

地理信息系统

地理信息系统是一门新兴科学技术，自从这门技术出现以来，其应用越来越广泛和深入。那么，地理信息系统的基本概念、功能及特点、组成是什么？数据结构、数据模型以及它与其他信息系统的集成如何开展？目前在电力系统的应用状况如何？这些问题是本章讨论的重点。因此，本章是全书的基本概念篇。

1.1 概 述

随着以信息高速公路为标志的信息时代和空间时代的到来，人们越来越意识到空间信息的重要性。信息高速公路（Information Super-Highway, ISH）又称国家信息基础设施（National Information Infrastructure, NII），是美国前副总统戈尔于 1994 年最先提出的，是由国家级、地区级和大企业级的数据库、声像设备以及覆盖全国全球的通信网络组成，并通过 Internet 连接全球的计算机通信网络，是一个能给全球广大客户随时提供大量信息及信息服务的计算机网络系统。作为信息高速公路的重要组成部分，地理信息系统（Geographical Information System, GIS）越来越成为人们解决各方面问题的有力工具，尤其是在空间信息方面。具有空间特征的地理信息系统是一门 20 世纪末迅猛发展起来的科学、技术和新兴产业，现已成为一个跨学科、多方向的研究领域，拥有广阔的发展前景。

1.1.1 相关概念

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，许多新的概念和术语也纷纷涌现。在理解地理信息系统的概念之前，了解一些与其相关的概念十分必要。

信息就是现实世界状态的反映，是用文字、数字、符号、语言等介质来表示事件、事物、现象等意义和内容，它不随载体的物理形式的改变而改变，具有客观性、实用性、可传输性和共享性等特征。而数据（Data）是信息的符号表示，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、图像等。信息用与物理介质有关的数据表达，而数据中包含的意义就是信息，信息和数据密不可分。

地理信息（Geographical Information）是指与所研究对象的空间地理分布有关的信息，是对表达地理特征与地理现象之间关系的地理数据的解释。它表示地上、地表、地下物体及环境固有的数量、质量、分布特征、联系和规律等。而地理数据（Geographical Data）则是各种地理特征和现象间关系的符号化表示，包括空间位置、属性特征和时间特征三部分，它们构成了地理空间分析的三大基本要素。

系统（System）是指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。信息系统（Information System）则指的是具有采集、存储、管理、分析和表达数据能力并且可以回答客户一系列问题的系统。计算机时代的信息系统都部分或全部由计算机系统支持，一般由计算机硬件、软件、数据和客户四个主要要素组成。如企业管理系统、图书情报信息系统、空间信息系统和其他一些信息系统等，其中，空间信息系统（Spatial Information System）是采集、管理、处理和更新空间信息的信息系统。

1.1.2 地理信息系统的概念

目前，由于受地理信息系统诞生历史不长、发展速度很快、应用领域广泛等因素的影响，在地理信息系统的定义问题上仍存在分歧，还没有形成一个普遍的共识。

地理信息系统在实际中为不同领域的客户所使用，因而不同客户对地理信息系统的理解都有所不同。制图界非常重视 GIS 快速生产高质量地图的能力，认为 GIS 是一种地图数据处理和显示系统，强调 GIS 作为信息载体和传播媒介的地图功能。计算机科学界认为 GIS 是一个完整的数据库管理系统，强调数据库系统在 GIS 中的重要地位。而普遍的 GIS 界则认为 GIS 是一门空间信息科学，强调 GIS 的空间分析和模型分析功能。总之，地理信息系统的外观表现为计算机软硬件系统，其内涵则是由计算机程序和地理数据组织而成的地理空间模型，一个逻辑缩小的、高度信息化的地理系统，从中可以提取地理系统各个不同侧面，不同层次的空间和时间特征，也可以将自然发生或思维规划的过程加在这个数据模型之上，取得对自然过程的分析和预测信息，用于管理和决策。所以，地理信息系统是一门集多学科为一体的新兴的交叉学科，是以整个地球或部分区域的资源、环境等空间数据为基础，在计算机软硬件支持下，采用地理模型分析方法，通过对地理空间数据的采集、存储、处理、检索、显示和综合分析，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技术系统。

地理信息系统的概念有三层含义：

（1）地理信息系统是一门新兴交叉学科。地理信息系统是集计算机科学、地理学、测绘学、遥感学、环境科学、城市科学、空间科学、信息科学、应用数学和管理科学为一体的新兴交叉学科。GIS 作为一门学科，主要研究以下内容：

- 1) GIS 的概念、定义和内涵；
- 2) GIS 信息处理方法；
- 3) GIS 理论系统；

的构成、功能、特点和任务；

- 6) GIS 发展历史；
- 5) GIS
- 4) GIS 发展方向；

7) GIS 技术系统设计和开发方法策略等。

(2) 地理信息系统是一项综合性的高新技术。地理信息系统是运用地理模型对地理空间数据进行综合分析、预测未来和模拟现实的一项技术。GIS 技术包括数据采集、存储、管理、空间分析方法、专题分析模型、系统集成技术和地理专家系统等。

(3) 地理信息系统是一个特殊的计算机信息系统。地理信息系统是在计算机软硬件支持下用于采集、处理、检索、模拟、分析和表达地理空间数据的计算机信息系统。它由计算机软硬件、地理数据、系统开发、管理和使用人员以及计算机网络组成，有数据采集和编辑、数据管理和组织、应用分析、结果显示输出和数据更新五大功能，能够由地理空间数据提供高层信息为管理和决策服务。

