在确定系统目标时，只能量体裁衣。当然，系统运行后会带来一定的收益，而且由于系统运行所显示的效果还可能引起新的投资兴趣，这些因素在确定系统目标时也应考虑在内。

3. 系统建设时间的要求

地理信息系统的建设是一项复杂的系统工程，一般需要较长的时间。但是如将系统建设时间规定得很长，不易为领导和用户所理解和接受。因此建设时间也就成了影响系统目标的一个因素。对于大的系统只能考虑分阶段实施的方案，通常将阶段时间划为 3～5 年是比较适宜的。

4. 技术条件

系统目标的确定受技术条件的明显制约，当前可利用的硬软件水平，特别是参加系统建设工作的人员素质与技术水平，都应当实事求是地进行评价，在此基础上才能恰当地确定目标。

5. 数据

数据是地理信息系统的核心。数据的状况对系统目标的影响很大。在考虑系统目标时，需要分析数据的拥有程度、数量、质量、更新频度、使用频率等。

三、近期目标和远期目标

一个完善的地理信息系统的建立需要较长的时间，通常持续几年的项目并不少见，为使系统能尽早地发挥其社会和经济效益，可以分阶段设立系统的近期目标和远期目标。应该说明，不同的行业、不同的要求以及不同的条件，对选择的系统目标肯定也是不同的。这里拟通过一个例子说明选择近期目标和远期目标的考虑因素和角度。

某土地管理部门拟建立土地信息系统，经过调查研究后决定其系统建设的目标可分为近期目标和中远期目标。

1. 近期目标

建成一个以土地信息的规范化管理为基础，以信息的存贮、处理、查询与分析为基本功能，为各级土地管理部门的管理工作服务的计算机网络系统，实现土地信息的手工作业

管理向计算机管理的转换。具体目标为：

(1) 土地信息管理的标准化和规范化，包括制定土地信息的指标体系、分类 编码体系、调整信息收集渠道和采集方式；

(2) 建立各级土地管理的共享数据库；

(3) 建立各行业的专业分析模型；

(4) 联网形成分布式土地信息系统；

(5) 实现对土地利用现状变化的动态监测。

2. 中远期目标

系统建设采用先进的技术，进行更广泛、更快捷的信息采集，对土地信息资源进行深度利用，为土地规划、计划和决策支持服务。

(1) 扩展和完善土地信息系统的网络化，建成对土地资源实施动态监测的业务运行系统；

(2) 建立和完善基础数据库和主题数据库、方法库和模型库；

(3) 建立面向土地全程管理的决策支持业务系统。

最后形成一个高度协调化、信息交流网络化和信息分析智能化的系统。

§5-3 系统分析

系统分析的基本思想是从系统观点出发，通过对事物进行分析与综合，找出各种可行的方案，为系统设计提出依据。它的任务是对系统用户进行需求调查，对选定的对象进行初步调查研究和可行性分析；在明确系统目标的基础上，开展对新系统的深入调查研究和分析；最后提出新系统的结构方案。系统分析是使设计达到合理、优化的重要步骤。这个阶段的工作深入与否，直接影响到将来新系统的设计质量和实用，因此必须给予高度重视。

一、用户需求分析

在地理信息系统开发的准备阶段，应该首先对用户的要求和对用户的情况进行调查和分析。通过调查，确定系统的用户结构、不同用户对应用界面和程序接口的要求，以及系统应具备的功能等等。

地理信息系统的用户有其特定的目的，对GIS有不同的要求，应用需求情况也各异。大体上可以分为三类情况：

1. 具有明确而固定任务的用户。这类用户希望用GIS来实现现有工作业务的现代化，改善数据采集、分析、表示方法及过程，并用以对工作领域的前景进行评估，以及对现有技术方法的更新改造等。通常这类用户多为那些测量调查单位、制图单位和应用科学部门等，有明确的待解决的问题。

2. 具有部分明确固定的工作任务，且有大量业务有待开拓与发展，需要建立GIS来开拓他们的工作。这类用户多为一些环境部门如土地、森林、水利等调查部门和一些可能要进行系列专题调查的单位，还有一些进行特殊项目调查和研究工作的单位。这些单位和部

门属于GIS的潜在用户，因为他们需要将空间数据组织在一起，形成统一的系统供各单位使用。对这类用户而言，由于各个部门和单位的使用侧重点不同，对数据的要求各异，因此在解决数据标准问题的数据结构，建立一致的精度标准等问题上较困难。通常的解决办法是各应用部门聘用自己的软件人员对通用的商业GIS软件进行二次开发。

3. 用户的工作任务不确定，由于各项工作的要求不同，对信息的需求是未知的或是可变的。一般情况下，高等院校和科研机关多属这种情况。他们将地理信息系统作为科学研究工具，或者是为了开发新的地理信息系统技术等。这类用户所需要的GIS差别较大，有的希望有功能全面的GIS来从事各种科研工作，有的则希望在功能一般的GIS基础上开发和研制成多功能的GIS。

综上所述，可将目前国内地理信息系统领域主要涉及的部门或行业大致分为三类。即：

——基层生产管理部门，用于对资源与环境信息和社会经济统计信息的存贮管理和规划决策；

——地学科研人员，用于对资源与环境信息进行系统综合分析和模拟实验，以发现自然规律，特别是空间规律；

——地理信息系统科研和教学部门，用于GIS的理论和方法研究以及教学实践活动等。

总之，对用户需求情况的调查和分析，是地理信息系统设计的基础，通过与系统用户进行的书面或口头的交流，将得到的信息根据设计要求归纳整理后，得到对系统的概略描述。对用户需求情况的调查和分析内容包括了用户的范围、领域、类型、数量、基础以及对GIS了解掌握的程度等。

二、可行性分析

可行性分析是在对用户需求分析基础上，根据社会、经济和技术条件确定系统开发的必要性和可能性，通常要考虑的因素有：①效益分析；②经费问题；③进度预测；④技术水平；⑤有关部门和用户的支持程度。

具体地讲，可行性分析就是从社会因素、技术因素和经济因素三大方面对建立地理信息系统的必要性和实现系统目标的可能性进行分析，以确定用户实力、系统环境、原始数据、数据流量、存贮空间、软件系统、经费预算以及时间分析和效益分析等。

实际工作中，这项工作是与对用户需求调查工作同时进行的。在进行大量的现状调查基础上论证地理信息系统的自动化程度、涉及的技术范围、投资数量以及可能收到的效益等，然后确定GIS的基本起始点，从这个起始点出发就能逐步向未来的目标发展。此外，这项工作还与数据源的调查和评估有密切关系。有关GIS技术人员在掌握了用户需求信息情况后，还应进一步掌握数据情况。例如，对GIS系统提供的数据结构、数据模型与应用所涉及到的专业数据的特征和结构进行适宜性分析；分析研究什么样数据能变换成需要的信息；还要对现有数据形式、精度问题、流通程度进行分析，以确定它们的可用性和欠缺数据的采集方法等。另外，还要对地理信息系统采用计算机系统的处理数据能力，数据库结构和数据大小以及输出形式和数据质量等问题进行分析。例如，微机型GIS与工作站GIS

在功能上有一定的差别，这主要取决于计算机CPU的运算速度、内存容量、存储介质等硬件技术条件。

现以国外早期某地理信息系统的设计为例，说明进行可行性分析的工作。该系统为进行开发可行性分析，先后两次进行原始数据的调研和经济效益的比较分析。通过对原始数据的调研，估算出可供输入系统的地图数量达到1600～3000幅（见表5-1）。

表5-1

地图内容	调研时间	
	1963年	1965年
农业	135～400*幅	135～500*幅
森林	200	200
野生生物	135～150	135～150
娱乐	135～150	135～150
人口统计	500**	500**
气候	2～14***	2～14***
土地利用	400～500	1200～1500**
估计总数	1600～1700幅	2500～3000幅

注：　*　指比例尺为1∶5万～1∶25万地图

　　**　指比例尺为1∶5万～1∶5万地图

　　***　指比例尺为1∶200万地图

　　无星号者均为大于1∶5万的大比例尺地图

这些图主要来自土地调查局，国家统计局和一些地理研究所，能源矿产和资源部门。不仅有地图资料、统计资料，还有可直接输入计算机的人口统计磁带。

为了对计算机系统进行经济可行性分析，分别采用了三种不同方法，测试在70万平方公里面积内由两种基本要素得出的30项统计数据所需要的费用、时间和人力（见表5-2）。证明了计算机地图量算和分析方法大大优于手工方法，而且原始数据经过编码，可以供进一步处理和分析应用。通过分析，估算出建立该系统的数据库大约需要2年（见表5-3）。系统每年的经费投资为50万美元，总投资额大约为100万美元。

表5-2

测试方法	费用（美元）	人工（人）
1. 在1∶25万地图上进行手工地图量算	879 292	58
2. 在1∶5万地图上进行手工地图量算	8 414 145	556
3. 在1∶5万地图上进行计算机地图量算	1 112 202	13～27

表 5-3

来　　源	图幅数量	比例尺	完成年限	速度
1963 年技术可行性报告	1600—1700	大比例尺	2 年	1000 幅/年
1965 年经济可行性报告	2500—3000	大比例尺	3 年	1200 幅/年
1966 年 GIS 项目讨论	22000	1∶5 万	5 年	4400 幅/年
1971 年地图生产估计	2365	1∶25 万	2 年	5 幅/年
1973 年加拿大环境部	2420	1∶25 万	—	60—70 幅/月

三、系统结构方案分析

在调查分析的基础上，明确系统的目标，弄清用户要解决什么问题，各个阶段达到什么要求，并确定系统数据关系的各项配置，提出系统结构方案，作为系统研制的基础和依据。

1. 对现行管理系统的分析

对本部门或其他有关部门的情况进行分析，例如，在对某国家级 GIS 的设计中从上至下分析本部门各级机构在目前和将来组织和发展业务上需要的功能和信息；从下至上地分析他们完成本部门专业活动所需要的数据和所采用的处理手段。从现行的管理体制划分可分为五级，即国家、省（区）、地（市）、县（市）和乡镇。从纵向管理联系上，逐级分析各级的管理职能、指导关系和监控联系以及各类数据的流向等。对数据采集体系和各种指标的规范化与否，数据和信息传递和反馈响应速度的快慢，信息处理能力的强弱，信息综合使用效果的好坏等方面进行分析。

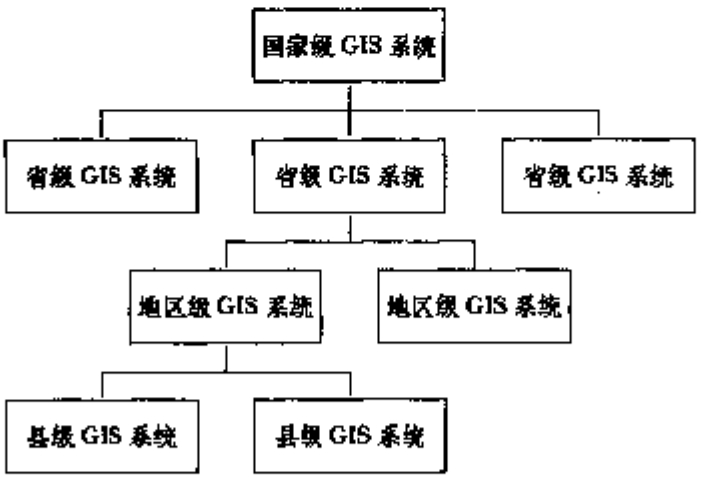


图 5-5　信息系统逻辑结构图

2. 系统总体逻辑结构

待建地理信息系统的结构构成应能适应现有管理体制的需求，并建立在对现行管理模式进行改革的基础上。建设一个结构合理、数据规范、信息全面、响应快速的计算机化的

GIS 系统，能够及时提供决策信息并向社会提供高质量的信息服务是对系统总体结构的原则性要求。

根据系统开发的目标和管理模式，制定系统数据流程图和数据组织方案，建立系统的逻辑结构。例如，对一个国家级的 GIS 系统，其系统逻辑结构可如图 5-5 所示。

3. 子系统的划分

系统总体方案的分析远达不到总体设计所需的系统分析资料那样详尽，其目的仅在于为子系统信息库的划分、子系统的功能以及与整个系统的功能设置提供有关的资料。

子系统的划分是通过分析总体管理模式与数据间的内在联系，合理地确定每个子系统业务运行范围和数据处理过程的一项系统分析过程。其目的是避免系统业务活动的重叠和数据处理过程以及存贮过程中产生的混乱，合理地实现系统目标。同时，子系统的划分给系统的逻辑设计和物理设计打下基础，为整个系统的运行提供保证。通常，子系统的划分应尽量遵守以下原则：

——子系统对其他子系统的数据依赖应尽可能小；

——子系统所包含的各个过程之间内在联系应尽可能强；

——子系统的划分应便于总的系统设计阶段实现。

应该说明，各级地理信息系统子系统的规模和功能因所处级别管理职能不同而有所差异，各级系统的功能主要从各级子系统的功能体现出来。例如，图 5-5 所示的国家级信息系统，在一般情况下，国家级和省（区）级主要是进行宏观管理，很少实施行业管理的具体业务，而县（市）和地（市）级则经常要大量处理最基础的管理业务，所以，对子系统进行划分时，可以划分成国家级、省（市）级信息系统的子系统和地（市）级、县（市）级的子系统两种方案。如图 5-6 (*a*)，图 5-6 (*b*) 所示为某土地信息系统的子系统结构示意图。

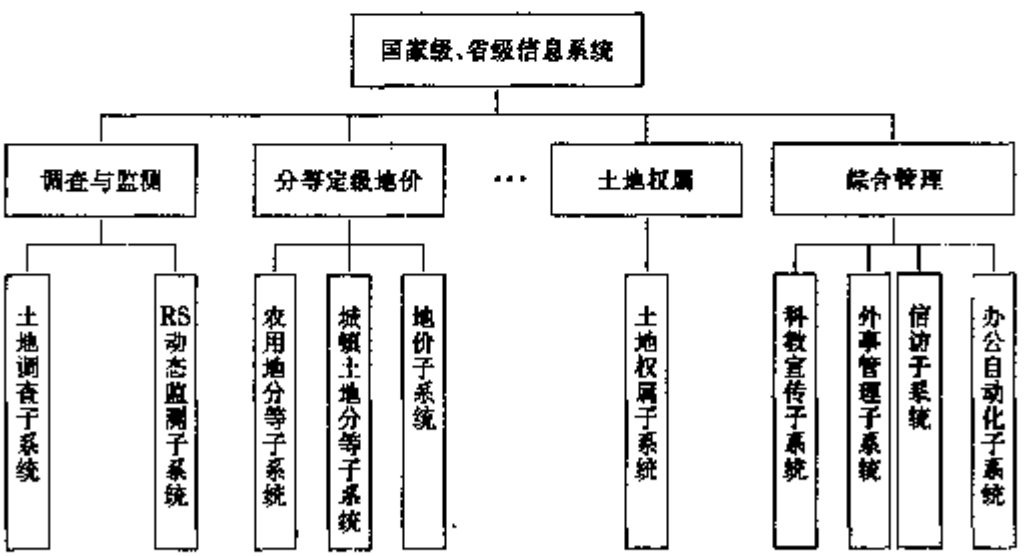


图 5-6 (*a*)　国家级、省级土地信息系统子系统结构示意图

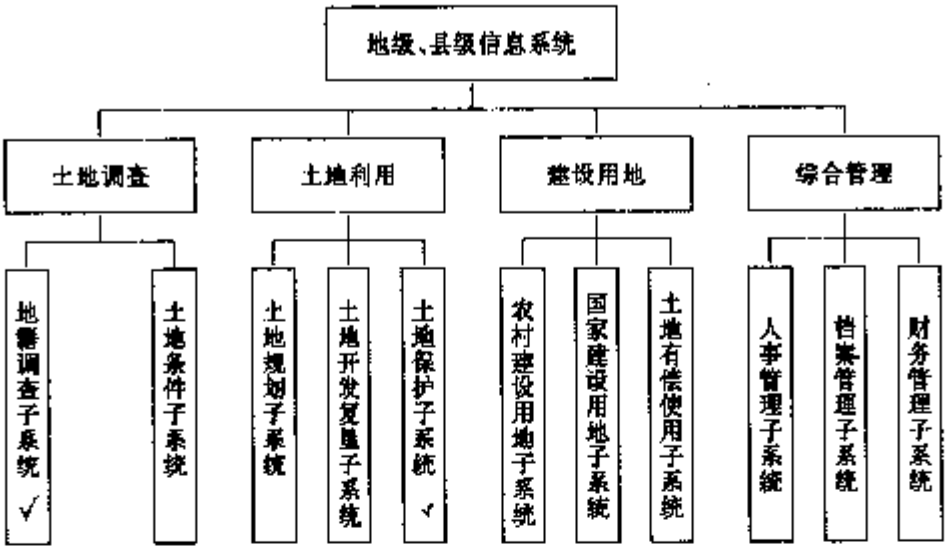


图 5-6 (b)　地（市）级、县（市）级土地信息系统子系统结构示意图

§5-4　系统设计

因为系统目标的不同以及所用数据的性质和系统功能的不同，地理信息系统设计的内容也有很大差异，但是其根本任务是将系统分析阶段提出的逻辑模型转化为相应的物理模型。一般而言，在系统设计阶段可以根据所研究对象的不同分成三个部分进行设计。首先应根据系统研制目标，确定系统必须具备的空间操作功能，称为功能设计又称为系统的总体设计。通常可以采用模块化程序设计方法；其次是对数据分类和编码的处理，完成空间数据的存贮和管理，称为数据库设计，含有数据采集设计、数据结构设计、数据存贮和检索设计等；最后是建立系统的应用模型和产品的输出，称为应用设计。

系统设计是地理信息系统整个研制工作的核心，不但要完成逻辑模型所规定的任务，而且要使所设计的系统达到优化。如何选择最优的方案，是系统设计人员和用户共同关心的问题。一般而言，一个优化的地理信息系统必须具有运行效率高、控制性能好和可变性强等特点。要提高系统的运行效率，通常的做法是尽量避免中间文件的建立，减少文件扫描的遍数，并尽量采用经过优化的数据处理算法。为增强系统的控制能力，在输入数据时，要拟定对数字和字符出错时的检验方法；在使用数据文件时，要设立口令，防止数据泄密和被非法修改，保证只能通过特定的通道存取数据。为了提高系统的可变性，目前较有效的方法是采用模块化的结构设计方法，即先将整个系统看成为一个模块，然后按功能逐步分解为若干个第一层模块、第二层模块等等。一个模块只执行一种功能，一种功能只用一个模块来实现，这样设计出来的系统才能做到可变性好和具有生命力。

一、功能设计

一个信息系统有无生命力，主要看系统对事务的处理是否满足应用的要求，即系统具有哪些功能以及这些功能处理事务的能力。因此，功能设计或总体设计的主要任务是根据

系统研制的目标来规划系统的规模和确定系统的各个组成部分，并说明它们在整个系统中的作用与相互关系以及确定系统的硬件配置，规定系统采用的技术规范，以保证系统总体目标的实现。

1. 总体构成

地理信息系统由硬件、软件（含系统软件与应用软件），数据和人员（管理人员，开发人员和用户）四部分构成（详见第二章）。如果从总体功能上划分，大致可分为数据输入子系统、数据处理子系统和数据输出子系统三大部分。图 5-7 和图 5-8 所示分别为两个地理信息系统的总体结构示意图和控制结构示意图。

2. 硬件配置

硬件设备的投资在地理信息系统总投资中往往占很大比重，因此，在选择硬件设备时主要是针对每级系统的功能和所要完成的工作来考虑。例如，对于国家级、省（区）级信息系统而言，由于涉及到的数据种类多和数据量大，通常都选择大而全的设备，而对地（市）级和县（市）级的系统而言，一般选择相对简单一些的设备。此外，对硬件设备的选择还要根据软件的要求和软件类型来确定。一般情况下，软件的设计是按一定的机型和外围设备设计的，常常只能支持一定型号的硬设备。因此，在考虑硬件选购时应知道打算购买的软件能否支持，否则可能会造成运行困难。还应该说明一点的是，随着计算机技术的飞速发展，硬件设备的更新换代速度亦越来越快，且价格亦越来越低。因此，在购买硬件设备时应有优先顺序，即首先购买工作开始时就必须使用的设备，今后有用而目前暂时不用的设备留待以后购置。表 5-4 所示为某系统的硬件设备配置情况。