1.1.3 地理信息系统的发展

地理信息系统技术的发展是与地理空间信息的表示、处理、分析和应用手段的不断发展分不开的。国内外发现的较早的关于地理空间信息的表示可追溯到中国宋代的地图（地理图碑）和罗马时代的地图。到 18 世纪，欧洲文明的昌盛，才使人类实现了图纸地图，进而到 19 世纪出现了各种不同的地图和专题图。这些地图和专题图可谓模拟的地理信息系统。到 20 世纪中叶，随着计算机的诞生和发展，出现了计算机化的数字地图，从而进入了数字的地理信息系统时代。我们现在所称的地理信息系统通常指的是以数字地图（或电子地图）为基础的地理信息系统。

20 世纪 60 年代初，加拿大的测量专家 Roger F. Tomlinson 和美国的 Durane F. Marble 在不同地方从不同角度最先提出了地理信息系统这一术语。之后，世界上第一个 GIS——加拿大地理信息系统（CGIS）建立并于 1972 年全面投入运行和使用。从此，地理信息系统得到了迅速的发展和推广。

地理信息系统的发展是与计算机软硬件的发展紧密相连的。GIS 发展分为以下几个阶段：

1. 萌芽期（60 年代）

随着计算机技术的发展，特别是专家的兴趣以及政府的推动，地理信息系统得以较快的发展。这一时期的 GIS 主要是关于城市和土地利用的，其软件功能有限，注重于空间数据的地理学处理。同时，许多与 GIS 有关的组织和机构纷纷建立，例如，国际地理联合会（IGU）于 1968 年设立了地理数据收集和处理委员会（CGDSP）。

2. 巩固期（70 年代）

随着计算机软硬件技术的飞速发展和 GIS 专业化人才的不断增加，以及资源开发和环保问题引起的社会需求的增多，许多不同区域不同规模和主题的各具特色的地理信息系统得到了很大发展。这一时期的 GIS 的应用和开发多限于政府性、学术性机构，其软件的数据分析能力仍然很弱，注重于空间信息的管理。

3. 突破期（80年代）

由于计算机的性价比的提高和计算机网络的建立，GIS的应用领域迅速扩大，数据传输速率极大提高，功能也得到了较大的拓展，注重于空间决策支持分析。同时，许多政府性、学术性机构和软件制造商大量涌现，市场上也出现了许多商用化系统。

4. 拓展普及期（90年代）

随着地理信息产业的逐步建立和信息产品在全世界的普及，社会对地理信息系统的认识普遍提高，社会需求大幅增加。

我国的地理信息系统的发展是从80年代初研究资源与环境信息系统开始的，大致经历了以下几个阶段：

(1) 1980年以前为理论准备和人才培养阶段。GIS得到了启蒙研究。

(2) 1980~1985年为起步阶段。完成技术引进，规范研究的制定以及区域实验等。

(3) 1986~1990年为初步发展阶段。GIS研究和应用已经有组织、有计划、有明确的攻关目标，并逐步与国民经济建设和社会生活需求相结合，为区域管理、规划和辅助决策提供服务。

(4) 90年代为快速发展阶段。GIS实用化、集成化和工程化，国产软件得到迅速发展。

1.2 地理信息系统的功能及特点

1.2.1 地理信息系统的基本功能

地理信息系统处理地理信息的功能强大，贯穿数据采集—分析—应用的全过程。GIS的基本功能至少包括以下几个方面：

1. 数据采集和编辑

GIS要对多种形式（影像、图形和属性）、多种来源（野外测量、航测遥感和地图数字化等）的数据，实现多种方式（自动、半自动或人工）的数据输入，建立空间数据库。

数据采集是对系统外部的原始数据（多种来源、多种形式）进行必要的编码和写入数据库的操作过程。

数据采集的方式与所使用的设备密切相关，常用的几种方式如下：

(1) 数字化方式。使用手扶跟踪数字化仪采集有关图形的点、线、面的位置坐标；使用光栅扫描数字化仪采集图像的网格数据。

(2) 鼠标及键盘输入方式。用鼠标绘制各种图形，用键盘输入有关图像、图形的属性数据。

(3) 磁盘、光盘方式。主要是对现有的原始数据（电子地图、属性数据库文件）以文件方式用磁盘、光盘作为媒介进行输入。

(4) 对电子地图等也可以用网络传输输入。

GIS 数据源及其采集方式如图 1-1 所示。

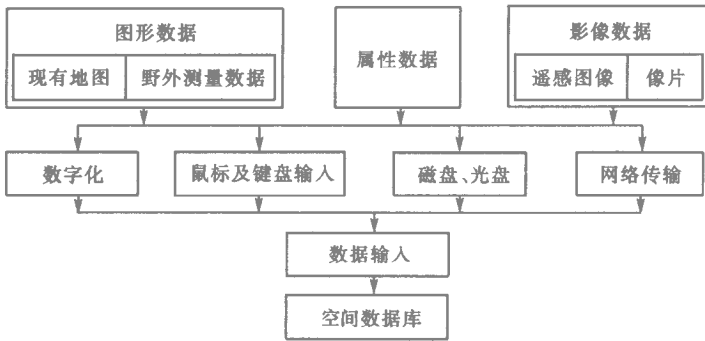


图 1-1 GIS 数据采集

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产生或引入一些差错，这就要求 GIS 对空间数据和属性数据应具备编辑功能以修正所出现的错误。通常，大多数 GIS 的数据编辑是比较耗时的交互式处理过程。对空间数据有图幅定向、文件管理、图形编辑（修正、增加、删除和更新）、生成拓扑关系、图形修饰与几何计算、图幅拼接等编辑功能；对属性数据则有修改、增加、删除和更新、数据库结构的修改以及与空间数据关联的编辑功能。无论是空间数据编辑，还是属性数据编辑，均需要建立简便易用、直观的对话框以便于人机交互。