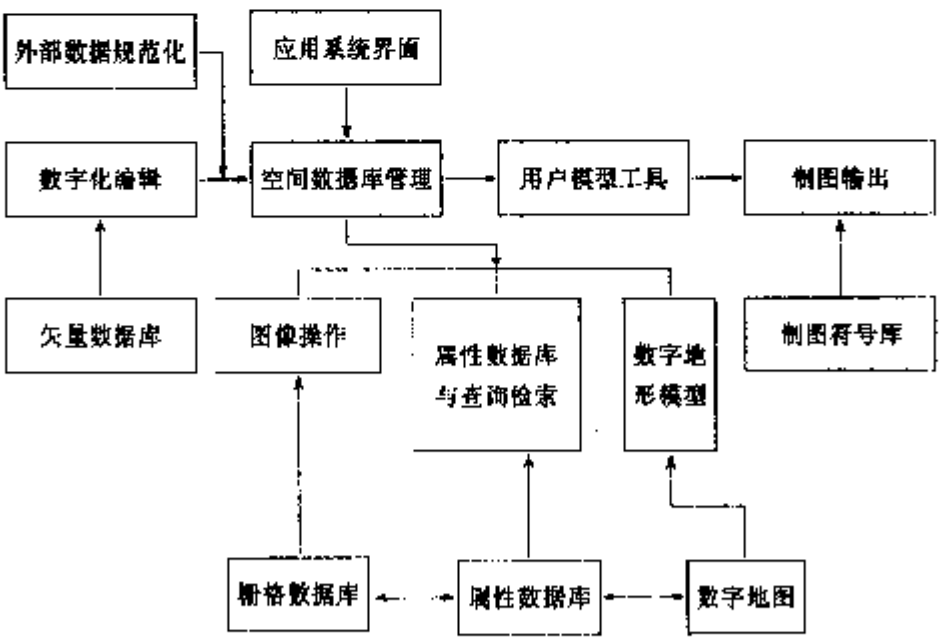


图 5-7　某 GIS 总体结构示意图

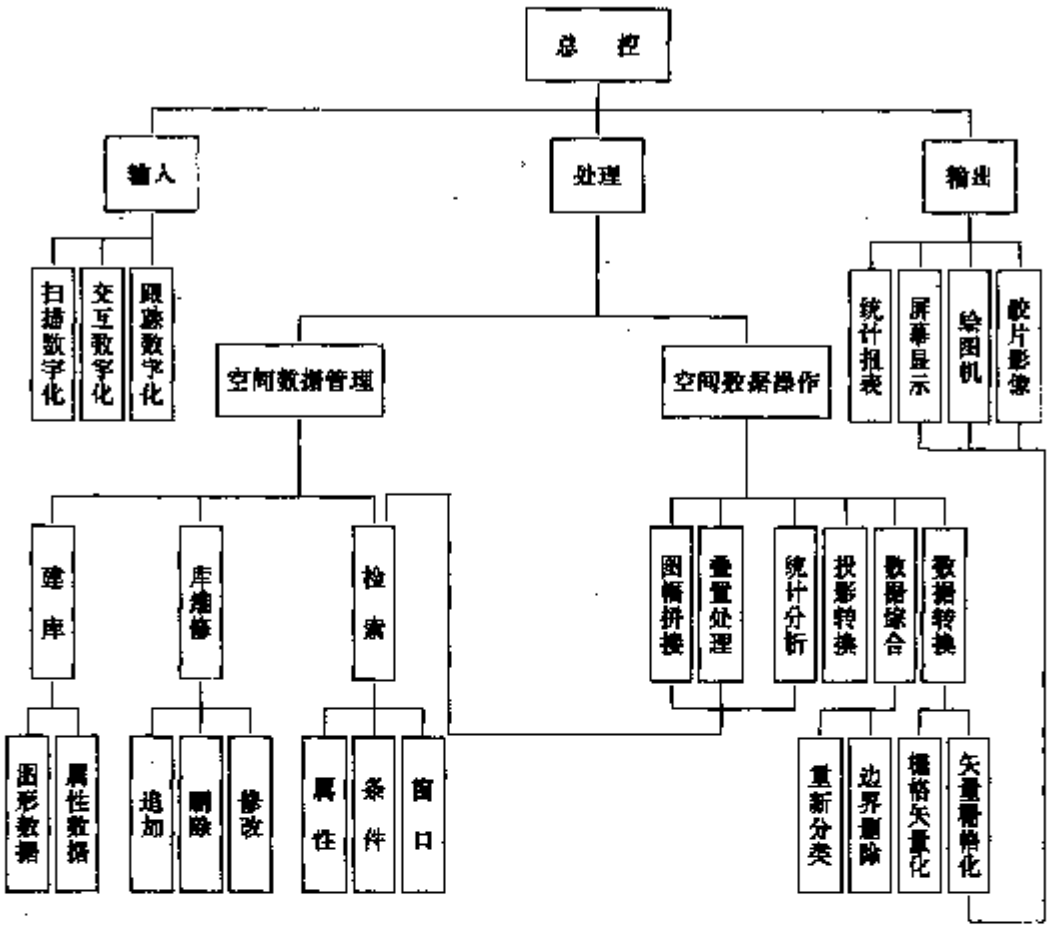


图 5-8 某 GIS 控制结构示意图

表 5-4 某系统硬件配置表

类 型	型 号	数量	性 能
服务器	SUN SPARCserver 470	1	内存 32 M CPU 33 MHz/22 MIPS 磁盘 1.3 GB×3+2.1 GB=6.0 GB
磁带机		1	1/2 in 150 MB 标准工业磁带
磁带机		1	1/4 in 150 MB
磁带机		2	8 mm 5 GB
光盘机	RICOH RS-9200EX	1	650 MB 可重写

续表

类，型	型　号	数量	性　　能
工作站	SUN SPARCstation 2	1	内存 32 MB　CPU 速度 28 MIPS　GX 图形加速器，19 in 彩显（1280×1024） 磁盘 210 MB + 2.1 GB（内置 SCSI）+ 1.3 GB（外置 SCSI） NewsPrint 卡（连 SPARCprinter）
工作站	SUN SPARCstation IPX	7	内存 16 MB　CPU 速度 28 MIPS GX 图形加速器，17 in 彩显（1024×800） 磁盘 210 MB+420 MB（外置 SCSI）
工作站	HUASUN station 1+	2	内存 16 MB　CPU 速度 22 MIPS 17 in 彩显（1150×900），磁盘 200 MB
数字化仪	Calcomp 9600	4	A0 幅面　高精度卡 分辨率最高可达 10000 线/in
数字化仪	Calcomp 33360	3	A1 幅面　分辨率 1000 线/in
数字化仪	长地 CD-9910B	2	A1 幅面　分辨率 2450 线/in
扫描仪	ScanMaker I	1	A4 幅面　分辨率 600 dpi
打印机	EPSON LQ-1600K	2	
打印机	CI-1000K	1	高速行打
打印机	SPARC printer	1	激光打印机 300 dpi
打印机	HP LaserJet 4	1	激光打印机 600 dpi
微　机	Compaq 486/33i	8	内存 4 MB 磁盘 120 MB VGA 显示卡（640×480）
微　机	Compaq 386/20e	3	内存 4 MB 磁盘 120 MB VGA 显示卡（640×480）
微　机	Olivetti M386/25	1	内存 10 MB 磁盘 120 MB VGA 显示卡（640×480）

续表

类　型	型　号	数量	性　　能
微　机	SUN/386/20	1	内存 2 MB 磁盘 40 MB VGA 显示卡（640×480）
微　机	AST 286	1	内存 1 MB 磁盘 40 MB EGA 显示卡
大屏幕 显示器	Mitsubishi（三菱） HJ 6905	5	分辨率 1280×1024（逐行扫描 4） （隔行扫描 1）
字符终端	STAR-500	6	福建实达公司产品 640×450 单色显示器
绘图仪	Roland DPX-3500	1	A1 幅面 8 笔 平板式
绘图仪	Calcomp 2024	2	A1 幅面 8 笔 滚筒式
绘图仪	Calcomp 1077	1	A0 幅面 4 笔 滚筒式

3. 软件模块功能

在进行地理信息系统设计时，由于各级系统的目标不同，因此要求的功能也不尽相同。第四章中已对系统的软件模块功能做过讨论，这里仅就一些地理信息系统的常见模块功能进行简单归纳和描述。

（1）数据采集模块功能

数据种类主要为空间（定位）数据、属性数据以及一部分管理数据。一般要求系统具有多种采集方式来采集空间数据，并且有精度要求。常用的采集方式有：

——手扶数字化输入；

——扫描数字化输入；

——GPS 接收机；

——航测仪器，全站型电子速测仪；

——卫星遥感影像数据；

——键盘输入；

——与其它系统的数据交换。

（2）图形处理模块功能

可以完成对图形的显示、查询、编辑、修改和管理工作，其主要处理功能为：

——人机对话，有友好的用户界面；

——图幅定向，将图幅坐标归化为地理坐标；

——图形窗口显示，提供修改、查询、编辑操作的区域，有缩放、漫游和分层显示功能；

——符号设计与图形整饰，建立符号库且有自动生成各种符号的工具；

——图形编辑，具有增删、连接、断开、移动、旋转功能和图形拷贝功能；

——图形的拓扑关系，建立图形元素之间的拓扑关系；

——属性数据的编辑，实现属性数据与空间数据的连接；

——几何图形计算，计算面积、周长、边长、点到线距离等；

——图形、属性之间的查询，实现由图形查属性，由属性查图形的功能；

——图形接边处理，可以消除几何裂隙和逻辑裂隙。

(3) 属性数据管理模块功能

——向用户提供定义各类地物的属性数据结构和用户自定义数据结构的功能；

——可以对数据结构进行修改、拷贝、删除、合并的功能；

——利用结构化查询语言（SQL）提供多种灵活的数据库查询；

——提供数据计算统计和统计分析功能；

(4) 制图输出模块功能

——在图形输出前，用户可以根据需要添加符号、颜色、注记、图例、并对图廓进行整饰；

——具备与多种输出设备的类别（打印、笔式、喷墨、热腊、静电、制版等）和型号相兼容的接口软件和绘图指令；

——能够向用户提供矢量图、栅格网、全要素图和各种专题图。

(5) 空间分析模块功能

——叠置分析，将同比例尺、同一区域的两组或多组图形要素的数据文件进行叠置得到新的图形和新的属性统计数据；

——缓冲区分析，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；

——空间集合分析，按照两个逻辑子集给定的条件进行逻辑交、逻辑并、逻辑差运算；

地学分析，如利用网络分析模块进行最佳路径分析、土地适应性分析、发展预测分析等。

(6) 地形分析模块功能

——数字地形模型（DTM），由等高线或不规则三角网（TIN）产生地面高程模型（DEM），可进行高程分级，地面参数计算（坡度、坡向辐照度、地面粗糙度等），三维立体模型多角度方位显示；

——地形分析，包括等高线分析、透视图分析、断面图分析、地形表面面积和挖填方体积计算。

(7) 图像处理模块功能

为保证系统的动态性和现势性，可利用遥感技术更新系统数据库的内容。其基本功能应包括：

——遥感数据的输入；

——画面显示、操作、坐标量测、色调变更等；

——几何校正，能从具有几何畸变的图像中消除畸变；

——图像增强，能使分析者容易地识别图像内容，按照分析目的对图像数据进行如灰度变换、彩色合成等处理；

——特征提取，把图像的特征进行量化处理；

——栅格数据矢量化处理；

——地面定位，能利用地理数据（三角点、地图数据、全球定位系统GPS）与遥感图像匹配；

——输出功能，具有胶片输出和数字输出功能。

二、数据库设计

数据库是地理信息系统的核心组成部分，根据不同的应用，数据库会有各种各样的组织形式。数据库设计就是把现实世界中一定范围内存在着的应用处理和数据抽象成一个数据库的具体结构的过程。具体地讲，就是对于一个给定的应用环境，提供一个确定最优数据模型与处理模式的逻辑设计，以及一个确定数据库存贮结构与存取方法的物理设计，建立能反映现实世界信息和信息联系，满足用户要求，又能被某个数据库管理系统（DBMS）所接受，同时能实现系统目标并有效存取数据的数据库。

1. 数据库设计目标

（1）满足用户要求。设计者必须充分理解用户各方面的要求与约束条件，尽可能精确地定义系统的需求。

（2）良好的数据库性能。数据库性能包括多方面的内容，在数据存贮方面既要考虑数据的存贮效率又要顾及其存取效率；在应用方面，不仅要满足当前应用之需要，又要能满足一个时期内的需求可能；在系统方面，当软件环境改变时，容易修改和移植。另外，还要有较强的安全保护功能。通常，上述性能往往有些冲突。因此数据库设计必须从多方面考虑，对这些性能作出最佳的权衡折衷。

（3）对现实世界模拟的精确程度。数据库通过数据模型来模拟现实世界的信息类别与信息之间的联系。数据库模拟现实世界的精确程度取决于两方面的因素：一是所用数据模型的特性；二是数据库设计质量。就目前情况而言，现有数据模型对于一般的信息系统能够表示现实世界中各种各样的数据组织以及数据之间的联系，所以能否精确描述现实世界的关键还在于数据库设计者的能力和水平。为了提高设计质量，必须充分理解用户要求，掌握系统环境，利用良好的软件工程规范和工具，充分发挥数据库管理系统的特点。

（4）能被某个数据库管理系统接受。数据库设计的最终结果，是确定数据库管理系统支持下能运行的数据模型和处理模型，建立起可用、有效的数据库。因此，在设计中，必须了解数据库管理系统的主要功能和组成。尽管数据库管理系统的功能因不同的系统而有所差异，但一般都应具有以下主要功能：

——数据库定义功能：提供定义概念模型、外部模型和内部模型的能力，勾通出数据库的框架；

——数据库管理功能：对整个数据库的运行控制、数据存取、更新管理、数据完整性及有效性控制以及数据共享时的并发控制等；

——数据库维护功能：数据库重新定义、数据重新组织、性能监督和分析以及发生故障时恢复运行等；

——数据库通讯功能：包括与操作系统的接口处理，与各种语言的接口以及与远程操

作的接口处理等。

为实现上述功能，数据库管理系统主要由如下三个部分的程序组成：

——语言处理程序组；

——系统运行控制程序组；

——建立和维护程序组。

2. 数据库逻辑设计

数据库逻辑设计的任务是运用数据库管理系统提供的工具与环境，将对现实世界抽象得到的概念性模型转换成相应的数据库管理系统的数据模型，并用数据描述语言描述出来。因此，逻辑设计是整个数据库设计的基础。其目的是要规划出整个数据库的框架、回答数据库能够做什么的问题。通过逻辑设计形成的相应数据库的数据模型应该独立于计算机的硬件和软件，并且是面向应用，易于为用户理解。

逻辑设计应该达到如下几点要求：

(1) 在共享数据资源方面，在降低数据采集、存贮和使用成本方面以及在数据维护的事务处理方面都应达到最大的效率。通常主要考虑的问题是：处理速度、吞吐量、响应时间、可维护性和存贮需求。

(2) 在数据质量方面要达到防止（尽量减少）数据冗余，保持数据内容与格式的一致。

(3) 要能最大限度地发挥系统的性能。考虑应能达到支持多种用户视图，有利于扩展用户应用开发的领域以及保持数据检索、分析和生成的灵活性。

(4) 维护数据的独立性。

如上所述，逻辑设计的核心任务是找出相关的数据模型，而该模型能够表达现实世界中的实体及其联系的模型。在§3-5中已经详细讨论了由“实体一联系”模型向数据模型，如层次模型、网状模型和关系模型的转换问题。应该说明，关系模型是潜力较大的模型，也是目前地理信息系统上采用较多的模型。它基于严格的数学理论，由于关系代数的引进，使得对数据库的各种检索运算要求得到较大满足，数据库易于更新，地理数据易于描述为二元关系表格，保持了系统数据与用户习惯的一致性。另外，从使用角度看，系统设计和管理人员与用户使用同一种结构，也便于与用户的交流。

地理信息系统具有处理数据量大、结构复杂等特点，为了便于管理和应用开发，经常在设计时将整个系统划分为一些子系统，与此相适应，数据库也被划分为若干子库，此外，对于一些比较大的或比较复杂的子数据库还要进一步划分。这种划分通常有两种途径：一是“纵向”划分，即按照数据的性质分类，将性质相同或相近的归为一类，形成所谓的数据层（Layer）的概念；另一种是“横向”划分，即按数据的空间分布将数据划分为规则的或不规则的所谓片(Tile)的概念。以上两种数据划分的途径可以同时应用于同一个数据库，也可以分别采用。

需要说明，层的概念既适用于栅格数据结构，也适用于矢量数据结构。栅格数据结构可以按每种属性数据形成一个独立的数据层来设置。栅格结构中新的属性就意味着在数据库中增加一个新的数据层，各个数据层在一起则形成栅格三维数据阵列。原则上，层的数量是不受限制的，但是由于存储空间的有限性限制了层的数量。矢量数据结构通常是用层

来区分空间实体的主要类别，目的是为了制图和显示，这种方法在计算机辅助设计系统中很频繁地使用。

数据分层可以按①专题，②时间，③垂直高度等方式来划分。按专题分层就是每层对应一个专题，包含一种或几种不同的信息服务于某一特定的用途或目的。例如，用于城市规划的数据层可以按街道、公交路线、交通工具、税收、给水排水、电力电讯、文化教育、金融、卫生、旅游、公安消防、区域经济、土地使用情况等方面的专题来划分。按时间序列分层则可以不同时间或时期进行划分。按垂直高度划分是指以地面不同高程来分层，如可分为地上、地下等。

作为参考，下面介绍在数据分层和分片过程中应注意的一些问题。

数据分层时应考虑的问题：

——按要素类型分层，性质相同或相近的要素应放在同一层；

——分层时要考虑数据与数据之间的关系，如哪些数据有公共边，哪些数据之间有隶属关系等，这些因素都将影响层的设置；

——要考虑用户视图的多样性；

——分层时要考虑数据与功能的关系，如哪些数据经常在一起使用，哪些功能是起主导作用的功能等；

——分层时应考虑更新的问题，因为更新一般以层为单位进行处理，所以应考虑将变更频繁的数据分离出来；

——分层时应顾及数据量的大小，各层数据的数据量最好比较均衡；

——尽量减少冗余数据。

数据分片时应考虑的一些问题：

——存取数据的要求，确定典型用户常用的查询范围；

——选择适当的数据量，每片中包含的数据量既不宜过大也不宜过小。过大则增加处理的时间，过长等待会使用户感到难以忍受，过小则会给管理和查询带来不便；

——分片时往往需要经过典型试验以确定最佳方案。

应该说明的是，关于空间数据库的分片类似于模拟地形图的分幅。但片的形状并不一定要求是矩形的或梯形的，所有的片合并起来应能覆盖整个数据的地理区域。由于空间数据库的检索常常要通过地理位置作为索引进行，所以将数据分片有利于建立优化的索引系统，达到数据空间位置的适配性和属性的一致性。此外，合理的分片还能实现较好的数据管理与维护。

3. 数据库物理设计

数据库物理设计的任务是使数据库的逻辑结构能在实际的物理存贮设备上得以实现，建立一个具有较好性能的物理数据库。数据库物理设计主要解决以下三个问题：恰当地分配存贮空间；决定数据的物理表示；确定存贮结构。

存储空间的分配应遵循两个原则：①存取频度高的数据存储在快速、随机设备上，存取频度低的数据存储在慢速设备上；②相互依赖性强的数据应尽量存储在相邻的空间上。

数据的物理表示可分为两类：数值数据和字符数据。数值数据可以用十进制形式或二

进制形式表示。通常，二进制形式占用较少的存储空间。字符数据可以用字符串的方式表示，有时，也可以利用代码值的存储代替字符串的存储。为了节约存储空间常常采用数据压缩技术，这在设计地理数据库时尤为重要。

存储结构的选择与应用要求有密切的联系。对批处理应用的数据，一般以顺序方式组织数据为好；对于随机应用的数据，则以直接方式或索引方法比较好，同时用指针链接法建立数据间的联系。物理设计在很大程度上与选用的数据库管理系统有关。设计中应根据实际需要，选用系统所提供的功能。

4. 数据字典

数据库逻辑设计和物理设计都属于数据库的结构设计。在数据库建设中，作为数据规范的数据字典也同样起着重要的作用。数据字典即所谓关于数据的数据，是对数据库的数据和应用程序的一种管理方法。在数据库的标准化方面，数据字典应该是重点考虑的一个问题。

数据字典一般包括数据类型的名称，关于数据的描述及其存贮的地址如何使用等内容。其描述的主要项目有：数据项、记录、文件、模式、子模式、数据库、数据用途、数据来源、应用模型和用户情况等。

数据字典的功能可以表现在以下几个方面：

——给管理者和用户提供关于可利用数据的线索；

——为系统分析人员提供数据是否存在的信息；

——为编程工作提供数据格式及数据位置。

数据字典的用途是多方面的，它在数据库的整个生命周期里都起着重要的作用。具体可归纳为以下几点：

(1) 在系统分析阶段，数据字典用来定义数据流程图中各个成分之属性与含义；

(2) 在设计阶段，数据字典提供一套工具以维护对系统设计说明的控制，帮助设计人员保证在早期阶段所确定的需求与实现一致；

(3) 在实现阶段，提供了元数据描述（数据库中的数据）的生成能力；

(4) 在调试阶段，辅助产生测试数据，提供数据检查的能力；

(5) 在运行和维护阶段，可帮助数据库的重组织和重新构造；

(6) 在使用阶段，可以作为“用户手册”。

例如，ARC/INFO 系统的数据字典有以下功能：

——数据标准化的出发点；

——辅助应用程序设计；

——辅助数据库设计；

——加强对数据的了解；

——消除冗余数据；

——改善数据的完整性。

ARC/INFO 的数据字典还具有由一系列相关文件构成的层次结构，其具体形式包括：层目录、项目录和编码描述，其关系如图 5-9 所示。

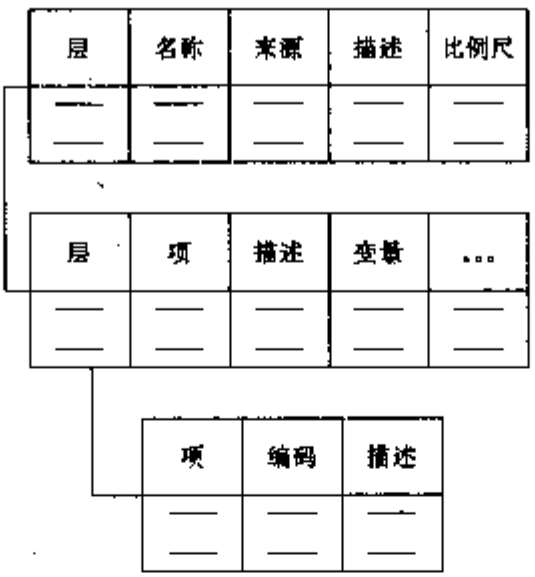


图 5-9 数据字典的层次结构

§5-5 系统实施和系统评价

一、系统实施

系统实施是GIS建设付诸实现的实践阶段。在这一阶段中，需要投入大量的人力物力并占用较长的时间，因此，应该作好细致的组织工作，制订出周密的计划。系统实施的主要内容是程序编制与调试和数据采集与数据库建立。此外还应包括人员的技术培训和系统测试等。

通常，为了保证程序编制和调试及后续工作的顺利进行，硬软件人员首先应进行GIS系统设备的安装和调试工作；然后在适当的开发软件提供的环境下将详细设计产生的每个模块的功能用某种程序语言予以实现；再进行程序调试、数据录入和试运行，以至建立一个能交付用户使用的实用系统。

程序编制工作要尽量做到标准化和通用化。对所编制的程序应该按统一的格式编写程序说明。一般可采用以下内容：

——程序名称；
——程序功能；
——程序设计算法；
——程序使用方法；
——需要的存贮空间、设备和操作系统；
——程序设计语言；
——程序使用的数据文件；
——其它有关说明等。

此外，为充分发挥系统的效率，在程序编制时还应顾及以下几个方面的因素：

——运算速度；

——对硬、软件环境的依赖程度，对计算机内存的要求；

——算法的复杂程度；

——程序的可靠性和适用性。

在对系统的硬、软件进行测试和安装调试过程中，还应该结合系统目标，通过采用模拟数据或选择出小块实验地区的数据，对各种功能进行全面试验。实验阶段不仅应测试系统各部分的工作性能，同时还应测试各部分之间数据传送性能的好坏、处理速度和精度的高低，从而保证所建立的系统能正常工作且各部分运行状况良好。

GIS 建设过程中还需要投入大量的人力进行数据的收集、整理和录入工作。GIS 规模大，数据类型复杂多样，数据的收集和准备是一项既繁琐、劳动强度又大的任务。通常要求数据库模式确定后就应进行数据的输入，并且应按数字化作业方案的要求严格进行，输入人员还应该进行相应程度的培训工作。

在这一阶段，由于有很多人员参加工作，包括系统开发人员、用户和领导管理人员，为了保证 GIS 的调试顺利和使用户尽快掌握使用方法，还应提前对有关开发人员、用户、软件操作人员等进行技术培训以及有关的系统管理维护工作，并制订出相应的工作规程与管理制度等。此外，对于一般人员和领导，也应给予一定的宣传和介绍，使其对新建 GIS 系统有所了解，关心和支持 GIS 的实施工作。

系统的调试与测试是指对新建 GIS 系统进行从上到下的全面测试和检验，看它是否符合系统需求分析所规定的功能要求，发现系统中的错误，保证 GIS 的可靠性。一般来说，应当由系统分析员提供测试标准，制订测试计划，确定测试方法，然后和系统设计人员、程序设计人员及用户共同对系统进行测试。测试的数据可以是模拟的，也可以是来自用户的实际业务，经过新建 GIS 系统的处理和检验，判断输出数据是否符合预期的结果，能否满足用户的实际要求，并对不足之处加以改进，直到满足用户的要求为止。

二、系统评价

所谓系统评价就是对所建立系统的性能进行考察、分析和评判，判断其是否达到系统设计时所预定的效果，包括用实际指标与计划指标进行比较，评价系统目标实现的程度。评价指标应该包括性能指标、经济指标和管理指标等各个方面，最后还应就评价结果形成系统评价报告。具体运作时可以从软件功能和系统总体功能两个方面进行评价。

1. 软件功能评价

GIS 的软件构成整个系统的核心部分，其功能的好坏与否决定了 GIS 系统功能的强弱，因此，正确全面地对软件功能进行评价是很重要的。为满足实际工作的需要，通常可以分五个部分对软件功能进行检验和评价。作为参考，下面具体列出对软件功能进行评价的纲要：

(1) 用户界面

——命令驱动界面有提示，应答界面有缺省应答；

——有下拉式或上托式菜单功能；

——具有交互式命令语言接口；

——可以使用命令缩写；

——可用各种简单功能的集合构成宏指令 shell 文件或批处理文件，以便自动地执行较复杂的功能；

——具有连机帮助屏幕概述各项命令、命令语法、功能等；

——能够取消或重新存储前面的输入值；

——具有与用户友好的错误提示信息；

——具有软错误恢复功能；

——具有口令存取保护功能。

(2) 数据库管理系统

一般功能：

——应能在空间及属性数据库中输入有关数据的世系（即来龙去脉）、位置精度、逻辑一致性、完整性等质量信息；

——具有跟踪数据事务处理的功能；

——支持按顺序，直接和按关键字存取数据文件的方式；

——具有定义文件内容和格式的数据字典；

——除顺序文件存取方式外，可直接存取专门要素；

——允许通过属性或空间数据字段，对表格或图形文件进行分类；

——可使用算术表达式或在相关文件中查表的方式对新的字段值进行计算；

——具有通过共享字段将几个数据文件联系起来的功能，并可将其结果集合当做一个整体用于所有表格处理功能（包括录入数据和生成报告）中；

——在空间和属性数据库中，提供读、写、访问的授权功能；

——提供建立、存储、检索、生成标准报告的功能；

——提供下列表格格式化功能：①在指定字段分行，②分项，③计算总和，④设计页和列的标题规格，⑤从单独记录中进行多行显示。

空间数据库构成：

——可通过以下方式组织空间文件：①位置，②项目，③专题，④地图单元；

——对永久数据文件可以进行多重访问，但只有被授权的用户才可修改数据库；

——对用户建立的工作文件提供完全的、只有该用户自己享用的增、删、改功能；

——对数据库中的全部数据，能够自动建立目录或索引，包括数据的质量、位置及最后日期；

——对数据库的内容和现状能够生成状态报告；

——能够不考虑文件大小或比例尺来添加数据文件。

(3) 数据库建立

数字化：

——数字化方法（手工数字化，扫描数字化，……）；

——加注记的方法；

——建立拓扑关系；

——输入属性数据。

编辑：

——空间数据编辑

① 对拓扑错误进行自动检查，在图形上显示错误，并有人机交互的纠错功能；

② 在数字化过程中或以批处理方式对矢量坐标或栅格像元进行格式检验、范围检验、数值检验；

③ 以单个地物或地物群为单位，人机交互地增加、删除、修改、移动矢量地物或栅格像元；

④ 在数字化过程中或以批处理方式自动检验线的定义处的过头和不足，并能以重新数字化或自动剪裁连接的方式进行改正。

——属性数据编辑

① 人机交互地插入、删除、改变、移动地物名称和编码；

② 可检验丢漏的地物名称或编码；

③ 在数字化过程中或以批处理方式检验非法属性值或属性值的组合；

④ 利用查询语言的选择函数，更新成组的图形地物名或属性记录；

输入和输出：支持多种数据格式的输入输出。

(4) 数据处理和分析

——数据检索：

① 可选择特定的数据类别；

② 通过矩形、圆形或多边形窗口选择空间或属性数据；

③ 从人机交互的屏幕数字化区或重新定义分类的数据类别选择空间或属性数据；

④ 通过单一地物名或一组地物名选择空间或属性数据；

⑤ 通过对属性的逻辑检索选择空间数据；

⑥ 通过图形连接选择空间或属性数据；

⑦ 可浏览空间或属性数据库。

——数据重组：

① 可进行从栅格到矢量，从矢量到栅格的数据转换，具有用户对点线面要素进行选择的优先权；

② 在缺省或用户指定的限差内，交互式地或自动地对几何位置相邻的数据进行连接、消除裂缝或重叠部分；

③ 可用行程编码或四叉树编码进行栅格数据的压缩和释放；

④ 通过重新取样修改网格尺寸；

⑤ 压缩不必要的多余坐标而同时保存角点、弯曲总貌和形状；

⑥ 对直线数据进行平滑，以恢复其弯曲总貌和形状；

⑦ 可从随机状态和网格化的高程值生成等高线，反之，也可从等高线生成网格化的高程值；

⑧ 可从随机状态、网格化的高程值、等高线生成不规则三角网（TIN）；

⑨ 可从不规则三角网生成网格或等高线数据；

⑩ 通过指定的阻挡层（例如断层）或约束条件（例如山脊线、汇水线）抑制等高线的生成；

⑪ 提供下列几何坐标功能：平行线、曲线及各种要素的延伸，产生相等的直线、弧段、相交直线、弧段，等分角，定位交叉切线和外切线等。

——数据变换：

① 矢量或栅格数据对于控制点的数学平差方法是：在 X，Y 方向的旋转、平移、缩放（4 个参数），在 X，Y 方向的旋转、平移、缩放（6 个参数），局部区域弹性图幅，多项式，其他类型的最小二乘法，投影变换、近似变换等；

② 从数字化像片数据中恢复地理坐标的方法是：结合数字高程数据使用单一像片后方交会或前方交会技术，使用模拟或解析测图仪测出立体像片带的坐标；

③ 对已知控制点的导线数据，使用最小二乘平差法将地面测量的方位和距离转换为地理坐标；

④ 辐射校准遥感数字图像或扫描像片数据；

⑤ 栅格数据值的重新定比（例如反差拉伸）。

——矢量或栅格数据叠加：

① 对矢量和栅格数据在下列范围中进行逻辑 AND、OR、XOR、NOT 叠加操作：多边形在多边形中，点在多边形中，点在线上，线在多边形中等；

② 在数据集间的叠加过程中能给某个数据类别的特征加权；

③ 在图形合成过程中自动或手工地合并属性信息。

——栅格数据处理：

① 指定二进制数据（1/0），离散数据（0～32768），连续实数数据存放到栅格数据集中；

② 对两个或多个栅格数据分类执行下列数学操作：加、减、乘、除、求最小值、求最大值等；

③ 对一个栅格数据分类执行下列数学操作：指数、对数、自然对数、绝对值、三角函数、反三角函数；

④ 以一个新的反映其相邻栅格某些数学综合值来取代原来数值的能力（例如平均值、最大值、最小值、合计值）；

⑤ 有监督或无监督聚类的能力。

——综合分析功能：

① 从点、线、多边形要素中指定距离缓冲区；

② 通过网络确定交替路径和最佳路径；

③ 自动标识排水网、集水区及视线遮蔽区；

④ 根据等高线数据进行挖、填方及断面分析；

⑤ 生成坡度、坡向及光线强度等数据分类；

⑥ 计算方位角、象限角及地形点位；

⑦ 具有定义、启动和关闭导线程序以及平差导线的功能。

——统计分析功能：

① 计算面积、周长、长度、体积；

② 对于同时出现在两种数据分类中的交叉分组列表，计算面积和平均值占总数的百分比；

③ 从表格数据中计算下列统计值：平均值、中值、四分位数、百分位数、中位数、标准差；

④ 对表格数据进行下列统计分析：相关分析、回归分析、方差分析、因素分析、判别分析、列联表；

⑤ 支持下列检验：T 检验、X^2检验、曼·惠特尼检验、串检验；

⑥ 支持下列分布：正态分布、泊松分布、二项式分布；

⑦ 计算置信区间和威尔科克森区间。

(5) 数据输出

——一般功能：

① 可用下列设备显示图形：图形终端、数字绘图仪、喷墨打印机、色带打印机、点阵打印机、激光打印机、静电绘图仪、字符打印机、胶片记录器等；

② 在栅格矢量显示设备上显示栅格或矢量的源文件；

③ 通过显示屏幕的拷贝生成地图；

④ 生成的地图可大于输出显示设备的物理尺寸，并可镶嵌；

⑤ 可生成网格表面的三维正射影像图和透视图，或其它具有高程数据的分类图；

⑥ 以交互方式或以缺省的地图格式组成显示的地图；

⑦ 可指定位置、大小、比例尺和视图的定向；

⑧ 可显示点、线和多边形数据集；

⑨ 以经纬度、国家基准面或 UTM 坐标参考系显示图廓线、网格线、晕线等并具有指定比例尺的注记；

⑩ 可从现存的表格中选择点状符号、线划类型、面积填充图案和字体。

——各种注记功能：

① 可以建立、命名、存储、检索及交互地定位下列内容：地图标题、图例、比例尺、南北箭头、单线或多线的文本字符串；

② 对所有文本记录指定符号类型、字符尺寸、颜色及字符串方向；

③ 在预先指定的点位（如多边形重心）自动地放置文字并补充以交互地移动或均匀配置注记的能力；

④ 可建立、命名、存储和选择缺省点位符号、线性类型、面积填充图案；

⑤ 可通过指定一地物名称或一组名称，指定显示颜色或颜色组，指定属性或属性组，以及用光标选择地物等方法来给图形特征设置点符号、线型、线宽、填充符号及颜色等；

⑥ 通过指定晕线颜色、线划类型、旋转角和距离间隔对面状区域进行填充。

应该说明，并非每一个地理信息系统软件都要一一具备上述各种性能。事实上对每一个特定用户而言，上述各种功能也并非都是必不可少的。因此，在实际工作中，往往是根

据上述各种功能从中选择出符合本单位部门要求的地理信息系统软件。

2. 系统总体功能评价

系统总体功能评价就是从技术和经济两个大的方面对所建立的系统进行评定。具体步骤可以对以下各项进行逐一审议和考核。

(1) 系统效率

地理信息系统的各种职能指标、技术指标和经济指标均是系统效率的反映。例如，系统能否及时地向用户提供有用信息？所提供信息的地理精度和几何精度如何？系统操作是否方便？系统出错如何？以及资源的使用效率如何？等等。

(2) 系统可靠性

所谓可靠性即指系统在运行时的稳定性，正常情况下应该很少发生事故，即便发生也能很快修复。可靠性还包括系统有关的数据文件和程序是否妥善保存，以及系统是否具有后备体系等。

(3) 可扩展性

任何系统的开发都是从简单到复杂的不断求精和完善的过程，特别是地理信息系统常常是从清查和汇集空间数据开始，然后逐步演化到从管理到决策的高级阶段。因此，一个系统建成后，要使在现行系统上不做大改动或不影响整个系统结构，就可在现行系统上增加功能模块，这就必须在系统设计时留有接口，否则，当数据量增加或功能增加，系统就要推倒重来，这就是一个没有生命力的系统。

(4) 可移植性

可移植性是评价地理信息系统的一项重要指标。一个有价值的地理信息系统软件和数据库，不仅在于它自身结构的合理，而且在于它对环境的适应能力，即它们不仅能在一台机器上使用，而且能在其它型号设备上使用。要做到这一点，系统必须按国家规范标准设计，包括数据表示、专业分类、编码标准、记录格式、控制基础等，都需要按照统一的规定，以保证软件和数据的匹配、交换和共享。

(5) 系统的效益

系统的效益包括经济效益和社会效益。目前地理信息系统还处于发展阶段，由它产生的经济效益不显著，可着重从社会效益上进行评价，例如信息共享的效果，数据采集和处理的自动化水平，地学综合分析能力，系统智能化技术的发展，系统决策的定量化和科学化，系统应用的模型化，系统解决新课题的能力，以及劳动强度的减轻，工作时间的缩短，技术智能的提高等等。从总的来看，地理信息系统的经济效益是在长时间内逐渐体现出来的，随着新课题的不断解决，经济效益也就不断提高。但是，从根本上说，只有当地理信息系统的建设走以市场为导向的产业化发展道路，商品经济的发展导致信息活动的激增，信息广泛而及时的交流，形成信息市场，出售地理信息系统产品和软件，才能为地理信息系统的发展提供契机，这时地理信息系统的经济效益才能真正体现，评价的目标也就自然地转向经济效益方面。

目前，一批以开发地理信息系统为目标的经济实体正在筹备和组建，地理信息系统的经济、科学和技术三统一的发展趋势是肯定无疑的。

第六章　GIS 工具软件及其应用

§6-1　ARC/INFO 系统概述

一、ARC/INFO 系统简介

1. 引言

ARC/INFO 是美国环境系统研究所（Environmental Systems Research Institute，缩写为 ESRI）研制的地理信息系统软件，也是世界上应用最广泛的 GIS 软件之一。ERSI 公司市场份额多年来一直雄踞 GIS 行业之首。据 GIS WORLD 杂志报道，ESRI 公司独占 1995 年世界 GIS 软件市场份额的三分之一，已连续几年稳居 GIS 市场占有率的第一位。1991 年，ESRI 公司对工作站 ARC/INFO 第 6 版进行了汉化，使其产品率先进入中国市场。

ARC/INFO 提供了用于地理数据的自动输入、处理、分析和显示的强大功能。目前已成功地应用到了包括自然资源管理、自动制图、设施管理、城市和区域规划、人口和商业管理、交通运输、石油和天然气、教育、军事等九大类别的一百多个领域。其主要特点为：