2. 数据存储和组织管理

空间数据库的数据量大，空间数据与属性数据不可分离，而且数据应用面广。因而，须对数据库进行有效地管理，使数据冗余量小，数据与应用程序相对独立，选择合适的数据结构和数据模型。如图 1-2 所示，空间数据一般有栅格结构、矢量结构和栅格矢量混合结构三种组织方式；而对属性数据则有层次模型、网络模型和关系数据模型等进行描述和表达，其中关系数据库系统应用最为广泛。

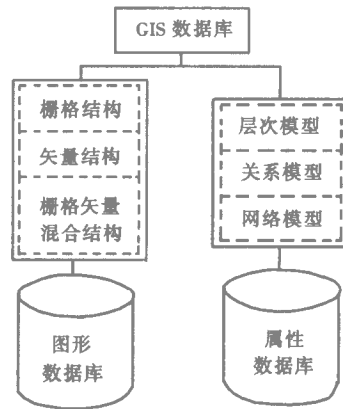


图 1-2 GIS 数据组织管理

由于空间数据往往关联诸多的属性数据地图，大量的属性数据通常采用关系型数据库与空间数据库分别存储的方法，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。

3. 应用分析

地理信息系统应用分析是在系统操作运算功能的支持下或建立专门软件来实现的，包括基本空间分析和模型分析，如图 1-3 所示。

GIS 最基本的分析功能有查询、检索、统计和计算功能，这些功能在其他自动化信息系统也具有。空间分析是 GIS 的核心功能，也是 GIS 与其他自动化信息系统的根本区

别，包括叠置分析、缓冲区分析、拓扑空间查询、空间集合分析等。而模型分析是指在GIS支持下，应用相应的数学模型分析和解决问题的方法，也是GIS应用深化的标志，如最佳网络分析、土地适应性分析和电网潮流分析等。

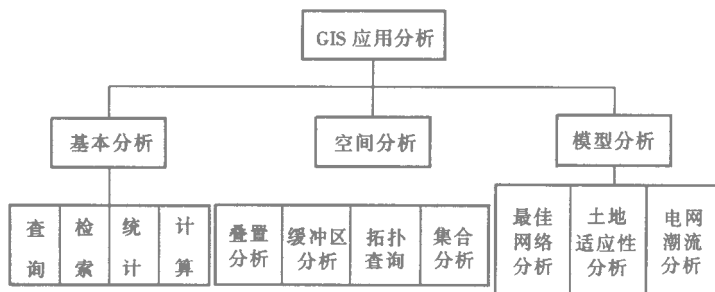


图 1-3 GIS 应用分析

4. 数据显示、结果输出

数据显示是中间处理过程和最终结果的屏幕显示，包括图形数据的数字化与编辑以及操作分析过程的显示；结果输出有专题地图、图表、表格和报告等各种类型的硬拷贝图形，其中屏幕显示也是结果输出的一种。GIS 应能提供一种良好的交互式的制图环境以人机对话方式来选择显示和输出，并能支持多种输出设备，如显示器、绘图仪、打印机、磁盘和光盘等，如图 1-4 所示。

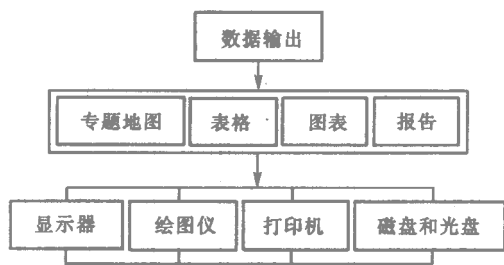


图 1-4 数据显示和结果输出

5. 数据更新

数据总是随着时间变化的，因而GIS应具有数据更新功能，才能较真实地反映现实情况，提供准确的决策依据，也可建立地理数据的时间序列，满足动态分析的要求。数据更新即通过删除、修改、再插入等一系列操作以新的数据或记录替换数据文件或数据序列中相对应的数据或记录。但现实中GIS的数据更新往往有时滞性，实时通信和3S（RS、GPS和

GIS）集成的研究正在努力缩小这种时滞性。

1.2.2 地理信息系统的特点

地理信息系统具备以下基本特点：

1. 统一的地理基础

GIS之所以区别于一般的信息系统就在于所处理的是地理信息。地理信息需要一个空间定位框架，即共同的地理坐标和平面坐标系统，才能更好地表达信息，从而支持空间问题的处理和决策。因此，GIS要建立统一的地理基础，包括统一的地理投影、统一的地理坐标系统和统一的地理编码系统。

2. 数据规范化

GIS 的数据来源多，形式多样，因而需将数据进行分级、分类、规范化和标准化，将其纳入一个特定投影和比例的参考坐标系统，使其适应于计算机的输入输出要求，便于进行社会经济和自然资源环境要素之间的对比和相互分析。目前许多 GIS 基础软件系统（如 Arc/Info、Mapinfo）都提供了多种常见的投影及其相互转换的功能。

3. 多维结构

在通常的二维数据结构中引入第三维（高程）和第四维（时间）以便为决策部门提供实时显示、多层次分析和动态分析等功能。

4. 空间分析

以地理模型分析方法为手段，采用各种空间关系运算进行空间分析和多要素综合分析，能够产生与这些要素相关的、综合的新信息，为决策提供服务。这是常规方法难以得到的。

5. 预测、模拟

在空间分析的基础上，采用数字和统计的方法，通过历史资料和数字模型的建立对事物进行定量分析，并对事物的未来作出判断和预测，尤其是对实时信息进行实时动态监测或预测。另外，还可以通过进行常规或非常规的数字模拟实验（空间过程演化模拟），为重大决策提供科学依据。