——采用地理关系数据模型；

——提供极强的空间操作和分析功能；

——采用模块式结构，使用灵活，易于扩充；

——提供宏命令语言和菜单方式，支持二次开发；

——提供 38 种地图投影方式，能够进行不同投影之间的坐标转换；

——具有存储和管理大数据量的能力；

——开放式结构，提供直接与多种数据库联接的接口；

——具有很好的兼容性，能与 25 种不同系统的数据格式相互转换；

——独立于硬件，支持最广泛的硬件平台；

——提供很方便的用户界面能力，用户能调用本系统的程序、系统外部的应用程序和操作系统的资源；

——广泛支持当今各种工业标准，包括操作系统、图形用户界面、商用数据库、网络协议标准、空间数据交换格式和图形输出格式。

2. ARC/INFO 软件的版本及硬件平台

ARC/INFO 的第一个产品建成于 1978 年，主要在小型机上运行。随着计算机硬件技术的飞速发展以及 GIS 技术的不断成熟，ARC/INFO 产品不断更新、扩大，形成适应各种用户、各种机型的系列产品。迄今为止，ESRI 公司已经推出的软件版本有：工作站和网络版的 ARC/INFO V7.1、工作站和微机版的 ARCVIEW V 3.0、微机版的 PC ARC/INFO V3.5 和 ARCCAD V11.4。到 1996 年底，ESRI 公司又把工作站版 ARC/INFO 的全部模块移植到 Windows NT 之上，使工作站环境下的 ARC/INFO 软件功能全部在微机上实现。

ARC/INFO 可运行在多种计算机和图形硬件上，这是 ARC/INFO 的一个重要特点，它能在如下的机器上操作：

——各种微机

——SUN 工作站及 SUN 多用户系列；

——DEC 工作站及 DEC 多用户系列；

——IBM 工作站及 IBM 多用户系列；

——HP 工作站及 HP 多用户系列；

——DG 工作站及 DG 多用户系列；

——PRIME 系列；

——INTERGRAPH 系列；

——APPLE 系列；

——TEKTRONIX 系列；

——NEC 系列。

ARC/INFO 既可与 TEKTRONIX 图形终端接口，也可以与仿真 TEKTRONIX4010、4014、4110、4200 等终端接口；此外，还支持 ENVISION 200(230)、ERDAS(400)、PRIME PT200 等多种终端接口。

ARC/INFO 支持的数字化仪包括：ALTEK，CALCOMP，IBM，KURTA 等厂家的各种型号数字化仪；支持的绘图仪包括：CALCOMP，NOVAJET，HP，IBM，ZETA，GRAPHICS 等厂家各种型号的笔式、静电式、热敏式、喷墨式绘图仪。

二、ARC/INFO 的基本模块与基本功能

ARC/INFO 采用模块设计的方法将整个软件按功能划分为多个逻辑子系统或称子模块。例如，所有的地图编辑功能组织在一个模块中，命名为 ARCEDIT；地图输出和绘图功能组织到另一个命名为 ARCPLOT 的程序中等。这样，ARC/INFO 就包含了许多子模块或程序包，而每一个模块都具有各自的一组命令和逻辑功能，本章将简要介绍工作站版 ARC/INFO 的基本模块与基本功能。

1、ARC/INFO 的基本模块

ARC/INFO 由描述地图特征和拓扑关系的 ARC 系统和记录属性数据的 INFO 数据库系统两部分组成。ARC 是 GIS 的总体管理模块，在 ARC 系统下又有若干个子模块组合，分别执行不同的功能。这些模块分别是：

——主模块：ARCEDIT、ARCPLOT、INFO、TABLES、LIBRARIAN

这是 ARC/INFO 软件的基础部分，用上述模块，即可进行图形及属性数据的输入、编辑、显示、查询、输出、管理等 GIS 的基本功能。

——扩展模块：NETWORK、GRID、TIN、COGO

扩展模块是满足用户特殊需求的工具，利用这些模块，可进行各种空间分析，如网络分析、三维表面分析、通视分析、最佳路径选择等等。

——新增模块：ARCSCAN、ARCSTROM、ARCPRESS、ARCEXPESS

在 ARC/INFO 第 7 版后，为了适应用户更高层次的需求，ESRI 公司又开发了这些新

增模块，可进行栅格矢量数据的转换，大型图库的管理，提高图形输出质量和速度。

此外，ARC/INFO 还提供称为 Arc Macro Language（简称 AML）的程序设计语言，供用户进行二次开发和功能扩充。

图 6-1 所示为 ARC/INFO 的基本模块结构。

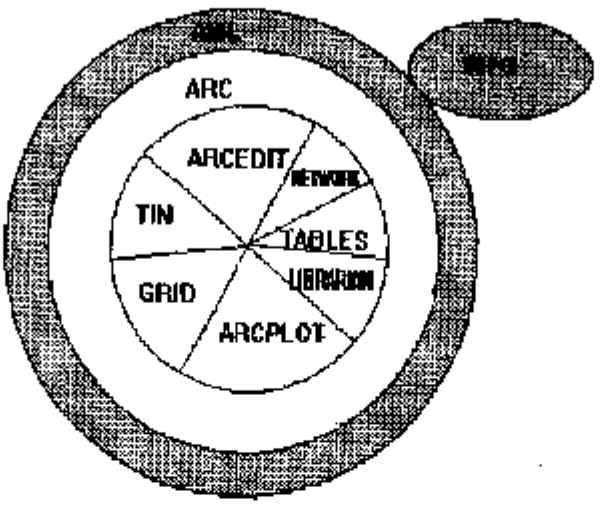


图 6-1 ARC/INFO 的基本模块结构

2. ARC/INFO 各模块的基本功能

下面简要列出 ARC/INFO 各模块的基本功能：

(1) ARC/INFO 主模块的基本功能

ARC：ARC 是 ARC/INFO 的主要程序环境，也是一个总体管理模块，其它子模块都要在 ARC 环境下进出。ARC 的主要功能有：

——对数据单元和工作空间进行操作和管理；

——进行数据格式转换和投影转换；

——建立拓扑关系及更新；

——进行空间数据分析；

——提出错误和检验错误；

——进行数据重组。

ARCEDIT：ARCEDIT 是一个交互式图形输入和编辑系统，提供了图形数字化的所有功能以及综合的图形和属性编辑功能。在 ARCEDIT 中，可以方便地输入点、线、面及文字注记等基本图形特征，可以对这些图形特征进行移动、拷贝、增加、删除、修改并提供影像和其它图层作为背景显示。在编辑图形特征的同时，ARCEDIT 还提供了对与特征相关的属性数据的编辑和维护功能，使用户不用退出 ARCEDIT 模块就能完成对属性数据的维护和更新工作。

ARCPLOT：ARCPLOT 是进行图形显示、查询和交互式制图输出的工具。在 ARCPLOT 中可以进行如下操作：

——显示图形和属性数据；

——查询图形和属性数据；

——建立符号库，设计线状、点状、面状符号；

——用地图制作器进行地图的设计、修改、制作；

——生成绘图文件并输出矢量数据地图；

——根据不同属性生成专题图；

——显示图像及栅格数据；

——绘制表面模型视图。

INFO 和 TABLES：INFO 是一个完整的关系数据库管理系统，它既是一个独立的子系统，完成对属性数据的管理和维护，又是 ARC/INFO 的一个有机组成部分，通过图形与属性的连接来管理与地理特征有关的属性数据。TABLES 是对 INFO 文件进行操作的表格管理系统，是 ARC 环境下的一个子模块，可以执行基本的 INFO 数据库操作，如建立、存贮、编辑和查询属性数据等。用 INFO 数据库可以进行如下操作：

——建立属性数据文件，包括定义表格、添加记录、显示和打印数据等；

——操纵属性数据文件，包括查询记录、更新记录、分类、统计、执行逻辑和算术运算等；

——建立报表和程序，包括生成报表、连接数据文件、用 INFO 编制程序等；

——管理 INFO 文件，包括改变报表和程序、维护文件记录、拷贝、删除文件等。

LIBRARIAN：LIBRARIAN 是对地理数据进行管理的模块，主要用于大数据量下的图形数据库的查询检索和管理，提供生成和管理图库的工具。可以将数据在水平方向划分为矩形或其它形状的片，称为 tile；在垂直方向划分为层，称为 layer。通过 tile 的索引结构可以加快数据的提取和查询检索速度。在对图库的管理上，LIBRARIAN 主要功能有：

——设置图库环境；

——管理图库结构；

——管理用户访问权限；

——建立图库索引；

——管理图库数据。

(2) ARC/INFO 扩展模块的基本功能

NETWORK：是进行线状网络和路径分析的工具。包括路径选择（Path—finding）和地址匹配（Address match），空间定位（Location），资源分区配置（Allocation）和动态分段（Dynamic Segmentation）等，为道路管理、公安及消防分析、城市设施（如地下管线、自来水、污水、电力、电讯等）管理提供了便利的工具。

GRID：是基于栅格数据的 GIS 分析、编辑、显示和处理模块，还是一个初级图像处理模块。它提供了将 Coverage 矢量数据转化为栅格数据的功能，具有图层及图层之间的地图代数运算、距离分析、三维表面工具、多变量统计分析以及路径代价计算等栅格分析功能。

TIN：是进行表面拟合与地形分析的工具。主要功能有：

——生成等高线，内插等高线数据；

——生成泰森多边形；

——执行体积分析、挖填方分析、流域分析；

——计算坡度、坡向、表面区域、表面长度；

——计算等高线间距；

——建立剖面图，生成纵断面图；

——指定两点间的通视情况；

——进行复杂情况下的可视分析。

(3) ARC/INFO 新增模块的基本功能

ARCSCAN：是进行扫描图预处理及矢量化的模块，能够实现栅格到矢量数据的自动或交互式跟踪、转换与编辑；可对扫描图像进行几何校正、空间配准及投影转换；可修改图像上的错误，如消除噪音、去除斑点等；可从图像中提取所需要的地理特征，建立 ARC/INFO 的 Coverage。

ARCSTORM：是空间数据库管理模块，是针对大中型 GIS 项目数据量巨大、用户众多、并发频繁的应用需求而设计的。它采用客户/服务器模式，使 ARC/INFO 的数据管理更为完善。ARCSTORM 的主要特点是：

——要素一级的事务管理；

——表格数据与空间数据的统一管理；

——网络上分布式数据的连续表示；

——支持数据库历史管理功能，可将数据库恢复到任意时刻以前的状态；

——提供系统恢复机制，当系统出现故障后，恢复机制能将数据库恢复到故障前的状态；

——客户机/服务器的体系结构。

ARCPRESS：是进行绘图和打印输出的工具。可将 ARC/INFO 的矢量地图文件或栅格和矢量数据并存的图形文件转换成纯净的栅格图像文件，加速栅格化输出过程，从而在不增加硬件投资的情况下完成高质量的、快速有效的绘图和打印输出；可产生支持多种绘图仪的影像文件。此外，ARCPRESS 还具有许多工具，包括比例尺控制、旋转、剪切、对比度控制和色彩校正等。

三、ARC/INFO 的数据结构及文件组织

1. ARC/INFO 的数据结构

ARC/INFO 描述两类地理数据：图形数据和属性数据。这些数据组织在一个地理数据库中。为了能有效地利用数据，不同类型的数据按照不同结构存储。属性数据存储在数据库中，图形数据按信息的类型和主题组织为一些层，每层称为一个 Coverage。

ARC/INFO 使用逻辑性较强的层状目录结构来组织数据。工作空间目录(Workspace)、Coverage 目录和 INFO 目录在 ARC/INFO 中尤为重要。

(1) 工作空间目录

在计算机中，操作系统以树状目录结构存储文件。信息按文件组织，而文件则按目录

组织。树形目录中，每个分支代表一个子目录或文件。ARC/INFO 即利用这一性能，指定一些专用的目录叫做用户工作空间。

用户工作空间是你在计算机上运行时的工作区域，它使用树状结构来组织你的工作。在 ARC/INFO 中，工作空间目录可以位于操作系统目录的任何位置，可通过指定路径名进入。一般的用户工作空间包含 coverage、grid、tin 等地理数据集和一个局部 INFO 数据库。工作空间包含的地理数据集作为工作区的子目录或文件存储。即：coverage、grid、tin 作为工作空间子目录存储，影像则作为文件存储。另外，工作空间还可能包含若干操作系统文件，如：AML 宏命令程序、菜单程序、绘图文件、投影文件等。

(2) Coverage 目录

Coverage 是 ARC/INFO 中的基本存储单元。它包含指定区域内地图要素的几何数据和属性数据，并可以实现它们的有效连接。通常，Coverage 只描述一种类型的地图要素，如道路、水系、居民地等，可以分为道路 Coverage、水系 Coverage、居民地 Coverage 等。有时，同一类地图要素又可以根据需要进一步细分，例如，水系要素可以分为面状水系（湖泊）和线状水系（河流）。总之，它可以根据使用者的需要灵活地划分。Coverage 是数字形式的图件，每个 Coverage 又包含多个地理数据文件，一些常见的 Coverage 文件有：AAT（弧段属性表）、PAT（多边形属性表）、ARF（弧段坐标和拓扑关系表）、BND（Coverage 最大最小坐标文件）、LAB（标识点坐标表）等等。

在层状目录结构中，Coverage 目录位于工作空间的下一层。在工作空间中可以包含一个或多个 Coverage，也可以没有 Coverage。它的多少取决于用户对图形数据的组织。

(3) INFO 目录

INFO 目录存贮与 Coverage 有关的属性数据，和 Coverage 目录一样，INFO 目录也位于工作空间的下一层；不同的是，每个工作空间只能有一个 INFO 目录。它包含了该用户工作空间内地理数据集的 INFO 文件，是工作空间的一个局部 INFO 数据库。

ARC/INFO 提供了一种机制，使位于 Coverage 目录下的数据可以被 INFO 访问。INFO 子目录中为每个 INFO 用户准备两个文件，一个是 ARCDR9，它是用户的目录文件；另一个是 ARCNSP，是默认的打印文件。此外，还有一系列的磁盘文件，它们是：

ARCnnnDAT——为相应 INFO 文件保存数据或记录外部文件路径名的磁盘文件。

ARCnnnNIT——包含为读取数据的数据项定义。nnn 表示顺序赋给内部 INFO 数据文件的一个三位数字（001～999）。

图 6-2 为某分省图数据库的目录结构，可以说法工作空间目录、Coverage 目录、INFO 目录的关系。

从图 6-2 可以看出，在这个名为 hainan 的用户工作空间中，包含了有关海南省地理要素的 coverage，它们是：水系（river）、道路（road）、建筑物（build）、境界（bound）及文字注记（anno）等。这些 coverage 均以子目录形式（目录下包含一组文件）存在于工作空间中，另外，工作空间还包含一个 INFO 目录（目录下也包含一组文件），它是该工作空间内的局部 INFO 数据库。此外，该工作空间中还包含一些操作系统文件，如用户编写的 AML 命令、菜单以及自动监视工作过程的日志文件 LOG 等等。

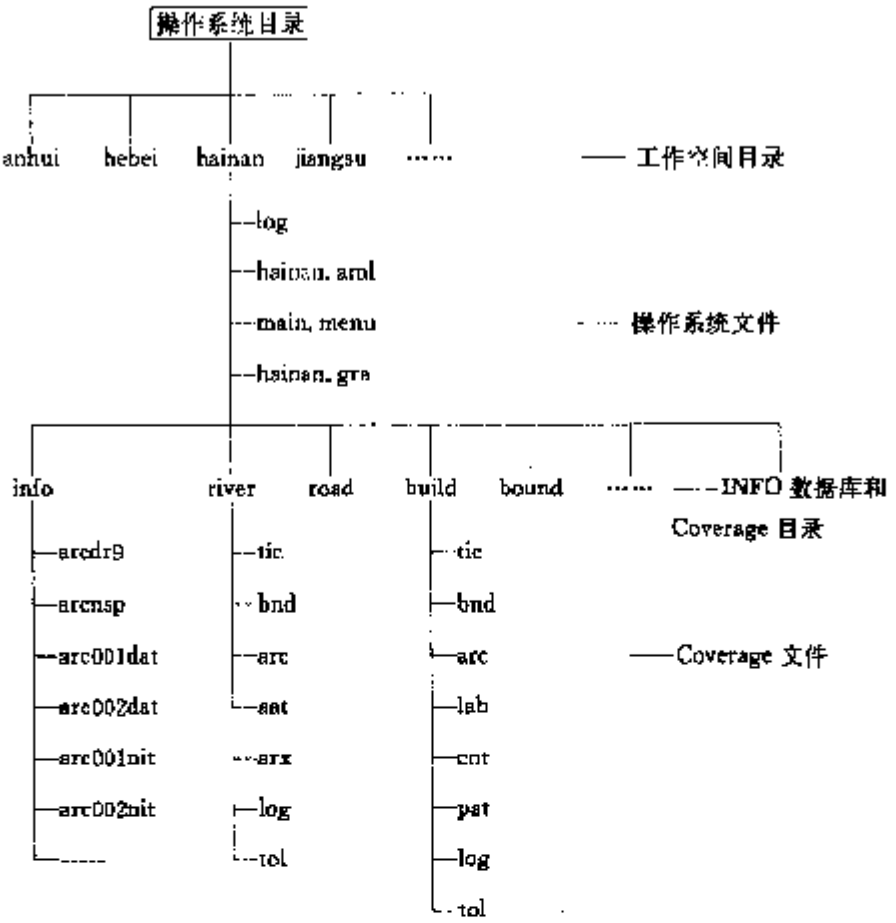


图 6-2　工作空间、Coverage、INFO 目录结构示意图

2. ARC/INFO 的文件组织

前面讲到工作空间目录和 Coverage 目录及 INFO 目录用于存放用户建立的图形数据和属性数据。而运行 ARC/INFO 软件所需的文件和程序则存储在专用的 ARCEXE 目录中。它是在安装软件时自动建立的。在这个目录下又有若干个子目录，分别存放运行 ARC/INFO 软件时所需的文件或程序。无论用户在哪个工作空间下进行操作，都会调用这些文件。表 6-1 列出部分子目录及其存放的文件。

表 6-1

目　　录	文　　件
ATOOL	存放 ESRI 编制的 AML 程序命令
DATABASE	存放支持数据库接口的文件
DIGFORM	存放数字化仪接口文件

续表

目　　录	文　　件
FONTS	存放文本字体文件
HELP	存放所有 ARC/INFO 联机帮助文件
INFOSYS	存放 INFO 数据库软件
LIB	存放 ARC/INFO 库文件
MENU	存放菜单文件
PLOTTERS	存放所有支持绘图仪的格式文件和绘图文件
PROGRAMS	存放 ARC/INFO 的可执行程序
SAMPLES	存放演示文件及其宏命令
STARTUP	存放 ARC、ARCEDIT、ARCPLOT 三个基本模块初始化宏命令
STATION	存放工作站初始化宏命令，也称站文件（station file）
SYMBOLS	存放标准符号设置文件及模板
SYSGEN	包含加密码，批处理及绘图队列信息
TABLES	存放 LIBRARIAN 所建立的图库名称及路径
TERMINALS	存放终端设备的格式预定义文件
TEMPLATE	存放 Coverage、INFO 和用户工作空间的模板
UTOOL	存放用户自定义工具文件

3. ARC/INFO 的数据模型

地理数据模型主要是用来反映现实世界的理论信息，它应该能让使用者对生成的数据进行编辑、显示、分析、维护和输出结果以及执行各种不同的应用。地理数据模型的设计应该做到简洁、容易理解和有充分的灵活性，用以管理大范围的不同的数据源。地理数据

模型还应该有健全的功能，能够涵盖各种地理特征，并且支持高级应用及能够产生灵活的应用界面。正是基于这样的设计思想，ARC/INFO 把地理信息提炼成最简单的点、线、面，每一空间要素能和相应的关系型数据库进行连接。自 6.0 版起又推出集栅格数据和关系型数据库为一体的 Grid 数据模型，所以，ARC/INFO 既是一个矢量的 GIS 软件，又是一个栅格的、具有图像处理功能的 GIS 软件。