1.3 地理信息系统的组成

完整的 GIS 一般有五个主要部分组成，即 GIS 硬件系统、软件系统、地理数据、系统的组织管理人员和开发人员以及计算机网络。其中，硬软件系统是 GIS 的核心部分，可谓 GIS 的骨肉；地理数据库可以用来表达和组织各种地理数据，也十分重要，可谓 GIS 的血液；而 GIS 的管理人员、客户以及开发人员则决定系统的工作方式和信息表达方式；另外，计算机网络为实现数据共享、建立网络 GIS 搭起了桥梁。

1. 硬件系统

GIS 的硬件系统包括计算机主机、数据存储设备、数据输入输出设备以及通信传输设备等。

(1) 计算机主机。为 GIS 的核心，是数据和信息处理、加工和分析的设备。其主要部分由执行程序的中央处理器和主存储器构成，包括大型机、中型机、小型机、工作站和微机等。

数据存储设备。包括软盘、硬盘、磁带、光盘、存储网络等及其相应的驱动设备。

(2) 数据输入设备。除键盘、鼠标和通信端口外，还包括数字化仪、扫描仪、解析和数字摄影测量仪以及全站仪、GPS 接收机等其他测量仪器。

(4) 数据输出设备。主要有图形 / 图像显示器、矢量 / 栅格绘图仪、行式 / 点阵 / 喷墨 / 彩色喷墨打印机、激光印字机等设备。

(5) 通信传输设备。即在网络系统中用于数据传输和交换的光缆、电缆及附属设备。其中大多数硬件是计算机技术的通用设备，而有些设备则在 GIS 中得到了广泛应用，如数字化仪和扫描仪等。

2. 软件系统

GIS 软件系统是 GIS 的灵魂，由计算机系统软件、GIS 基础软件、GIS 二次开发软件和其他应用分析程序部分组成。

(1) 计算机系统软件。它是 GIS 日常工作所必需的，包括操作系统、系统库编程语言和库程序等，以及一些标准软件，如图形处理程序、数据库管理系统等。

(2) GIS 基础软件。能够提供给客户进行二次开发的 GIS 基础平台，为 GIS 核心软件，包括数据输入、数据处理、管理、结果显示输出和空间分析等部分。目前市场中主要有 Arc/Info、ArcGIS、MapInfo、MGE、Geomedia、Geostar 和 MapGIS 等商用 GIS 基础软件。

(3) GIS 二次开发软件。指针对不同客户，不同功能需求，不同管理和运作方式，基于 GIS 基础软件平台上的二次开发软件可为实现客户的特定要求提供开发环境（或语言），如 ArcView 的 Avenue 语言、Arc/Info 的 AML 语言、Mapinfo 的 MapBasic 语言等。许多 GIS 基础软件平台还支持用现代高级语言（如 Visual Basic、Visual C++、Java 和 Delphi 等）编程实现 GIS 应用功能。

(4) 其他应用分析程序。是系统开发人员或客户根据地理专题图与区域分析模型编制的用于某种特定应用任务的程序，是系统功能的扩充和延伸，一般可挂靠于原系统。

3. 地理数据

地理数据是 GIS 研究和作用的对象，是指以空间位置为存在和参照的自然、社会和人文经济景观数据，包括空间数据和属性数据，可以是图形、图像、文字、表格和数字等。空间数据表达了现实世界经过模型抽象后的实质性内容，即地理空间实体的位置、大小、形状、方向以及拓扑几何关系等；属性数据是与地理实体相关的地理变量和地理意义，是实体的属性描述数据。空间数据和属性数据密切相联，共同构成地理数据库，用于系统的分析、检索、表示和维护。地理数据库的建立和维护是一项非常复杂的工作，技术含量高，投入大，是 GIS 应用项目开展的关键内容之一。

4. 系统开发、管理和使用人员

仅有系统的软硬件和数据还不能构成完整的 GIS，需要人进行系统组织、管理、维护和数据更新、完善功能，并灵活采用地理分析模型提供多种信息，为研究和决策服务。同时还需要整个组织进行全盘规划，协调各部门内部的相关业务，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，同时周密规划 GIS 项目的方案及过程以保证项目的顺利实施。GIS 专业人员是 GIS 应用成功的关键，而强有力的组织则是系统运行的保障。

一个完整的 GIS 项目应包括项目负责人、系统分析设计人员、系统开发人员、系统维护人员、系统管理人员和客户等。其中项目负责人主要负责系统的规划设计、实施、资金预算以及协调项目相关方面的关系等；系统分析设计员负责系统需求分析、功能分析、需求开发的功能规划设计以及应用功能开发方案设计等；系统开发和维护人员负责应用软件

的编程开发和系统的维护工作；系统管理员负责系统硬件和其他外设的运行、物资管理、客户权限管理和请求支持等；而客户则为系统的最终使用者，包括决策层客户、管理层客户、技术层客户等。

5. 计算机网络

进入 90 年代以来，随着支持多客户网络操作系统的发展，以局域网 (Local Area Network, LAN) 和广域网 (Wide Area Network, LAN) 为主的计算机网络系统以及星地一体化的通信网络系统已经形成人类社会信息共享的有效体系。计算机网络利用通信线路将分布在不同地理位置上的具有独立功能的计算机系统或其他智能外设有机地连接起来，它包含下面三个主要的组成部分：

- (1) 若干台主机。用于向客户提供服务。
- (2) 通信子网。由一些专用的节点交换机和连接这些节点的通信链路组成。
- (3) 一系列协议。这些协议是为在主机之间或主机和子网之间的通信而用的。

计算机网络常见的拓扑结构 (连接方式) 有星形 (Star)、环状 (Ring)、总线 (Linear OR Bus) 和树形 (Tree) 等。

地理信息系统利用计算机网络技术可以实现空间数据的分布式存储和管理、网络资源的共享、重要数据的转移和备份。利用远程通信技术，还可实现跨国、跨大陆的 GIS 连网，获得更为广泛的共享资源和信息服务。

1.4 空间数据分类与特点



空间数据 (Spatial Data) 是指在 GIS 中用来表达地理空间信息的各种数据。狭义上，空间数据指的是图形实体数据，是相对于非空间的属性数据而言的；而广义上指的是 GIS 数据，包括图形、图像、属性以及相关的多媒体数据等，用于构建 GIS 数据库，作为 GIS 处理分析的对象。

1.4.1 狭义空间数据

狭义空间数据定义了图形和制图特征所处的空间位置，包括对象的空间位置以及空间拓扑关系等信息。在属性数据的支持下，空间数据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空间实体。通常，可用分类码和识别码来描述某一空间实体。“分类码”是对具有某一特征的实体子集的描述，而“识别码”是对单个实体的识别描述，在系统中必须是惟一的，其表现形式可以是数字的，也可以是文字的。此外，对空间实体的描述还有位置、实体特征、实体角色、行为或功能以及实体的空间特性等。

空间实体可以抽象为点、线、面、体以及组合类等几种基本地形要素。

1. 点状实体

点状实体为 0 维矢量，是一个数据点，用某种维度空间下的惟一的实数对坐标 [二维欧氏空间为 (X, Y) ，三维欧氏空间为 (X, Y, Z)] 来表示，逻辑上不能再分，至少有一个属性；在数学上，点状实体没有大小和方向。点状实体有点 (Point)、实体点 (Entity)

Point)、注记点 (Label Point)、节点 (Node)、角点(Vertex) 几种类型。城市道路交叉点、电力系统的杆塔、开关、变压器等地理实体均为点状实体。

2. 线状实体

线状实体为一维矢量，是一个线状要素或空间对象的边界，用一组离散化实数点对序列 [二维欧氏空间为 $(X_1, Y_1) (X_2, Y_2) \dots (X_n, Y_n)$ ，三维欧氏空间为 $(X_1, Y_1, Z_1) (X_2, Y_2, Z_2) \dots (X_n, Y_n, Z_n)$] 来表示。线的形状决定于坐标对序列的排列顺序，线上每个点有相同的公共属性并至少存在一个属性。线状实体具有方向（起始于起始节点，结束于终止节点），还具有长度、弯曲度等特征。线状实体包括线段、边界、链、弧段和网络等。河流、道路、地形线、区域边界和地质界线等均为线状实体。

3. 面状实体

面状实体为二维矢量，是一组闭合弧段所组成的空间区域，用一组首尾节点相同的离散化实数点对序列 $(X_1, Y_1) (X_2, Y_2) \dots (X_n, Y_n) (X_1, Y_1)$ 来表示。凡是面的内部点可以多于三个邻点，面内每个点至少具有一个相同属性。面状实体有面积、周长、凹凸性、独立性、包含性、相邻性和重叠性等空间特性。面状实体又称为多边形，如湖泊、岛屿、植被和行政区划等地理实体均为面状实体。

4. 体状实体

体状实体为三维矢量，是一组或多组闭合曲面所包围的空间对象，具有长度、宽度及高度等属性，还具有体积、周长、断面图和剖面图等空间特性，如建筑物、三维工业物件等。

5. 组合类实体

组合类实体是点、线、面、体四个基本地形要素的某种组合，以表示现实世界的各种复杂的自然实体或关系。如不规则楼房、含岛的湖泊、电力线的交叉跨越等。

1.4.2 广义空间数据

广义空间数据，又称 GIS 数据或地理数据，主要有以下几种类型。

几何图形数据。即前述的狭义空间数据，它来源于各种类型的现有地图（野外测量地图、手工绘制的草图和其他图件）和实测几何数据。它不仅反映空间实体的地理位置，还反映实体间的空间关系（拓扑关系）。

影像数据。影像数据主要来源于航空和航天像片、航空和航天数字图像或其他数字形⁽³⁾₍₂₎为遥感影像等。它能较快地、直观地、形象地表达地理现状。

(1) 属性数据。属性数据主要来源于实测数据、几何图形数据的属性描述信息、调查统计数据（包括社会经济数据）、图例符号以及遥感解译信息等。其空间特征更加隐含（往往与图形数据密不可分），是空间数据的重要组成部分。

(4) 地形数据。地形数据主要来源于地形等高线图的数字化、已建立的格网状数字高程模型（DEM）或三角网地面模型（TIN）等。地形数据能有效地辅助几何图形数据来表达地理现状，如地形的高低起伏变化情况。

(5) 多媒体数据。多媒体数据主要来源于与地形实体相关的声音、图片、录像和文本

等形式的各种数据，是图形数据的一种补充，可用来丰富和完善地理实体的表达形式。

(6) 其他，如专家知识数据等。

1.4.3 空间数据的特点

空间数据主要具有以下特点：

(1) 空间特征。空间数据的空间特征是指地理实体或现象的空间位置及其相互关系，如大小、形状、位置等几何特征以及相邻实体的拓扑关系。空间数据须用统一的坐标系统进行空间定位，要求不同坐标系间可相互转换。空间数据用空间拓扑信息来表示空间位置关系。