最近几年，ARC/INFO 的数据模型有了非常大的发展，如增加了“区域”(Region)、“动态分段”(Dynamic Segmentation)，以适应各种高级空间数据及快速增加的数据量的要求。基本的图层可以组织成高层次的面向用户的数据组视图 (VIEWS)，用来帮助用户高效率地组织存取和使用地理信息。

四、ARC/INFO 新的技术发展

ESRI 公司在软件开发方面的投资是所有 GIS 软件厂商中最高的，其技术发展和产品更新速度也是最快的，继工作站 ARC/INFO 第 6 版后，已于 1994 年 11 月正式发行 ARC/INFO 第 7 版，目前工作站 ARC/INFO 最新版本为 7.1 版。新版本的 ARC/INFO，以其新的技术模型、新的数据自动化功能、客户/服务器体系结构、分布式处理技术和开放系统的战略体现了 GIS 及其相关领域的最新科技成果。新的技术发展主要体现在以下几方面：

1. Regions

Regions（区域）是 ARC/INFO 第 7 版中新的更有效的数据模型，它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，而不是具有相同属性的多个要素。如夏威夷州是由几个岛屿组成的，在数据库中只使用一条记录，而不是多条记录。这一新特征类代表了由一个或多个多边形组成的面状特征。Region 类型具有更加复杂的应用，特别用于模拟叠加区域，如野生动物活动范围以及时间区域上的历史变迁等。

2. 数据转换功能的改进

ESRI 公司以支持数据转换能力最强而著称。在此基础上，ARC/INFO 第 7 版又全面支持美国军方常用的 5 种 GIS 数据格式，包括 VPF（矢量数据格式）、SLF（标准线状要素格式）、DFAD（数据要素分析数据格式）、ADRG（先进数字光栅图形数据格式）、DTED（数字地形高程数据格式），同时支持 IMAGINE、SDTS、ADS、MOSS 等格式。

3. 图形输出方面的增强

增加了新的输出格式，如 VCGL2 和 RTL，RTL 支持 HP650C 喷墨绘图仪和 HP7600 静电绘图仪的图像混合输出，且大大提高了输出速度。笔式绘图仪绘制非单线体的文字注记很慢且效果较差，为此，ARC/INFO 第 7 版增加了勾勒文字外轮廓的功能，减小了绘图文件和绘图时间。

4. ARCEDIT 模块的改进

可在 ARCEDIT 下直接生成和维护多边形的拓扑关系；进行 REGION 的编辑；流方式 (Stream mode) 数字化；在 ARCEDIT 中使用 ARCPLOT 的命令；在更换编辑要素时，保持原来选择的要素；可进行图形的动态平滑漫游，而不必全屏刷新。

5. ARCPLOT 模块的改进

可自动绘制坐标格网；进行注记的自动布设（当注记过密时，能自动调整注记的位置，

使之不致重叠)；当注记背景过多过杂，注记看不清楚时，可为注记设置不透明的背景；可用复合的扇形或方形表示各要素的多种属性；可用点的分布密度表示某专题的多少；可在ARCPLOT中执行ARC的命令。

6. GRID模块的改进

GRID模块的改进很多，形成了很强的栅格GIS分析、编辑、显示系统；增加了很多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、交互式栅格数据编辑、NODATA的透明显示等。

7. NETWORK模块的改进

在NETWORK模块中，路线的建立和编辑更为方便，在编辑Route和Section时，不必频繁切换编辑要素；能自动检查路线属性的错误；提供多种网络选址/分配模型，包括：最短距离(Mindistance)，有限制的最短距离(Constrained Mindistance)，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。

8. 应用程序之间的通讯

ARC/INFO第7版可与本机或网络上其它节点的其它应用程序之间相互通讯，包括发出申请、执行请求、发送或接收数据、返回分析结果等。应用程序之间的通讯可在多种协议标准下实现，ARC/INFO既可作客户Client，又可作服务器Server。例如：用Arcview作为客户，为最终用户显示ARC/INFO分析的结果，ARC/INFO作为GIS分析的服务器，为Arcview提供后台支持。有了应用程序之间的通讯，ARC/INFO可以和操作系统的应用程序相互补充，和Arcview相互配合，形成功能完善、体系完整的GIS应用系统。

9. 提供菜单化的联机帮助

ARC/INFO第7版改进了联机帮助文件的界面，提供美观、友好、容易检索的菜单式联机帮助。这些帮助按功能分为不同的专题，并按字母顺序排序，使用户在遇到问题时不必离开计算机，就可以得到直接而迅速的技术支持。

10. 全面汉化

ARC/INFO第7版目前在绝大多数流行的UNIX工作站上实现了全面彻底的汉化，包括SUN、DEC、SGI、HP、IBM等。

11. ArcTools

ArcTools是ESRI的开发人员利用ARC宏命令语言（AML）编制的用户界面。它由一系列图标和按钮组成，它们可以用来进行数据生成、数据管理、显示、查询、地图建立、矢量数据处理以及栅格数据处理等。ArcTools作为一个现成的工具嵌入到ARC/INFO第7版中，使用户可以更快地完成常用的GIS分析工作；同时，对于那些经验丰富的用户，ArcTools也会给他们提供新的工具。ArcTools是利用面向对象的方法开发而成的，它使得这些工具可以作为用户进行二次开发的范例和基础。应用这些方法简化了应用开发，各个工具可以嵌入到其它应用程序中，而增加用户自己开发的工具也很简单方便。面向对象的方法用在ArcTools的开发中代表了GIS应用开发的一个突破。

§6-2 ARC/INFO 系统基本操作

ARC/INFO 软件是一个功能强大的GIS 软件，它的几个基本模块分别执行图形输入与编辑、图形查询与显示、空间分析、数据库操作、图形输出等基本操作，用户可以根据自己的应用需求选择不同的模块执行不同的操作。

一、图形输入与编辑

1. 图形输入操作

图形输入操作即是将各种地理信息数据转换为 ARC/INFO 的基本图形数据单元——Coverage。地理信息数据源的种类很多，如模拟地图、扫描数据、坐标文件或其它标准数据格式，对于不同的数据源，ARC/INFO 采用不同的方式进行处理。它们的最终结果都是以一系列(x，y)坐标值的数字形式获取空间数据，并存储于 ARC/INFO 的图形 Coverage 中。

(1) 手扶数字化

对于传统的模拟地图，最常用的方式就是用手扶数字化仪进行数字化。ARC/INFO 的 ARCEDIT 模块就可以进行手扶数字化工作。

利用 ARCEDIT 进行数字化，首先就要对图纸进行对点。对点的过程就是对图纸进行定位的过程。在每一幅模拟地图上都有其四角的控制点坐标，ARCEDIT 利用这些控制点坐标对图纸进行定位。在 ARC/INFO 中称这些控制点为配准点（TIC 点），如图 6-3 所示。

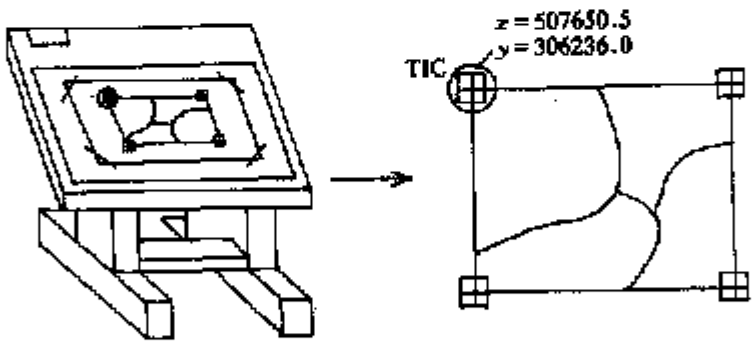


图 6-3 手扶数字化的对点

经过定位后的图纸，所有数字化的空间数据即被赋予真实世界坐标系的值，(通常是米或英尺)。在数字化板上即可用鼠标沿着图纸的线划方向进行手扶跟踪数字化。有时，为了工作上的方便，也可以先采用数字化仪的相对坐标进行配准定位和数字化（通常是厘米或英寸），然后再统一转换到真实世界坐标系中。

(2) 扫描数字化

对于地形较复杂的模拟地图，可采用扫描数字化方式。例如对等高线图采用扫描数字化比手扶数字化节省时间，提高效率。ARC/INFO 软件新增加的 ARCSCAN 模块就可以进行扫描数字化。

ARCSCAN 集成于 ARCEDIT 的环境中，利用 ARCSCAN 可以对扫描图像进行编辑和

矢量化。

(3) 坐标文件输入

除手扶数字化和扫描数字化外，也可以通过建立标准格式的数据文件进行数据输入。例如，在野外测量中直接记录了被测物体的 x、y 坐标，即可将这些坐标信息转换成标准格式的数据文件，利用 ARC/INFO 提供的数据转换命令将它们直接转换成 Coverage，或通过键盘直接输入坐标数据。

(4) 其它标准数据格式输入

ARC/INFO 可以接受多种通用的标准数据格式，在 ARC 主模块下通过格式转换命令，可将它们直接转换成 ARC/INFO 图形 Coverage。国内用户最常用的是 AutoCAD 的 DXF、DLG 文件，美国人口普查局的 TIGER 文件，此外还有其它 GIS 软件的数字交换格式。表 6-2 列出了 ARC/INFO 软件支持的几种常用的数据格式：

表 6-2

数据格式	内　容　描　述	ARC/INFO 转换命令
DXF	AutoCAD 文件和一些扫描产品的转换格式	DXFARC，ARCDXF
DLG	美国地质调查局的数字线划图交换标准	DLGARC，ARCDLG
IGDS	Intergraph 设计文件	IGDSARC，ARCIGDS
SIF	Intergraph 交换文件	SIFARC，ARCSIF
MOSS	美国联邦机构的 GIS 格式	MOSSARC，ARCMOSS
DIME	美国 1980 年人口普查街道图	DIMEARC，ARCDIME
TIGER	美国 1990 年人口普查街道图	TIGERARC，ARCTIGER

2. 图形编辑操作

无论采用何种方法将数据输入到 ARC/INFO 系统中，都不能直接进行应用分析，因为这些数据中还存在着错误或误差，例如：多边形未闭合，线段未连接，交叉的线段过头或不足，遗漏标识点或过多的标识点等，如图 6-4 所示。因此首先要对输入的图形进行编辑，修改错误，然后建立拓扑关系，才算完成了图形数据的输入和编辑。这个过程往往不止进行一次，它取决于原始数据的质量、数字化方法和改错的方法等。

图形编辑工作在 ARCEDIT 模块中进行，通常要进行以下步骤：

(1) 设置编辑环境

这步工作是设置要编辑的 Coverage 或 INFO 文件，在 ARCEDIT 中可随时设置当前编辑环境，以便于修改不同的图层。

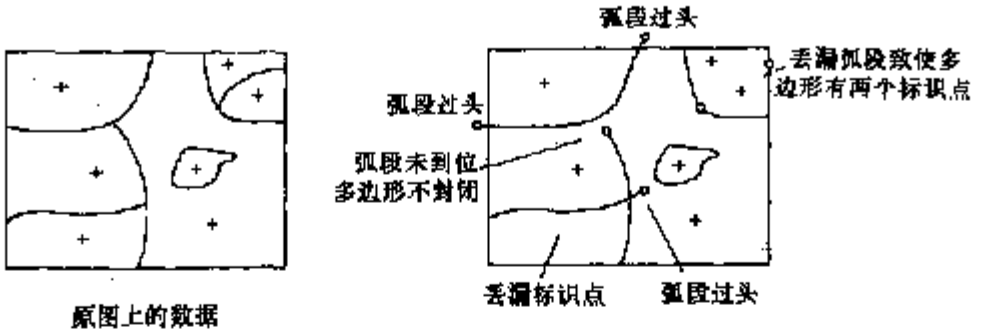


图 6-4 常见的数字化错误

(2) 指定特征类型

ARCEDIT 是基于特征 (feature) 进行编辑的，因此要指定当前编辑哪种特征类型，一次只能编辑一种特征类型。这些特征有：

配准点 (tic) ——是对图形进行定位的控制点；

弧段 (arc) ——是由一系列 (x, y) 坐标对组成的具有长度和方向的线段；

节点 (node) ——一条弧段的两个端点或多条弧段的交点称为节点；

标识点 (label) ——表示点要素，或用于给多边形指定用户标识；

注记 (annotation) ——是在图形中进行属性描述的一系列文字信息，如河流、道路的名称等。

(3) 选择特征

定义了特征类型后，要用 SELECT 命令选择要编辑的特征。SELECT 命令可以选择一个 (select one)、多个 (select many) 或一个区域 (select box) 来对选择的特征进行编辑。

(4) 进行编辑

进行了必要的设置后，即可对 Coverage 特征进行编辑，修改空间位置及其属性。ARCEDIT 提供以下基本的编辑命令：

DELETE——删除所选择的弧段或标识点

EXTEND——延长所选择的弧段

MOVE——移动所选择的弧段、节点或标识点

COPY——拷贝所选择的弧段或标识点

PUT——将所选择的弧段或标识点拷贝到另一个 Coverage 中

GET——从指定的 Coverage 拷贝所有特征到当前编辑的 Coverage

FLIP——改变所选择弧段的方向

SPLINE——对所选择弧段按指定的限差进行平滑或综合

ROTATE——对所选择的弧段进行旋转

RESHAPE——改变所选择弧段的形状

SPLIT，UNSPLIT——对所选择的弧段截开或合并

UPDATE——修改属性数据

CALCULATE——计算属性数据值

MOVEITEM——移动属性数据项

在 ARCEDIT 中，可以利用窗口系统的缩放（Pan/Zoom）功能，任意放大需要修改的区域，对地理特征的操作可在任一窗口或多个窗口同时进行，修改后所有窗口都更新。

二、图形显示与查询

查询操作的目的主要有两个：一是了解数据的性质和范围；二是选用查询的结果作进一步的空间分析。ARC 主模块和 INFO 数据库都提供一些查询命令，使用户可以在没有图形显示的情况下查询数据库。通常，更多的查询操作在 ARCPLOT 子模块中进行。在 ARCPLOT 中允许用户查询任何位置的图形和属性数据，同时显示查询结果。

你可以很快地在 ARCPLOT 中了解 Coverage 的有关内容，还可用如下命令查询满足一定条件的数据：

ITEMS——列出数据库中属性表的数据项

LIST——列出数据库中属性表的记录

RESELECT——从一个 Coverage 中选择满足一定条件的图形及属性数据

CLEARSELECT——清除所选择的结果

用以下命令你能很快地知道某一地物特征的性质并可在 ARCPLOT 中对数据库进行统计操作：

IDENTIFY——通过图形屏幕来选择或通过给定 x，y 坐标来列出 Coverage 中某一特征的属性

STATISTICS——从 INFO 文件的数据项中产生概括性的统计并把结果存到一个新的 INFO 文件中

还可以在屏幕上进行简单的坐标、长度、面积的量算：

MEASURE WHERE——测定地图上任意点的坐标

MEASURE LENGTH——测定两点或更多点之间的长度

MEASURE AREA——测定由三点或更多点定义的多边形面积

在 ARCPLOT 中，用户可以进行从图形查询属性或从属性查询图形的双向查询。从图形查询属性是选择一个或多个几何图形，并显示与之相对应的属性数据。例如：在屏幕上用鼠标点中几栋房屋的位置，立刻从属性数据库中查出这几栋房屋的面积、周长、房主、地址等属性，并同时显示在屏幕上。这一过程可以用如下命令来实现：

reselect build poly many——选择任意指定的楼房

polygonshades build 2——用红色填充选择结果

list build poly——列出所选楼房的属性数据

clearselect——清除选择

从属性查询图形是根据属性数据的关键字或某一限定条件，查找出相应的目标。例如要查找楼房层数在6层以上的楼房，可以用如下命令：

reselect build poly layer>6——选择层数在6层以上的楼房

polygonshades build 3——用绿色填充选择结果

list build poly——列出所选楼房的属性数据

clearselect——清除选择

根据上述表达式即可在屏幕上显示出所有 6 层以上的楼房。

三、空间分析操作

地理信息系统的空间分析功能使其区别于一般的数据库软件和图形软件。在我们今天面对的自然和环境问题中，空间分析是解决问题的关键。通过 GIS 的空间分析，我们可以用新的创造性的方法综合提取并显示地图数据，通过组合多个专题来找出它们之间的空间关系。

ARC/INFO 软件的扩展模块 GRID、TIN、NETWORK 可以执行多种复杂的空间分析操作，ARC 主模块可以进行常用的多边形叠加分析和缓冲区分析。在这里我们将这两种分析方法加以介绍。

1. 多边形叠加分析（Polygon overlay）

多边形叠加是指将两个层中重叠的多边形合并在一起，在一个输出层中建立新的多边形，多边形的属性也被合并在一起。为解决地理问题，我们分析与感兴趣区有关的专题数据。在 ARC/INFO 中，专题层存储为 Coverage，包含图形信息和属性信息。通过多边形叠加，不仅把专题层的图形叠加，而且把不同层的数据库也复合在一起，由此可以进行详细的分析。

ARC/INFO 提供三个命令执行多边形叠加：UNION、INTERSECT 和 IDENTITY，如图 6-5 所示。

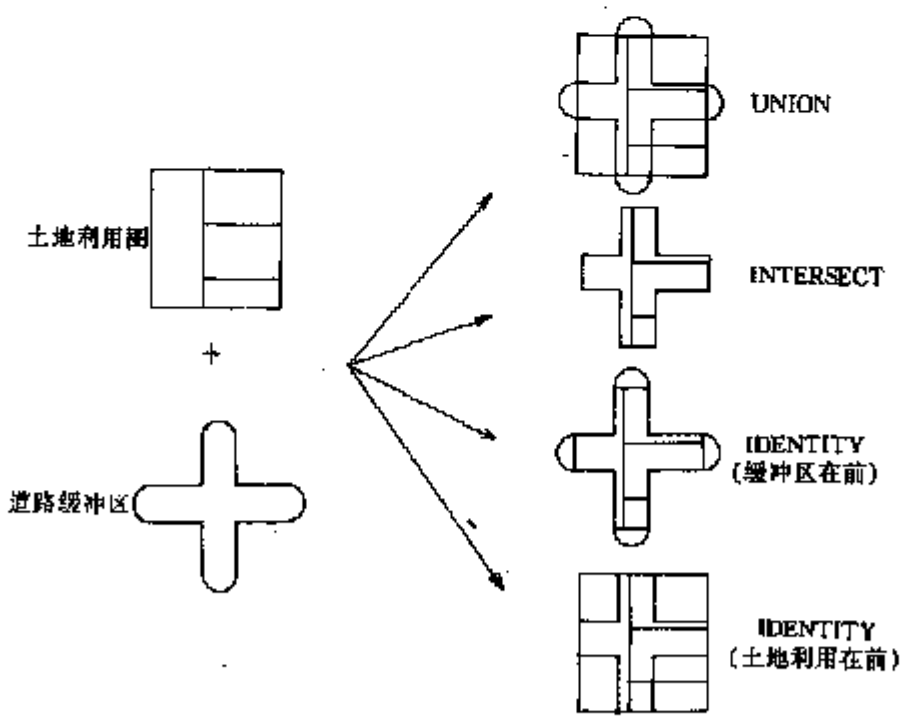


图 6-5　ARC/INFO 的三种拓扑叠加运算

——UNION：结合两个 Coverage 的所有特征；

——INTERSECT：把两个 Coverage 的公共重叠区部分的特征进行组合，而在公共区以外的数据将被删除；

——IDENTITY：对输入 Coverage 图幅范围之内的两个图层的特征进行叠加，而在输入 Coverage 图幅之外的区域将被删除。输入图层的顺序会使得输出结果发生变化。