属性特征。空间数据的属性特征是指实体或现象的非几何特征，如空间实体的名称、类别、数量等属性。例如一栋房屋的属性可能包括门牌号、单位名称、层数、建筑结构类型和修建时间等。属性数据作为空间数据的重要组成部分，只有同几何图形相结合，才能表达空间实体的全貌。

(2) 时间特征。空间数据的时间特征是指实体的空间特征和属性特征随时间而变化的动态特征，可以同时随时间变化，也可以相互独立随时间变化。例如，行政区界的变更表示空间位置数据的变化，而该区界的人口情况的变化在区界未变更时仅表示属性数据独立随时间变化。空间数据的时间特征说明了 GIS 是一个动态的、复杂的空间管理系统。该特征往往是隐含的，在 GIS 数据表示中加上时间轴会增加空间数据处理的难度和复杂度，因而在一般情况下是将空间数据的时间特征隐含在数据中，用时间属性来表示空间数据的时间特征。

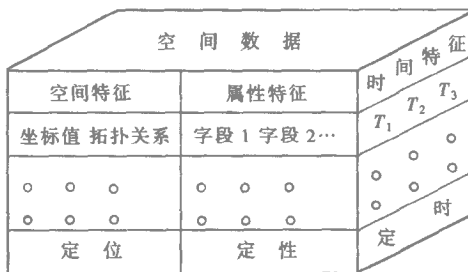


图 1-5 空间数据的特点

空间数据还具有图形、属性数据共存，图形、图像、多媒体数据共存等特点，总之，地理空间数据有着定位、定性、定时的特点，如图 1-5 所示。

1.5 地理信息系统的数据结构

数据结构 (Data Structure) 即数据组织的形式，是适合于计算机存储、管理和处理的数据逻辑结构。对空间数据而言，则是地理实体的空间排列方式和相互关系的抽象描述。

GIS 主要有两种空间数据结构：栅格数据结构和矢量数据结构。前者是将研究区域划分成一系列正方形或矩形栅格 (格子单元) 而形成的，每个栅格或像素的位置由栅格所在的行列号定义，栅格的值为栅格所表达内容的属性值。矢量数据结构是用一系列 X、Y 坐标来描述点、线或多边形要素的组织结构。

1. 栅格数据结构

栅格数据结构实际上就是像元阵列，每个像元由行列号确定它的位置，且具有表示实

体属性的类型或值的编码值。点实体表示为一个像元，线实体则表示为一组相邻像元集合，而面实体由一片相邻像元集合表示。

栅格数据是二维平面上地理数据的离散量化值，例如下述的简单影像：

A	B	B	B
B	B	A	A
A	A	A	B
A	A	B	B

在计算机内是一个 4×4 阶的矩阵，但在外部设备中，没有矩阵存储器，通常是以左上角开始逐行逐列存储，上例的存储顺序为 **ABBB BBAA AAAB AAB B**，通常是一个文件中存储多层信息，记录每个像元的行列号及其相关信息。像元值一般为整数型。

地理数据一般有较强的相关性，相邻像元的值往往是相同的，因此可以用某种编码对数据进行压缩，如行程编码、变长编码和二叉树编码等（具体编码方法请参阅有关书籍）。以行程编码为例，对于上例有：**1A3B, 2B2A, 3A1B, 2A2B**。如果行之间不间断编码则有：**1A5B5A1B2A2B**。显然，编码技术节省了不少空间。数据相关性越强，节省空间越多。

虽然栅格数据结构具有结构简单、易与遥感结合、空间叠置和空间分析易于进行、速度快、有些操作效率高等优点，但也存在许多不足之处，如数据量大、输出不精美、精度低、难以建立地物间的网络拓扑关系等。

2. 矢量数据结构

矢量数据结构直接以取样点坐标为基础，尽可能将目标表示得精确无误，通过记录坐标的方式精确地表示点、线、面实体位置。点实体表示为空间的一个坐标点： (X, Y) ，线实体表示为多个点组成的矢量弧段： $(X_1, Y_1) (X_2, Y_2) \dots (X_n, Y_n)$ 而面实体则表示为封闭的曲线段（组）： $(X_1, Y_1) (X_2, Y_2) \dots (X_n, Y_n) (X_1, Y_1)$ 。

矢量数据除了记录空间目标的几何位置外，还要考虑与这个目标相关的属性信息以及空间目标之间的相互关系（拓扑关系），以满足空间查询和空间分析的需要。其最突出和最具特色的优点就是能够完全显式地表达节点、弧段、面块之间的所有关联关系。在矢量数据中，用拓扑信息来描述空间的“相邻”、“包含”等关系，从而清楚地表达空间地物之间的结构。另外，根据属性特征的不同，点可用不同的符号、颜色以及大小来表示，线可用不同的样式、颜色以及粗细来描绘，而面则可以填充不同的图案和色彩加以分别。

矢量数据具有较高的空间分辨率，结构紧凑，冗余量小，易于网络拓扑分析且输出精美，但也有不足之处，如结构复杂，不易与遥感结合，信息复合难度大（多边形叠置，空间均值处理）等。

3. 矢量栅格一体化数据结构

矢量和栅格数据结构各有所长，它们之间可以通过一定技术相互转换。现在许多 GIS 软件中，既有栅格结构又有矢量结构，形成一种混合数据结构。但点、线状地物在栅格结构中难以独立表达，往往会与相邻的面状地物发生矛盾，因而限制了它们与栅格影像数据直接交互使用的能力，这就要求一种数据结构，它不是矢量与栅格的简单混合，而是既有矢量特点又有栅格特点的综合数据结构，即矢量栅格一体化结构。