2. 缓冲区分析（BUFFER）

空间操作中有一类非常重要的操作是关于确定不同地理特征的空间接近度或邻接性的操作，这种操作即为缓冲区分析，在 ARC/INFO 中用 ARC 下的 BUFFER 命令来进行。BUFFER 通过产生围绕一个特征（线、面、点或节点）的缓冲区来建立多边形 Coverage，如图 6-6 所示。BUFFER 操作总是产生一个多边形 Coverage，其属性表文件中有一项为 INSIDE，它的值有两个：100 表示落在缓冲区内的区域，1 表示落在缓冲区外的区域。缓冲区的大小可以由以下三种方式定义：

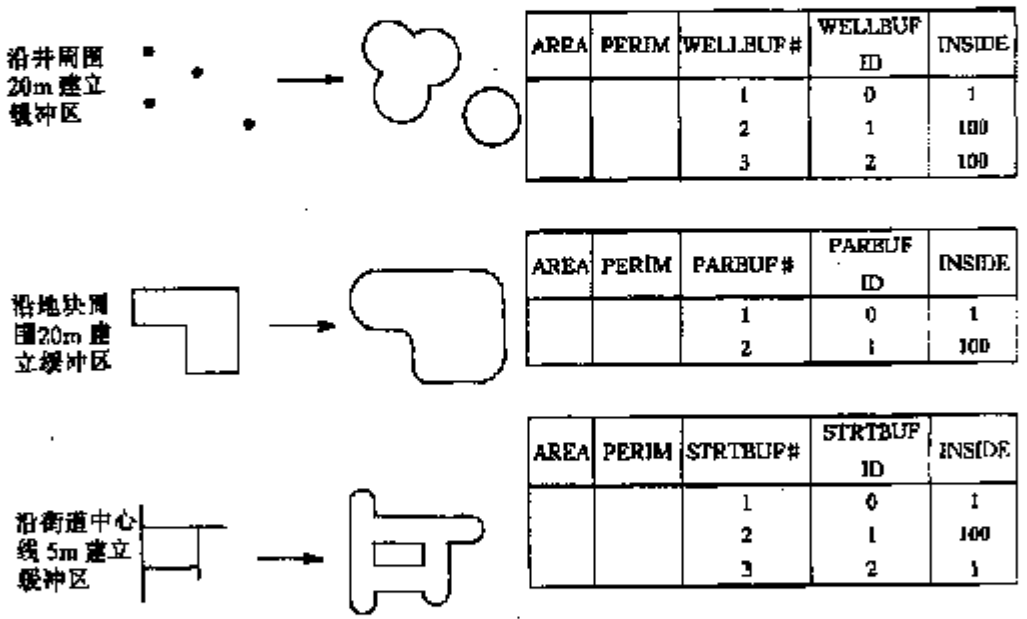


图 6-6 缓冲区分析

——通过指定缓冲区距离把所有缓冲区的大小定义为统一的大小；

——通过指定一个数据项来直接产生不同大小的缓冲区；

——通过查找表中的数据项值来产生不同大小的缓冲区。

利用多边形叠加与缓冲区分析可以进行多种应用。例如：有一街区内的道路要进行扩建，把原来 5 米宽的道路扩建为 15 米，道路两侧分别要加宽 5 米，必然要破坏道路两侧的居民住房或公共设施要拆迁，因此要计算出拆迁面积和需要赔偿的拆迁费用。这一应用的方法是：

(1) 沿道路中心线作一 15 米宽的缓冲区；

(2) 将街区图与缓冲区图进行叠加，所有在扩建范围内的房屋或公共设施用地都将落在缓冲区内；

(3) 落入缓冲区内的房屋产生新的多边形属性表并自动计算出房屋面积，根据面积和土地价格即可计算出拆迁的费用。

四、数据库操作

像操作空间数据一样，GIS 能够建立、存储、查询、显示属性数据。ARC/INFO 软件利用 INFO 数据库和 TABLES 表格管理系统进行数据库的操作。INFO 是一个专用数据库，它提供了数据库管理系统的功能；而 TABLES 是 INFO 功能的一个子集，仅提供数据库管理系统的基本功能。但是，用 TABLES 操作要比用 INFO 操作速度快，且对大小写也不像 INFO 要求得那样严格。因此，用户可以根据对数据库操作的内容来决定用 INFO 还是用 TABLES。表6-3列出了 INFO 与 TABLES 的相同和不同之处及其主要命令。

表 6-3

INFO	TABLES	主要命令
建立新文件	建立新文件	define
添加数据	添加数据	add、add from
显示数据	显示数据	list
提取数据	提取数据	select、aselect、nselect
操作数据	操作数据	calculate、move、update
删除数据	删除数据	purge
删除文件	删除文件	erase
连接文件	连接文件	relate、joinitem
报表打印		report
表格形式输入数据		forms
创建数据项		additem
重新定义数据项		redefine
有限制地改变项定义		alter

除了 INFO 和 TABLES 外，ARC/INFO 软件还提供了数据库接口，可以和几种关系型数据库连接，其中国内常用的有：ORACLE、INFOMIX、INGRES、SYBASE。

五、图形输出操作

地图是地理现象的图形表示，用于地理信息的交流。一个 GIS 应用项目，不但要在计算机屏幕上进行查询、显示、分析，还要将这些结果制成各种形式的专题图、地形图，通过图形输出设备输出，以作进一步的应用。ARCPLOT 模块提供了较强的图形输出功能，主要

包括以下两方面：

1. 符号设计

设计符号的目的是为了进行图形输出。在GIS中输入的基本数据是点、线、面，而输出时要成为一张完整的、符合用户要求的图形，就要套上各种不同的符号，填充各种不同的颜色等等。这些都需要符号设计。ARC/INFO软件把符号存在符号集文件中，一个文件可存放999个符号的定义。用户可以使用软件提供的内部符号集，也可以建立自己的符号把它们存在符号集中。ARCPLOT提供以下符号类型：

——SHADE符号用来填充多边形；

——LINE符号用来显示线或生成图廓；

——MARKER符号用来显示点、节点或任何一点（x，y）的位置；

——TEXT符号用来显示图上各种类型的注记或说明文字。

用户可以在ARCPLOT中利用符号编辑界面定义上述类型的符号库，建立、存储和修改符号。

2. 交互式地图制作

制作一张完整的地图，除了地图本身的内容外，还需要添加一些制图要素，用以帮助理解图中的信息。制图要素包括：标题、图例、指北针、比例尺、图廓、描述性文字等等。在ARCPLOT模块中，可以通过建立地图制作器（map composition）在屏幕上交互地制作地图。一个地图制作器是一个目录，在这个目录下可存储999个图形文件（或图素），用地图制作器可以多次任意地处理每个图素，可以增加、编辑、删除地图元素，可以将制图要素放在地图的任意位置，当对设计结果满意时，可从图形输出设备中输出硬拷贝地图。用地图制作器制作地图的常用命令有：

map——指定要生成的地图制作器的名字

mbegin——打开地图制作器

mselect——选择制图要素

mcopy——复制当前选中的制图要素

mmove——移动当前选中的制图要素

mfit——把当前选中的制图要素放入矩形框中放大或缩小

mscale——按指定的比例系数对选中的制图要素进行缩放

mrotate——对当前选中的制图要素进行旋转

mdelete——删除当前选中的制图要素

mposition——重新定位当前选中的制图要素

malign——将一组选中的制图要素沿着指定的水平或竖直轴对齐

mfresh——刷新屏幕，并重新绘制已生成的制图要素

mend——关闭地图制作器

交互式地图制作完成后，可生成绘图文件进行输出。ARC/INFO软件支持多种图形输出设备，如彩色喷墨绘图仪、静电绘图仪、笔式绘图仪、激光打印机等。

§6-3 ARC/INFO 系统应用

ARC/INFO 软件以其强大的查询显示和空间分析功能广泛应用于各个领域，成为一种有效的辅助决策支持工具。在美国及其它发达国家，ARC/INFO 软件的应用遍及环境保护、资源保护、灾害预测、投资评价、城市规划建设、政府管理等众多领域。近年来，随着我国经济建设的迅猛发展，加速了地理信息系统应用的进程，国内使用 ARC/INFO 软件的用户也在逐年增加。在城市规划管理、交通运输、测绘、环保、农业、制图等领域发挥了重要作用，取得了良好的经济效益和社会效益。本节将通过几个实例介绍 ARC/INFO 软件在不同领域中的应用。

一、ARC/INFO 在城市消防指挥系统中的应用

近年来，随着城市建设的不断发展，城市人口的不断增加，也导致了城市火灾呈上升的趋势。而消防队对火情的反应速度，也是衡量城市消防能力的重要指标。因此，充分利用现代化的技术手段，包括计算机技术、通信技术、地理信息系统技术，建立快速反应的自动化报警、消防系统，可以减少人工干预次数，提高消防抢险的效率，从而节省时间，有效地将损失减少到最小。为此，某市选用工作站版 ARC/INFO 软件和微机 ARCVIEW 软件，建立了该市消防指挥自动化系统，并在实际应用中取得了良好的社会效益和经济效益。

1. 城市消防指挥系统的构成

整个消防指挥系统由四大部分构成，即报警系统、地理信息系统、调度处理系统及网络传输系统，如图 6-7 所示。这四部分分别执行以下功能。

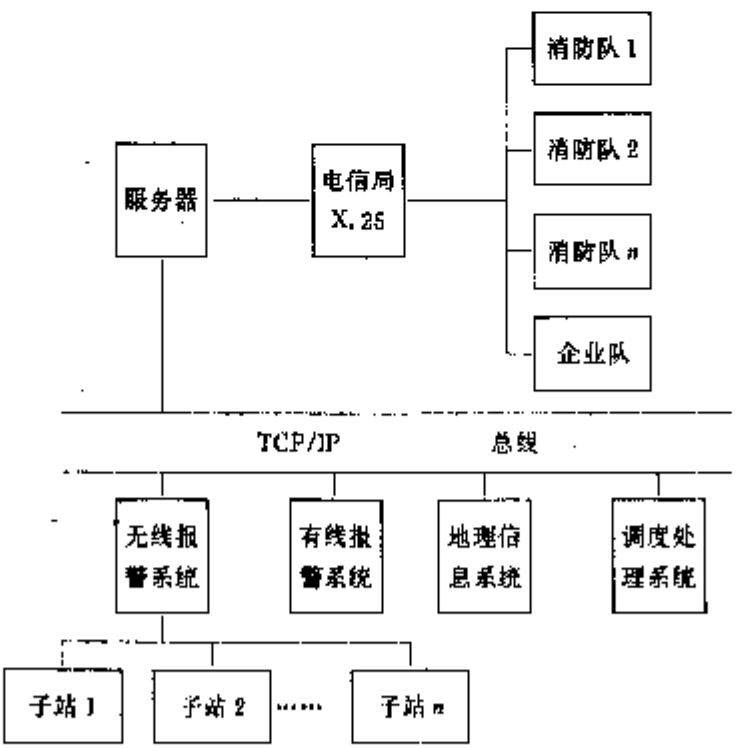


图5-7 城市消防指挥自动化系统构成

(1) 报警系统

这是火情信息的接收部分。它可以受理无线报警和有线报警火情。无线报警是通过安装在单位、重要建筑物内的子站自动报警；有线报警主要是受理通过 119 火警电话的火情报警。通过这两种方式报警的火情信息经确认和预处理后传输到地理信息系统。

(2) 地理信息系统。

这是消防指挥系统的关键部分。主要有两大功能。一是日常的维护管理功能，系统中存有与消防管理工作有关的城市街区图、交通图以及重点单位内部详细图等，在无火灾情况下进行数据的输入、显示、分析、输出等业务流程。二是对火情的分析功能，一旦发生火灾，系统即可以迅速判断火灾位置、着火单位信息、相邻单位信息、确定可能危险单位、周围水源，同时利用系统的空间分析功能确定各消防中队到火灾现场的最佳路径等等。这些分析成果通过网络传输到调度处理系统。

(3) 调度处理系统

调度处理系统的作用是运用地理信息系统的分析成果，综合各消防中队的实力和各相应单位的灭火方案，形成各中队的调度指令，通过电信网络下达到有关中队。

(4) 网络传输系统

网络传输系统是连接各个部分的纽带。包括两部分：一是连接报警系统、地理信息系统、调度处理系统之间的 TCP/IP 网络协议构成的局域通信网；一是通过电信局 X.25 网将调度指令传输到各消防中队。

2. 消防指挥地理信息系统的设计

在整个消防指挥系统中，地理信息系统的设计起着举足轻重的作用，整体系统的反应速度、动态控制的可视化程度均体现于地理信息系统的结构与功能设计上。

(1) 系统硬软件环境

系统以 SUN SPARCstation 20 工作站为主机，ARC/INFO 7.0 为软件开发环境，PC 机用于数据维护和 PC ARCVIEW 的运行，外围设备配有数字化仪、扫描仪、打印机、投影仪，通过连接 TCP/IP 协议下的网络软件进行数据的通讯。整个系统硬软件环境如图 6-8 所示。

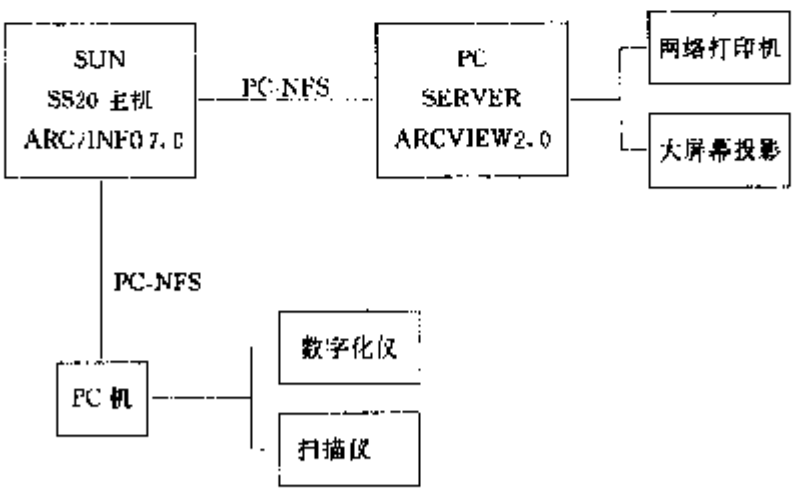


图6-8 系统软硬件环境示意图

(2) 数据结构设计

数据是整个系统的关键所在，要达到使系统简捷、实用、快速反应的目的，数据结构设计是至关重要的。在该系统中充分利用了 ARC/INFO 管理图形数据和属性数据的特点，确定数据层及所包含的空间及属性数据的内容。为增强可视化效果和便于分析，该系统以1:1万城市平面图为基础数据，根据用途分为两大类，即显示用图层和分析用图层。在这两大类图形数据中又连接相应的属性数据。

显示用图层包括以下几层：

——街区层：以街道网为基础构成的多边形封闭区域。属性按照街道内建筑物覆盖程度分为城区用地和郊区用地；

——铁路层：用于显示市区铁路的基本情况；

——水系层：该层既是显示图层，又具有分析意义。水系作为消防中重要的水源，要及时显示它的位置。在属性数据库中提供市区河流、湖泊、人工水池的水容量及动态变化情况，以及时补充消防水源；

——注记层：标识街道单位建筑物等的名称信息及特殊符号的定位显示。

分析用图层是指对分析中所需要的城市结构及设施等要素的分层，主要有：

——街区道路层：用于进行最佳路径分析，属性数据包括路况、路宽及动态信息，如路况分为主干道、一般道路、菜市场、规划路、土路等，动态信息则是根据城市进行的集会、市政建设，实时更新街道的畅通、阻塞情况；

——单位层，是用于消防或防火分析的基础层。其属性数据区分单位的性质，如一般事业单位、企业单位、商店、集贸市场、文化活动场所及住宅小区等；

——建筑层：主要指有一定高度的楼房层。属性数据中提供建筑物类型及楼层高度信息；

——市政消火栓位置层：反映全市消火栓分布情况，属性数据提供消火栓口径、水压及状态信息；

——消防单位及责任区层：显示消防单位在市区的分布及管辖的责任区，属性数据提供消防单位实力信息；

——无线电报警点位层：显示报警装置安装单位在市区的分布情况，属性数据提供这些单位的基本信息。

为配合日常的业务管理及消防抢险中的指挥分析，在系统中还存有重点单位平面分布图及有关照片，包括：单位内建筑物及道路分布图；单位水源状况、消防器材设施及关键点位分布图；特殊而重要的建筑物图片等。

(3) 系统功能设计

消防指挥系统的关键就是要快速、准确。当火情发生时，要迅速判断事故现场，立即组织消防救援。因此，在地理信息系统中提供数据的快速查询、显示、分析是非常重要的。ARC/INFO 软件提供了强大的查询、显示、分析功能，在消防指挥系统中，利用 ARC 宏命令语言（AML）将地理信息系统所要实现的功能编制成菜单式的界面，极大地提高了查询、显示、分析的速度。系统实现的主要功能有：

——显示查询功能，包括：

① 给定条件查询。如根据单位名称查询有关单位的各项属性信息；

② 屏幕点位查询。根据屏幕菜单提供的条件，利用鼠标在屏幕上定位，对所需图层信息进行查询；

③ 屏幕区域查询。根据屏幕菜单提供的功能，利用鼠标进行圆形区域、任一多边形区域的查询。

——定位分析功能，包括：

① 定位分析。通过 119 火警电话提供的信息，利用鼠标在屏幕上进行直接定位，例如，根据输入两条道路交叉口及方向、距离定位；根据输入某参考单位及距离、方位定位；根据对某特征物的距离、方位定位。

② 路径分析。确定各消防中队到着火地点的最佳行车路线。最佳行车路线根据系统提供的路况、路宽及动态信息，确定各路段的阻抗值（影响行车时间），计算行车路线的里程和时间。

——数据维护功能，主要是对图形和属性数据进行实时维护。图形和属性数据的正确与否，直接影响着查询和定位分析的准确性。因此要做到只要数据发生变化就及时进行更新，包括增加和删除多边形、弧段、点等图形数据，增加、删除属性数据项和记录，并要做到图形与属性数据修改的一致性。

——数据操作与输出功能，包括：

① 对当前屏幕实时进行漫游、缩放、旋转；

② 实时进行屏幕图形的截取打印和数据传输；

③ 对图形数据进行各层之间的组合输出，对属性数据进行列表打印输出。

3. 技术难点与实现方法

在系统实现过程中，重点解决了路径分析以及和其它系统接口等技术难点。所有这些技术难点都用 ARC/INFO 提供的宏命令语言（AML）编制程序加以解决。

(1) 路径分析中的关键技术

影响报警速度的关键因素是地理信息系统中的路径分析速度。为了提高路径分析速度，在图形显示上采用图像格式代替矢量地图，这样，既具有地图的效果，又大大提高了显示速度。在路径分析时，为了防止路径分析结果在地图上标出时发生和底图的错位，利用 ARC/INFO 中设置显示范围的命令 MAPEXTENT 对显示范围进行控制，以此保证两者的一致。在数据结构方面，利用 INFO 数据库管理着火单位的属性信息、拓扑信息及路径信息，在读取时，通过无线接警代码用 SQL 语言从 INFO 数据库中提取数据信息，可使得系统速度大大加快。

(2) 系统接口中的关键技术

① 无线报警系统的接口。消防指挥地理信息系统和无线报警系统、指挥调度系统是通过 TCP/IP 网络构成一个局域网系统的。但是，在 TCP/IP 网络协议层上 ARC/INFO 和无线报警系统不能同时共享同一个文件的读写权，为此，利用 AML 语言编制一个小程序来判断报警系统是否读写完毕，当判断已读写完毕时，即打开火警信息文件进行读取。

超星阅览器提醒您：
使用本复制品
请尊重相关知识产权！

② 调度系统的接口。路径分析的结果要用关系表的形式传到调度指挥系统中，但调度系统所用的数据库为 SYBASE，而地理信息系统中用 INFO 数据库，在文件的转换和传输中，利用 ARC/INFO 的 DBMSEXECUTE 命令，直接将地理信息系统分析的内容写入相应的 SYBASE 数据库中，同时，ARC/INFO 的数据库指针功能 DBMSCURSOR 又保证了数据的一致性。