矢量栅格一体化结构本质上是一种以栅格为基础的数据结构。采用细化格网、填满线状目标路径和充满面状目标空间的表示方法，将线用元子填充法表示，若采样点正好落在栅格单元之内，则数据既有矢量特征（ X, Y ），同时也具有栅格特征，每个面状目标除了周围的 X, Y 坐标外，还记录面状的栅格，所以以栅格为基础，避免了传统不足（精度不够）。

(1) 细分格网。由于栅格的精度低，为了提高精度，需要在有点、线通过的网格内再细分成 256×256 个子网格（又称细分格网），基本栅格和细分格网均采用线性四叉树十进制地址码（Morton 码）进行编码。这样一个点位置用两个 Morton 码表示，第一个表示基本格网中的位置，第二个表示细分格网的 Morton 码。对于组成线状和面状地物的弧段而言，不仅记录原始采样点的位置，而且记录每条弧段通过每个基本格网边的交叉点，以解决线状目标与栅格数据交互叠置问题。

(2) 点状目标的数据结构。它与矢量拓扑结构相似，使用两个 Morton 码表示如下：

NODE ID	M_1	M_2	关联弧段
---------	-------	-------	------

M_1, M_2 分别表示基本格网和细分格网中的 Morton 码。形式上， (M_1, M_2) 类似于 (X, Y) 坐标，当然实质含义迥异。

(3) 弧段的数据结构。它亦与矢量拓扑结构相似，但除了采用 Morton 码表示位置特征外，还特别强调应记录弧段通过的所有格网。即

ARC ID	起始节点	终止节点	左多边形	右多边形	中间节点 (M_1, M_2)
--------	------	------	------	------	---------------------

(4) 面状地物的数据结构。在综合数据结构中，面状地物的数据结构不仅包含了组成该面状地物周边的弧段，而且要包括该面状地物中的面域栅格。这些面域栅格由二维行程编码组织，并且用循环指针将每个面状地物所属的叶节点串起来，然后再建立面状地物的数据结构，即

面块标识号	周边的弧段标识号	中间面域叶节点信息
-------	----------	-----------

矢量栅格一体化数据结构既有矢量特点，精度高，易建立拓扑关系，便于表达实体属性，又有栅格特点，具有整体性的表达能力，易进行叠置分析和易于与遥感影像结合。

1.6 地理信息系统的数据库模型

数据模型是数据库系统中关于数据和联系的逻辑组织的形式表示。从 GIS 数据管理的角度看，数据模型主要是指用来管理和存储空间数据的数据库模型。换句话说，空间数据模型是关于现实世界中空间实体 (Spatial Entity) 及其相互间联系的概念，是描述空间数据内容以及数据关系的工具。

GIS 数据模型不仅要表达和处理空间实体及其相互之间的关系，而且要考虑计算机存储和处理的性能。数据模型的选择对系统的有效性和适用性有很大影响。GIS 常用的数据模型有层次数据模型、网络数据模型、关系数据模型和面向对象数据模型。

1. 层次数据模型

层次数据模型采用的是树型数据结构，即将数据组织成一对多的关系，一个父节点对应多个子节点，而一个子节点只对应于一个父节点。在层次模型中，各种数据之间存在一种依次从上而下的层次关系。例如，图 1-6 所示的地图的层次模型如图 1-7 所示。

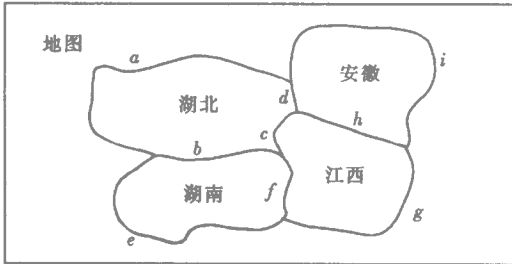


图 1-6 地图

层次数据模型反映了现实世界中实体间的层次关系，层次结构是众多空间对象的自然表达形式，并在一定程度上支持数据的重构。其优点是：结构清晰，易理解，存取方便，速度快，数据修改和数据库扩展较易等；不足之处为：结构呆板，没有灵活性，不能表达多对多关系，数据重复存储，冗余量大，不适合表达数据的拓扑关系。

关系。

2. 网络数据模型

网络数据模型采用图数据结构，即将数据组织成多对多关系，且数据之间具有显式的连接关系，但没有明显的层次关系，一个父节点对应多个子节点，一个子节点也可能对应多个父节点。如图 1-8 所示。

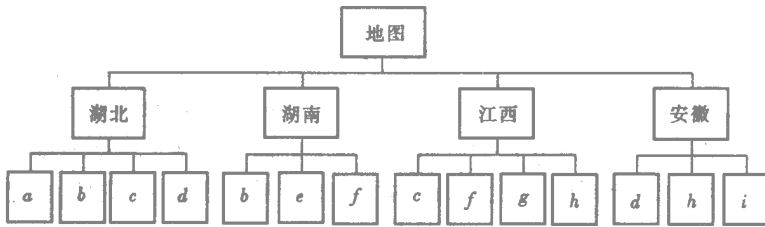


图 1-7 层次数据模型

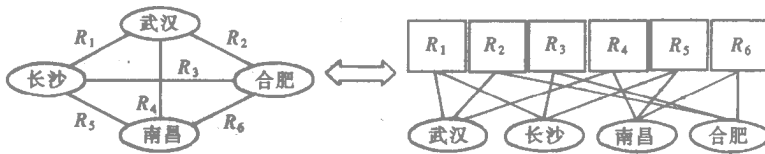


图 1-8 网络数据模型

网络数据模型反映了现实世界中常见的多对多关系（网络拓扑关系），在一定程度上支持数据的重构，具有一定的数据独立性和共享性。其优点是：极大地压缩了数据量，易于表达复杂的拓扑关系；不足之处为：指针数据项使数据量额外有所增加，指针随着数据库中数据的修改而变化，其建立和维护较为麻烦。