二、ARC/INFO 在市政设施管理中的应用

市政设施包括城市的各种管线，如地下的给水、排水、污水、煤气等管线，地面上的还有各种通讯线路、照明线路等。这些管线错综复杂，尤其是地下管线，在地面上是看不见的，因此在城市建设施工中往往会出现破坏地下管线的情况。这不仅会带来极大的经济损失，像北京这样的国际化大都市，还会产生不良的政治影响。

为了解决上述问题，掌握有大量地下管线信息的北京市某测绘部门利用 ARC/INFO 软件建立了地下管线信息系统，从而实现了对地下管线信息的全面的现代化的管理。这对本单位的资料管理部门，实现了从落后的手工查询检索方式转为计算机管理及查询检索；对外，可为城市规划设计与管理部门、市政工程设计与管理部门、城市交通部门与道路建设部门等提供地下管线及其它测绘信息的查询服务。

该系统以城市大比例尺地形图为基础图形数据，在此基础上综合叠加地下及地面的八大类管线（包括上水、污水、电力、通讯、煤气、热力、燃气、工业管线）以及测量控制网、规划路等基础测绘信息，形成一个基于测绘数据的城市地下管线信息系统。图6-9为该系统总体功能框图。

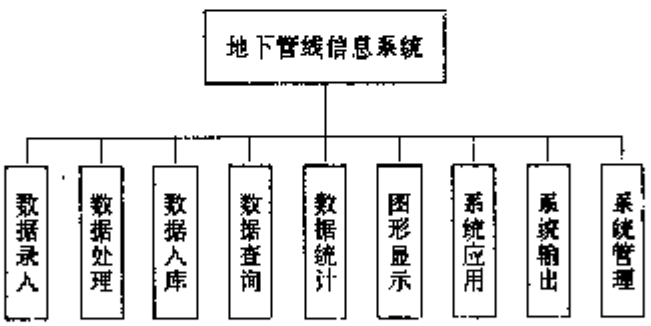


图6-9　地下管线信息系统总体功能

系统的主要功能有：

——数据输入：具有输入文字信息、线划地图，接收由航空像片测量的数据和野外测量数据的能力；

——数据管理：具有数据维护、各种图形与属性数据文件维护、数据库备份及安全性维护等功能；

——数据查询：具有从图形到属性或从属性到图形的信息查询能力；

——数据输出：具有按用户要求输出各种地形图、管线图和各种报表的功能；

——数据分析：具有空间/非空间数据分析的能力。利用该系统，可以将各种空间信息进行叠加处理，并进行许多有助于决策的分析。例如在道路扩建工程中，可利用本系统获得

超星阅览器提醒您：
使用本复制品
请尊重相关知识产权！

征地范围、面积及涉及单位等信息。

以下是该系统进行数据查询和输出的一些例子：

(1) 按指定区域查询地下管线信息

将地下管线与地形图套合在一起，在计算机屏幕上任意指定某一区域，就可以很清楚地查到该区域内每一条管线的分布情况。查询的结果可以显示在屏幕上，也可以输出到图纸上。

(2) 查询管线段或井位的属性信息

可以查询任意一段管线或该管线上各井位的属性信息，也可以查询整条管线或某个范围内几类管线的属性信息，如管线的位置、深度、管径、埋设时间等等。其结果既可以在屏幕上显示，也可以用表格形式输出。

(3) 绘制管线剖面图

只要在屏幕上任意指定某一位置，即可绘制出该位置的管线剖面图。

(4) 查询并输出地形图

利用该系统可以按区域、地名、图幅编号等多种条件来查询有关的地形图，查询的结果可以按用户要求进行全要素或部分要素的地形图输出，比例尺也可灵活地放大或缩小。

上述这些功能的实现主要用到 ARC/INFO 软件的 ARCEDIT、ARCPLOT 模块及动态分段技术，并利用 ARC/INFO 提供的 AML 语言有机地将这些功能结合在一起，形成灵活的易于检索的用户界面。其中主要的技术关键有：

——设置范围：通过图形或文字的逻辑表达确定所选专题的地理范围，并进行专题图形显示；

——信息查询：采用定性定位的检索方法查询地物信息；

——制图参数：对系统当前地图制作参数信息的描述或设置；

——地图制作：建立当前显示专题地物的 MAP 文件，及对此文件的各项操作如打开、存储、删除等；

——地图编辑：编辑地图要素及地图整饰；

——工具箱：为用户提供便利的服务工具。包括图形重绘、清屏、平移、缩小、开窗、漫游、点、线、面信息的量测，绘制统计图表和管线剖面图等；

——系统维护：对系统数据文件、数据库的安全性、日常管理与帐目登录等信息进行管理操作。

三、ARC/INFO 在公路管理中的应用

公路是一个连续的、网络形的实体，它与地理位置、地理环境密切相关。地理信息对公路的构造和形成有着决定性的影响。为了有效地掌握公路资源，全面了解公路布局，某省利用 ARC/INFO 软件建立了省级公路数据库。它利用 ARC/INFO 软件提供的 AML 语言开发成菜单式的用户界面，将与公路有关的信息以图形、图像或文本的方式形象、直观地显示在用户面前，达到图文并茂的效果；它可以对公路有关信息进行空间的查询、统计和分析，并可根据用户需要迅速绘制和生产出最新的交通图。图6-10是该公路数据库的系统功能结构图。

超星阅览器提醒您：
使用本复制品
请尊重相关知识产权！

公路数据库的特点是图形和属性数据量大，且公路情况是经常变化的，需要随时进行更新，因此对数据库的维护和管理是其主要内容和技术关键。在这里，涉及到 ARC/INFO 软件的关键技术有：

1. 利用 ARC/INFO 的图库管理功能（MAP LIBRARIAN）管理地图数据

该公路数据库选取1:25万地形图作为底图，分为城镇、道路、山峰、河流、行政区划五层地图要素。由于管辖区域广、地图要素数据量大，仅仅用 Coverage 进行管理是不行的，因此在对图形数据的管理中使用了 LIBRARIAN 模块，将经过接边处理后的各幅地图分块存放，根据用户需要，有关的各幅地图从图库中提取出来进行显示和查询。对用户而言，显示的地图就如同一个整体一样，通过空间位置和属性的连接关系保证了各幅地图数据的连续性和一致性。

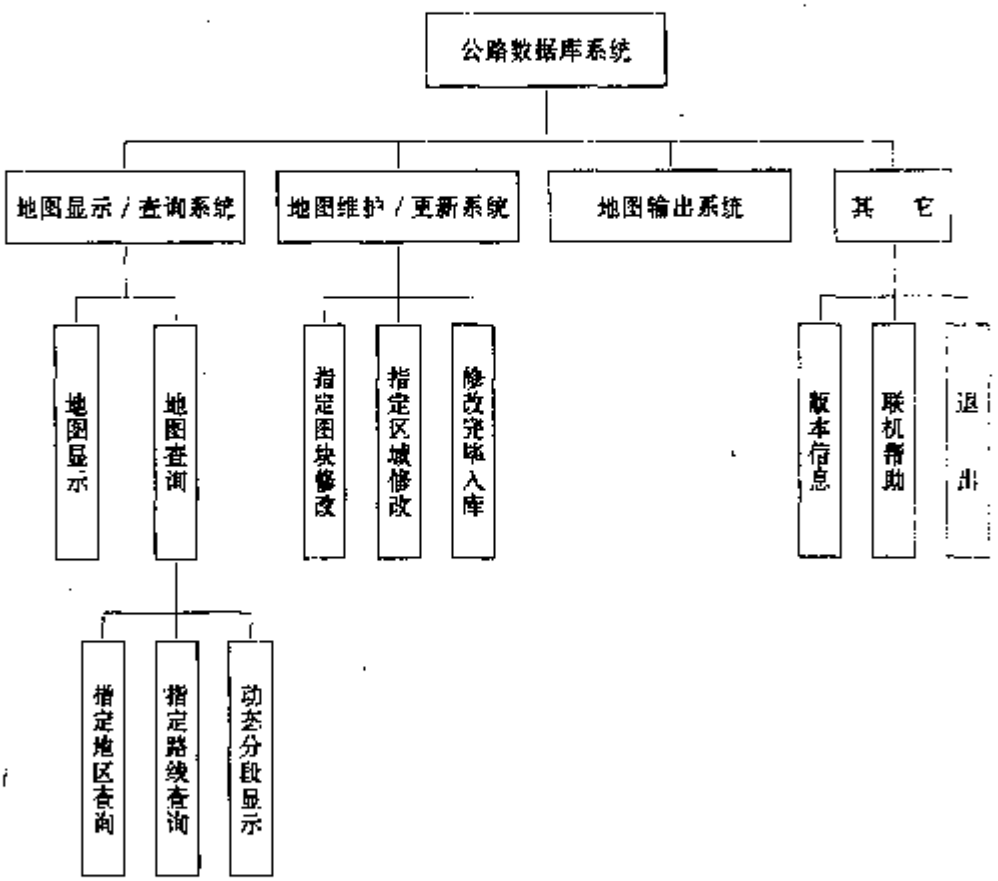


图6-10 公路数据库系统功能结构图

2. 利用 ARC/INFO 的动态分段技术管理公路属性数据

地理数据采用地理坐标（经纬度）或经投影后的直角坐标来表达其在空间的位置信息。然而，在我们日常公路管理中通常使用公路里程桩进行定位，一条公路同时具有多种属性，例如技术等级、路面等级、路面类型等，通过这些属性可将公路分为若干路段进行管理，分段位置便是使用里程桩来反映的。但是，公路上这些属性的分段位置一般是不一致的，因此，

超星阅览器提醒您：
使用本复制品
请尊重相关知识产权！

要在公路数据库中实现对公路各种属性的查询和分析，必须处理好整条路线与各段路线的连接关系及属性数据的相互匹配。ARC/INFO 软件提供的动态分段技术（Dynamic Segmentation）有效地解决了这些问题。

动态分段是 ARC/INFO 提供的一种数据模型，可将用户数据与 Coverage 中线性特征的任何一段相匹配。按照传统的方法，要存储不同的属性，必须用伪节点将线段断开。如果在一条线上有很多属性，则需要将一条线路截成很多段。这样，就造成了数据的冗余及查询检索上的不方便。而利用动态分段技术可以将多种属性与线性特征或其中的任何部分相联系，而不必采用增加节点将线段断开的方法来改变线性特征的属性。多种属性集可以连接到弧段或弧段的一部分上，以形成路径（route）。

在公路数据库中，利用动态分段技术的这一特性，根据公路的路线编号，对国道和省道建立名为 MILE 的路线系统，即 route。这是动态分段的基础。另外选取部分公路参照点（如公路的交叉路口、公路出入行政区域的边界点）建立名为 KNOWPOINT 的点状 Coverage，将这些参照点所在公路的路线编号和里程桩作为属性数据输入到该 Coverage 的点属性表中。利用这些属性数据建立起 KNOWPOINT 和 MILE 路线系统的连接和匹配关系，这样就可以将空间数据的分段和属性数据的分段独立开来，利用关系数据库中公路属性表的值，通过计算动态地得到空间数据的相应路段，以进行显示、查询和分析。

四、ARC/INFO 在地图制图中的应用

地理信息系统的发展是从地图制图开始的，因而 GIS 的主要功能之一是用于地图制图，建立地图数据库。传统的手工制图方式投入大，周期长，更新慢。而利用 GIS 建立起地图数据库，可以达到一次投入、多次产出的效果。它不仅可以为用户输出全要素地形图，而且可以根据用户需要分层输出各种专题图，如行政区划图、土地利用图、道路交通图等。ARC/INFO 软件的生产厂家 ESRI 公司与国内外许多数据采集部门、制图部门合作，用 ARC/INFO 软件开发了 1∶100 万世界数字航图库（Digital Chart of World）。该图库包含了世界各国的地貌、水系、交通、城市、行政区划、植被等基础地理要素，以及人口、气候、土地利用等社会、环境信息，用户只需购买该产品光盘，通过指定国名或区域即可在计算机中查阅、输出所需的图形及属性信息。

可以预言，地理信息系统的发展将使测绘、制图领域产生一场深刻的变革。传统的手工制图方式，输出的多是纸质的、平面的图形，而用 GIS 制图，图的内容形式及数据载体都发生了变化。由于 GIS 是一种空间信息系统，它所制作的图也能够反映一种空间关系，可以制成多种形式的立体图形，而制作立体图形的数据基础就是数字高程模型（Digital Elevation Model）。ARC/INFO 软件的 GRID 和 TIN 模块可以生成数字高程模型，在这里介绍几种利用数字高程模型制作的立体图。

1. 地貌晕渲图

在传统的制图方法中，为了增加丘陵和山区描述高差起伏的视觉效果，制图工作者发展了许多有关的制图技术，其中一种就是地貌晕渲法。用这种方法绘制的地图具有表现高差起伏的立体效果，但费用太高，且晕渲的质量和精度很大程度上取决于制图人员的主观意识。而利用 DEM 制作地貌晕渲图，能自动精确地产生晕渲效果。其方法是，首先根据 DEM

超星阅览器提醒您：
使用本复制品
请尊重相关知识产权！

数据计算坡度和坡向，然后将坡向数据与光源方向比较，一般将光源方向确定为西北45°，面向光源的斜坡得到浅色调灰值，背向光源的斜坡得到深色调灰值，界于中间坡向的斜坡得到中间灰值，灰值的大小则按坡度进一步确定。利用此方法制作的地貌晕渲图，立体感更强，精度更高，山区、丘陵、平原的高低起伏一目了然。国家测绘局所属的中国测绘科学研究院就利用ARC/INFO软件制成1:100万中国地貌晕渲图，在环境保护、水利工程等重大项目中得到了广泛的应用。

2. 地貌晕渲图与其它专题地图的叠加

晕渲图本身在描述地表三维状况中已经很有价值，在地形定量分析上的应用也不断扩大，如果把其它专题信息与晕渲图叠加组合在一起，其实用价值将会更高。例如，规划设计的运输线路图与晕渲图叠加后大大增强了直观感，从而使设计者可以根据地面的高差起伏情况来分析规划设计的正确与否。如果将各种彩色的专题图与地貌晕渲图叠加在一起，则产生的图形既有晕渲效果又反映专题内容。例如，用不同颜色表现行政界线的彩色政区图与地貌晕渲图叠加，可以清晰地看到每一政区内的地貌情况。在这方面，国家测绘局所属的国家基础地理信息中心已经有了成功的经验，他们将某省的基础地理底图数据包括水系、公路、铁路、居民地、行政区划以及文字注记等，在ARCPLOT模块中配以各种颜色和点状、线状符号，制成彩色政区图，并从1:100万数字高程模型中裁剪出同一地区的数据，形成栅格形式的数据文件，二者进行叠加套合，生成了具有地貌晕渲效果的彩色政区图。

3. 视线图

在地理信息系统的应用中，确定地面景观中点与点之间相互通视的能力具有很广泛的应用价值。它对于军事活动、各种电讯台站的规划及娱乐场所和风景旅游点的规划设计都是十分重要的。按照传统的等高线图来确定通视情况较为困难，因为在分析过程中必须提取大量的剖面数据并加以比较，而数字高程模型的建立为这类分析提供了极为方便的基础，例如，无线电通讯站的站址一般要选定在不受地物或高山阻挡的地方，要选择这样的站址，首先可以确定一处观察点，在DEM中辨认出观察点所在的位置，从这个位置引出所有的射线，比较射线通过的每个点的高程，将不被物体隐藏的各点进行特殊编码，从而得出一幅精确的视线图，可确定站址的最佳位置。

§6-4 其它GIS工具软件简介

一、国外其它GIS软件

目前世界上商用的GIS软件已多达400多种，它们大小不一，风格各异，占领着不同的应用市场。除ARC/INFO外，在国内外应用较多的还有GENAMAP、MGE、MAPINFO等几个大型的GIS软件。本节将对这三个软件作一简要介绍。

1. GENAMAP

GENAMAP是由澳大利亚GENASYS公司开发的GIS软件，该软件可以在工作站和微机上运行，操作系统均为UNIX和X—WINDOWS界面，且微机版本与工作站版本的源代码完全一致，易于移植。GENAMAP是产品的核心模块，子模块还有GENACELL和

超星阅览器提醒您：
使用本复制品
请尊重相关知识产权！

GENAREF。这三部分一起组成一套完整的地理信息系统。GENAMAP 主模块用于向量式(VECTOR)的地理信息数据的输入、管理、分析等，它所提供的功能大致可分为以下十个方面：

(1) 地图的定位校正；

(2) 图形数据的输入；

(3) 生成点、线图元；

(4) 属性数据的输入；

(5) 制作符号及图式；

(6) 向量式的分析功能；

(7) 属性数据库管理系统；

(8) 应用模式的建立和编辑；

(9) 图形显示功能；

(10) 系统管理功能。

GENACELL 模块是一套栅格式(RASTER)的地理信息系统，它与 GENAMAP 使用相同的空间数据库管理系统及相同的用户界面。但它的主要作用是将 GENAMAP 的向量式数据转换为栅格数据，提供栅格式地图分析和显示的能力。在 GENACELL 中，使用者可以指定格网的大小形状和颜色。所有 GENAMAP 的显示功能都能在 GENACELL 中使用。GENACELL 模块还具有基于栅格的空间分析功能，如数据的重新分类、叠加分析、距离分析、邻近分析等。

GENAMAP 产品除了使用自己的格式外，也可以接受其它软件的数据，或提供数据给其它软件使用。但这些数据首先要转换成标准格式，GENAREF 模块便提供这种功能，它可以将 DXF、DLG、IGES、SIF 等标准数据格式转换成 GENAMAP 的格式，以降低重新数字化所花费的时间与费用，提高系统的生成能力。

2. MGE (Modular GIS Environment，模块化的 GIS 环境)

MGE 是由美国 INTERGRAPH 公司开发的 GIS 软件。该软件最大的特点是具有很强的图形处理功能，因而更多地应用于地图制图出版领域。MGE 由三大部分构成：

(1) 面向矢量数据的软件

这部分主要是建立在数据库基础上的 MGE 家族各模块。它们是：

MGNUC (MGE Basic Nucleus) ——这是 MGE 的基本核心模块

MGAD (MGE Basic Administrator) ——这是 MGE 的基本管理模块

MGMAP (MGE Base Mapper) ——这是 MGE 的基本制图模块

用上述三个模块可实现基于数据库的 GIS 数据建立、管理、分析以及地图编辑工作流程。其它应用模块都建筑在它们之上。MGE 通过关系型数据库界面系统 RIS (Relational Interface System) 与各种商品数据库连接，用户只需对 RIS 进行操作，而无须进入数据库内操作数据。

(2) 操作栅格数据的软件

这部分主要用于栅格与矢量数据的相互转换。从矢量数据转换为栅格数据用 IPLOT

模块，从栅格数据转换为矢量数据用 I/GEOVEC 或 I/VECMS。

(3) 图形输出软件

MGE 软件最大的特点是图形处理功能强、作图时间短、精度高，可输出各种高质量的地形图、专题图、综合影像图等。Map Publisher 模块可以进行地图的分色分版。SRIF (Scanner/Film Recorder Interface) 模块可以输出分色软胶片。

最近，INTERGRAPH 公司又推出了新的 GIS 应用软件产品——Geomedia。它是一种通用的地理信息客户工具，可以使用户从各种不同的数据源中进行数据访问和集成。无论是 MGE 的数据、AUTOCAD 的数据或栅格形式的卫星影像数据，Geomedia 都可以将它们作为一个综合图件进行处理而不需任何转换。Geomedia 通过标准的 Windows 界面，提供简单的查询选项、专题绘图工具和图例制作功能，使初学者很容易掌握。由于 Geomedia 在 Windows 界面中，为用户创造了良好的二次开发环境，用户可以使用自己熟悉的编程语言如 Visual Basic、Visual C++、Excel Foxpro 等进行自己的功能设计。

3. MapInfo

MapInfo 软件是由美国 MapInfo 公司开发的 GIS 软件。它是一个面向中小型用户的桌面地理信息系统。因此其价格要大大低于 ARC/INFO 和 MGE。MapInfo 公司把 MapInfo 软件的市场定位总结为一个等式，即：MapInfo = Mapping + Information。正是基于这一思想，MapInfo 软件充分体现了小型、灵活、简单的特点。在用户界面上，MapInfo 利用 Windows 的功能，提供符号化的菜单和开发工具，使用户易于掌握。在数据库接口上，MapInfo 可以直接接收 DBASE 和 FOXBASE 的数据格式而无须中间加工，在图形上还可与 AutoCad、ARC/INFO、Intergraph MGE 的数据实现共享。在数据的可视化方面，MapInfo 具有很大的灵活性，每一张地图都可以用不同层次的图叠加而成，并且通过窗口的缩放功能观察整体与局部的细节。

除了一般的查询、显示、绘图功能外，MapInfo 软件也具备地理信息系统的空间分析功能，如叠加分析、缓冲区分析、数值及统计计算等。

MapInfo 软件的开发语言为 MapBasic，它将 MapInfo 的操作与标准的 Basic 语言融为一体，用户可以根据自己的要求，利用 MapBasic 开发适合自己使用的应用软件包。通过使用 MapBasic 进行二次开发，能够扩展 MapInfo 的功能，实现程序的自动重复操作并使 MapInfo 与其它应用软件集成，MapBasic 可以与用 Visual Basic、C++、Visual C++ 等语言编写的应用软件集成。

二、国内 GIS 软件

我国地理信息系统的发展开始于80年代初，其间经历了起步阶段、准备阶段、发展阶段，到目前已逐步向产业化的方向迈进。国产 GIS 软件更加注重于实用性，由于国内大多数单位都以微机为硬件平台，因此国产 GIS 软件也基本为微机版本，它们以价格低廉、使用方便同时又具备地理信息系统的基本功能而在一些中小城市、中小型应用项目上受到用户的欢迎。在此我们也选择三个有代表性的软件作一简要介绍。

1. 城市之星 (CITYSTAR)

CITYSTAR 由北京大学和陕西三秦企业集团联合研制。该软件的特点是将地理信息系

统（GIS）、遥感（RS）、全球卫星定位系统（GPS）结合在一起，具有把图形、声音、图像，数据、文本、模型量等进行综合处理显示的功能。根据不同的应用需求，CITYSTAR 分为以下几个模块：

（1）CITYSTAR-OFFICER——办公管理型地理信息系统

这是一套桌面 GIS 工具，该系统可建立矢量地图和对图像建立空间索引形成空间数据库，并在此基础上建立文档管理系统，将图像信息、多媒体、音像信息、地图信息、数据库及模型进行综合管理。

（2）CITYSTAR-PROFESSOR——资源与环境管理型地理信息系统

该系统基于 Windows95及 WindowsNT 将 GIS、GPS、RS 综合起来，具有矢量数据输入、编辑，空间分析及查询，遥感图像处理分类，遥感信息模型分析，网络分析及 GPS 数据与 GIS 地图叠加、量测，信息查询，多媒体信息复合，三维地形模型制作等功能。

（3）CITYSTAR-ENGINEER——城市管理型地理信息系统

该系统以矢量数据为主，与办公多媒体管理结合，直接为城市管理各部门服务，面向对象的数据结构及分层管理模式可以有效地满足用户对海量数据管理的需要。

（4）CITYSTAR-RS——遥感图像处理系统

该系统提供一整套遥感影像分析处理功能，包括：几何纠正配准、图像增强、图像漫游、多层图像运算、非监督分类与监督分类、GIS 矢量数据叠加统计及多媒体信息叠加检索等，并具有遥感分析结果、多元信息表达及办公信息查询等功能。

2. 吉奥之星（GeoStar）

GeoStar 是武汉测绘科技大学 GIS 研究中心研制的地理信息系统软件，其特点是采用面向对象技术，一类对象对应着相应的一系列操作，各类对象的联系便建立了吉奥之星的面向对象的空间数据模型。

GeoStar 是一个模块化集成环境，它将空间信息的录入、属性数据库定义、图形编辑、建立空间数据拓扑关系、地图符号化、地图输出、图像处理、空间查询分析等模块集成在同一环境中，并支持多媒体技术，能接收多种数据源，能提供多种空间数据输出格式，能满足不同用户的要求。

GeoStar 采用栅格——矢量一体化的数据结构及基于工程和工作区的空间数据组织方法，具有友好的中文用户界面、灵活的图幅接边与地图整饰能力。使用工程和工作区的概念，用户既可以根据需要切出任意大小的区域，又能保证较高的操作效率，因为实际操作并不是大的工程，而仍是一小块工作区。在一个工程内，根据地物的内容和特性划分为若干地物层，每层可能包含几种类型。将地物划分为类型和专题层，有助于快速查询、显示和输出。

3. MAP/GIS 地理信息系统

MAP/GIS 地理信息系统是由中国地质大学计算中心华地图形数据公司开发的地理信息系统软件。该系统利用先进的计算机图形图像处理技术及地学空间信息的处理方法，采用了矢量数据和栅格数据混合结构，将不同来源、不同类型的数据和相关的属性信息进行了有机的集合和综合分析与查询，并将结果以图形、报表等形式进行输出，绘制出供分析决策用的图件，实现计算机的信息处理，完成为管理人员提供辅助决策的手段。MAP/GIS 地

理信息系统主要由数字化子系统、图形编辑子系统、拓扑结构处理子系统、数字高程模型子系统、地图接边建库子系统、专业属性定义及管理子系统、数据库管理子系统、空间分析子系统、图像分析子系统、图形输出交换子系统等几大系统组成。各系统既互相独立、完成不同的功能，又共享图形信息和数据，实现综合查询和分析，从而成为一个完整的地理信息系统，既便于扩充，又便于维护。

MAP/GIS 地理信息系统在结构上采用了矢量数据和栅格数据混合结构，有效地解决了信息来源多种多样、数据类型各不相同的矛盾，实现数据信息的共享。

附录　国外八种主要GIS软件的基本情况与基本功能一览表

基本功能 \ 软件名称		ARC/INFO	GENAMAP	MGE	MAPINFO	SYSTEM9	SICAD	IDRISI	ERDAS
公司基本情况	所在国家	美国	澳大利亚	美国	美国	美国	德国	美国	美国
	公司创建时间	1969	1986	1969	1986	1986	1990	1987	1978
	公司职员	1700	145	10000	250	——	900	35	700
	当前版本	7.1	6.1	——	4.1	5.3	1.2	4.1	8.3
	软件发展时间	1982	1986	1987	1987	——	1979	1993	1992
	是否卖源码	NO	NO	NO	NO	YES	NO	NO	NO
操作系统	UNIX	√	√	√	√	√	√	×	√
	DOS	√	×	√	×	×	×	√	×
	Machintosh	×	×	×	√	×	×	×	×
	Windows NT	√	×	√	×	×	×	×	√
	OS2	×	×	×	×	×	×	×	×
	Windows	√	×	√	√	×	√	√	√
数据库管理系统	Access	×	√	×	√	√	√	×	×
	内部数据库	×	×	√	×	×	√	×	×
	DB2	√	×	√	×	×	√	×	×
	Dbase	√	√	√	√	×	√	√	×
	DS	×	×	×	×	×	×	×	×
	Foxbase	×	×	√	√	×	×	×	×
	OODBMS	×	×	√	×	×	×	×	×
	IMS	×	×	√	×	×	×	×	×
	INFO	√	×	√	×	×	×	×	√
	Infomix	√	√	√	×	√	√	×	×
	Ingres	√	√	√	×	√	√	×	×
	Oracle	√	√	√	√	√	√	×	×
	OS/DB2	×	×	×	×	×	×	×	×
	Paradox	×	×	√	×	×	×	×	×
	RBase	√	×	×	×	×	×	×	×
	SyBase	√	√	√	×	√	√	×	×

基本功能 \ 软件名称		ARC/INFO	GENAMAP	MGE	MAPINFO	SYSTEM9	SICAD	IDRISI	ERDAS
数据结构	矢量	√	√	√	√	√	√	√	√
	栅格	√	√	√	√	√	√	√	√
	三维	×	√	√	×	√	×	×	√
	TIN	√	√	√	×	√	×	×	√
	其它	×	√	√	×	×	×	×	×
数据集成	栅格—矢量	√	√	√	×	×	√	√	√
	矢量—栅格	√	×	√	√	×	√	√	√
	矢量与栅格叠加	√	×	√	√	√	√	√	√
数据获取	手扶跟踪数字化	√	×	√	√	√	√	√	√
	扫描数字化	√	×	√	√	√	√	√	√
	GPS 数据接收	√	×	√	√	×	×	√	√
	摄影测量	√	×	√	×	√	×	×	√
	鼠标	√	×	√	√	√	√	×	√
	野外测量	√	√	√	×	√	×	×	×
数据编辑	建立拓扑关系	√	√	√	√	√	√	√	√
	栅格数字化	√	√	√	√	×	√	√	√
	线生成	√	√	√	√	√	√	×	√
	接边	√	√	√	√	√	√	×	×
	拓扑误差检查	√	√	√	√	√	√	×	×
	属性域检查	√	√	√	√	√	√	×	×
	悬挂点检查	√	√	√	√	√	√	×	×
	节点匹配	√	√	√	√	√	√	√	√
数据显示	多幅图叠加	√	√	√	√	√	√	×	√
	阴影透视图	√	√	√	×	×	×	√	√
	网状透视图	√	√	√	×	×	√	√	√
	专题图叠加	√	√	√	√	×	×	√	√
	用户格网	√	√	√	√	×	×	√	√
	制图元素	√	√	√	√	√	×	√	√
	改变注记大小	√	√	√	×	√	√	×	√

基本功能 \ 软件名称		ARC/INFO	GENAMAP	MGE	MAPINFO	SYSTEM9	SICAD	IDRISI	ERDAS
数据显示	设置注记角度	√	√	√	×	√	√	×	√
	沿地物注记	√	√	√	×	√	√	×	√
	动画	√	×	√	×	×	×	√	√
	多媒体	×	√	√	×	√	√	×	√
几何测量	直线距离	√	√	√	√	√	√	√	√
	弧段距离	√	√	√	×	√	√	×	√
	面积	√	√	√	√	√	√	√	√
	频率	√	√	√	×	√	×	√	√
查询检索	由光标查询	√	√	√	√	√	√	√	√
	由键盘查询	√	√	√	√	√	√	√	√
	由属性查询	√	×	×	×	×	√	×	√
	支持 SQL 查询语言	√	√	√	×	×	√	×	√
多边形操作	多边形叠加	√	√	√	√	√	√	×	√
	点在多边形内	√	√	√	√	√	√	×	√
	线在多边形内	√	√	√	√	√	√	×	√
	通过属性合并与分割	√	√	√	√	√	√	×	√
缓冲区分析	点缓冲	√	√	√	√	√	√	√	√
	线缓冲	√	√	√	√	√	√	√	√
	多边形缓冲	√	√	√	√	√	√	√	√
	加权缓冲	√	√	√	√	√	×	√	√
网络分析	网络最短路径	√	√	√	×	√	√	×	√
	累计属性值	√	√	√	×	√	×	√	√
	邮路分配	√	√	√	√	√	×	×	×
	空间邻域搜索	√	√	√	√	√	√	×	√
	最近邻域搜索	√	√	√	√	√	√	×	√
	地址匹配	√	√	√	√	√	×	×	×
	动态分段	√	√	√	×	√	×	×	×

基本功能 \ 软件名称		ARC/INFO	ENAMAP	MGE	MAPINFO	SYSTEM8	SICAD	IDRISI	ERDAS
地图分析	重新分类与编码	√	√	√	×	√	√	√	√
	多层叠加	√	√	√	×	√	√	×	√
	平均格网值	√	√	√	√	√	×	√	√
	最大最小格网值	√	√	√	√	√	×	√	√
	逻辑组合	√	√	√	√	√	√	√	√
	加减地图	√	√	√	√	√	√	√	√
	乘除地图	√	√	√	√	√	×	√	√
	窗口内平均值	×	×	×	×	×	×	×	√
	窗口内最大最小值	×	×	×	×	×	×	×	√
	窗口内总值	×	×	×	×	×	×	×	√
	窗口内最频繁格网值	×	×	×	×	×	×	×	√
	聚块	√	√	√	√	√	×	√	√
	形态分析	√	×	√	×	√	×	×	√
地形分析	坡度	√	√	√	×	√	×	√	√
	任意点内插高程	√	√	√	×	√	×	√	√
	点通视分析	√	√	√	×	×	×	√	√
	线或面通视分析	√	√	√	×	×	×	√	√
	生成等高线	√	√	√	×	√	×	√	√
	用户定义断裂线	√	√	√	×	√	×	×	√
	计算最优路径	√	√	√	×	×	×	√	×
	生成断面	√	√	√	×	×	×	×	×
	充填方计算	√	√	√	×	×	×	×	×
混合分析	泰森多边形	√	√	√	×	√	×	√	×
	流域盆地	√	√	√	×	×	×	×	×
	制表报告	×	√	√	√	√	×	√	√
	输出概略统计	√	√	√	√	√	√	√	√
	集成数据描述信息	√	√	√	×	√	×	√	×

基本功能 \ 软件名称		ARC/INFO	GENAMAP	MGE	MAPINFO	SYSTEM9	SICAD	IDRISI	ERDAS
混合分析	产生随机采样	×	√	√	×	×	×	√	√
	近似分析	√	√	√	×	√	×	×	√
	加权近似分析	√	√	√	×	×	×	×	√
图像处理	数字正射影像	√	×	√	×	×	×	×	√
	辐射纠正	×	×	√	×	×	×	√	√
	传感器纠正	×	×	√	×	×	×	√	√
	合并数据	√	×	√	×	√	×	√	√
	几何纠正	√	√	√	√	√	√	√	√
图像增强	高低过滤波	√	×	√	×	√	×	√	√
	用户定义滤波	√	√	√	×	×	×	√	√
	反差拉伸	√	√	√	×	√	×	√	√
	颜色域变换	√	√	√	×	×	×	√	√
	密度分割	√	√	√	×	×	×	√	√
	直方图	√	√	√	×	√	×	√	√
	直方图均衡	√	×	√	×	×	×	√	√
	镶嵌	√	×	√	×	√	×	√	√
信息提取	主成份分析	×	×	√	×	×	×	√	√
	波段比值	√	×	√	×	×	×	√	√
	监督分类	√	×	√	×	×	×	√	√
	非监督分类	√	×	√	×	×	×	√	√
地理坐标	地理经纬度	√	√	√	√	√	×	√	√
	平面坐标	√	√	√	√	√	×	√	√
	UTM	√	√	√	√	√	√	√	√
	坐标变换	√	√	√	√	√	√	√	√
	投影变换	√	√	√	√	√	√	√	√
	地图投影	√	√	√	√	√	√	√	√
	用户定义	√	√	√	√	√	√	√	√
	其它	√	√	√	×	√	×	×	√

基本功能 \ 软件名称		ARC/INFO	GENAMAP	MGE	MAPINFO	SYSTEM9	SICAD	IDRISI	ERDAS
基本类型	自动制图	√	√	√	√	√	√	×	√
	数据采集	√	√	√	√	√	√	×	√
	格式转换	√	√	√	√	√	√	×	√
	文本管理	×	×	√	×	×	√	×	√
	桌面制图	√	×	√	√	×	√	×	√
	影像处理	√	×	√	×	√	√	√	√
	数据库管理	√	×	√	√	√	√	×	×
	遥感处理	×	×	√	×	×	×	√	√
	地理编码	√	√	×	×	×	√	×	×
用户界面	用户命令语言	√	√	√	√	×	√	√	√
	下拉式菜单	√	√	√	√	×	√	√	√
	用户定义菜单	√	√	√	√	√	√	×	√
	用户产生宏命令	√	√	√	√	√	√	√	√
	联机帮助	√	√	√	√	√	√	√	√
	命令提示	√	√	√	√	×	√	√	√
	多用户	√	√	√	√	√	√	√	√
	Windows 3.1	√	×	√	√	×	×	√	×
	X-Windows	√	√	√	×	√	√	√	√
	Windows-NT	√	×	√	×	√	×	×	√
	MOTIF	√	√	×	√	√	×	×	√
输出设置	笔式绘图仪	√	√	√	√	√	√	√	√
	静电绘图仪	√	√	√	√	√	√	×	√
	喷墨绘图仪	√	√	√	√	√	√	√	√
	摄像输出	√	√	√	√	√	×	×	√
	激光打印机	√	√	√	√	√	√	√	√
	点阵打印机	√	√	√	√	√	×	√	√
图形输出格式	打印到硬盘	√	√	√	√	√	√	×	√
	PS	√	√	√	√	√	√	√	√
	EPS	√	√	√	√	×	√	×	√
	PTCT	√	×	×	√	×	×	×	×
	HPGL	√	√	√	√	√	√	√	√
	TIFF	√	√	√	×	×	√	√	√

主要参考文献

[1] 黄杏元，汤勤. 地理信息系统概论. 北京：高等教育出版社，1989

[2] 邬伦，任伏虎等. 地理信息系统教程. 北京：北京大学出版社，1994

[3] 李德仁等. 地理信息系统导论. 北京：测绘出版社，1993

[4] Burrough, P. A.. Principles of Geographical Information Systems for Land Resources Assessment. Clarendon Press, 1986

[5] 王峥，任伏虎等. 微机 GIS 系统软件设计的原理与方法. 环境遥感，1989，4 (1)

[6] 郑若忠，王鸿武. 数据库原理与方法. 长沙：湖南科技出版社，1983

[7] Chrisman, N.R.. The Accuracy of Map Overlay: A Reassessment. Landscape and Urban Planning, 14, 1987

[8] Goodchild, M., Gopal, S.. Accuracy of Spatial Databases. Taylor and Francis, 1991

[9] 黄杏元，陈丙咸. 地理信息系统发展趋势. 地理学报，1989，44 (2)

[10] 史九林. 数据库系统初步. 北京：人民邮电出版社，1986

[11] 陈述彭. 地理信息系统的探索与试验. 地理科学，1983，3 (4)

[12] 邵进达. 地籍数据库的建立. 工程勘察，1996 (1)

[13] 张家庆等，九十年代 GIS 软件系统设计的思考. 测绘学报，1994，23 (2)

[14] 李伟生等. 适合共享的面向对象的地理数据模型. 武汉测绘科技大学学报，1996，21 (1)

[15] 朱光. GIS 中的误差与不确定性问题，武测科技，1994 (2)

[16] Mathew Kroguleeki. A Prototype Decision Guide and Audit Log for Preparation of Spatial Databases. Photogrammetric Engineering and Remote Sensing, 1990, 56 (4)

[17] 程朋根译. 三维地理信息系统的数据结构与应用问题. 武测译文，1995 (3)

[18] 李晓燕编译. 地理信息系统软件功能评定标准综述. 北京测绘，1991 (3)

[19] 汪正祥等. 城市 GIS：UIS 信息系统硬件配置设想. 北京测绘，1993 (3)

[20] 王大武. 多用途地籍信息系统的总体结构设计. 测绘通报，1994 (4)

[21] 程维译. 地理信息系统中的数据质量管理. 武测译文，1995 (2)

[22] 张仁霖等译. 一个面向对象的地理信息系统外壳. 武测译文，1995 (2)

[23] 邸方，苏亚芳等编译. ARC/INFO 地理信息系统初级教程. ARC/INFO 中国技术咨询与培训中心，1994

[24] 樊红编译. ARC/INFO 地理信息系统. 武汉测绘科技大学，1996

[25] 刘建中，田智慧等. 城市消防指挥地理信息系统的设计与实现. 中国 GIS 协会第二届年会论文集，1996

[26] 熊允泰. 北京市地下管线信息系统简介. ARC/INFO 中国通讯，1996 (2)

[27] 徐新，邓碧凤等. 地理信息系统在公路数据库中的应用，第二届 ARC/INFO 中国用户大会论文集，1996

[28] 龚健雅，夏宗国. 国外地理信息系统软件评述. 中国 GIS 协会第二届年会论文集，1996