3. 关系数据模型

关系数据模型，采用线性表数据结构，即把数据的逻辑结构归结为一定条件的二维表

中的元素，这种表就称为关系。一个实体由若干个关系组成，而关系表的集合就构成了关系模型。图 1-6 的关系模型如图 1-9 所示。

关系数据模型以简单、灵活的方式表达了现实世界中各种实体及其相互间关系，并支持数据的重构，数据描述具有较强的一致性和独立性。其优点是：数据结构简单、清晰、灵活，能够直接处理多对多的关系；可满足所有用布尔逻辑运算和数学运算规则形成的查询要求；还能搜索、组合和比较不同类型数据；数据的增删改方便，易维护和理解；另外，数据独立性强，便于数据集成。而不足之处为：许多操作要求在文件中顺序查找满足特定条件的数据，当涉及的目标很多时，操作时间长，效率低。

地图	湖北	湖南	江西	安徽
湖北	a	b	c	d
湖南	b	e	f	
江西	c	f	g	h
安徽	d	h	i	

图 1-9 关系数据模型

4. 面向对象数据模型

上述三种数据模型都是以记录为中心的，语义比较贫乏，数据类型少，从而难以表达和处理复杂的空间地理实体，难以实现快速查询和复杂的空间分析，不能很好地面向对象。因而，研究更接近于人类通常的思维方式，便于操作管理空间实体数据的模型，成为 GIS 研究的重要课题。

面向对象的数据模型就是针对上述三种数据模型的缺点，以接近于人类思维方式设计的，将现实世界的实体都抽象成对象，即将模拟实体状态的数据和操作封装成相对独立的模块，然后进行分类和概括，以人们认识问题的自然方式将所有的对象构建成一个分层结构来描述问题领域中各实体之间的相互关系和相互作用，从而建立起一个比较完整的结构模型，使得对现实世界的构成与人们认识问题的方式直接对应，如图 1-10 所示。

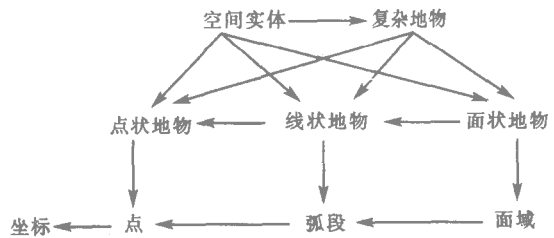


图 1-10 面向对象数据模型

面向对象模型引入类的概念，即相同对象的集合，允许定义任意复杂的数据类型，模型化真实世界的静态特征，并提供与对象类型相关联的操作，记录对象的动态特征。一个空间实体对应相应类的一个对象。面向对象的数据模型具有表示和构造复杂对象的能力。其表示自然，易于理解；通过概括、聚集、联合形成“超类”及通过继承形成子类，使面向对象模型更加灵活，增强了系统扩充能力，其封装性则利于系统的模块化设计。

1.7 空间数据库的组织和管理

地理信息系统数据库主要涉及图形数据和属性数据。如何组织好两者的关系，即图形数据和属性数据的连接，就成了空间数据库组织和管理的核心问题。

单纯选用一种商品化数据模型，如关系数据模型来组织管理图形数据和属性数据并不理想，尤其对于复杂目标的空间数据组织和管理。但是鉴于一些通用数据库管理系统（DBMS）通用性强，在技术方面较为完善，效率高，因此，目前许多 GIS 均是以关系型数据库管理系统（RDBMS）为基础加以扩展来组织管理图形数据和属性数据。

1. 文件与关系数据库混合管理

关系数据库中的记录一般是定长的，空间数据往往是变长的，因而很难表达清楚复杂的空间数据及其关系，另外以单纯的文件形式，数据文件之间又往往不能发生联系，难以有效地组织管理相互关联的图形数据和属性数据，因而采用混合管理方式，分别用两个子



系统关联图形数据和属性数据。其中图形数据存储在文件系统中，属性数据存储在关系型数据库中。这样图形数据和属性数据独立管理，通过标识码 OID 联系，如图 1-11 所示。在分析处理目标时分别访问两个子系统，再把它们结合起来得出结论，如图 1-12 (a) 所

图 1-11 图形和属性数据

性数据的关联这种组织方法结构简单、通用性强，但分析处理时需同时查询访问两个子系统，使得查询难以优化，效率较低。另外，数据完整性较差，例如可能会出现图形数据访问中已被删除的目标，其属性数据在属性数据库中还存在，或出现相反的情况。

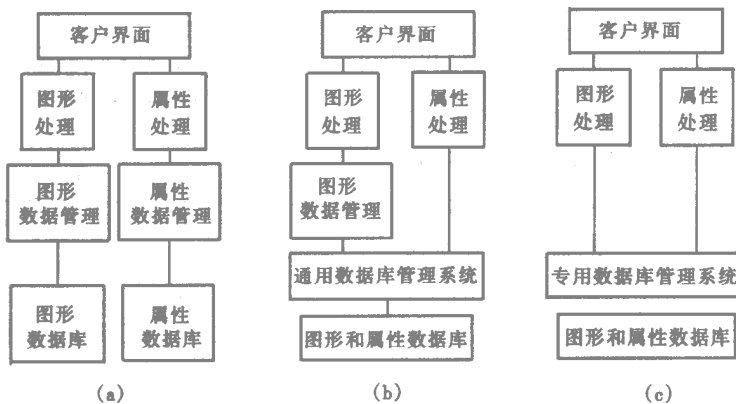


图 1-12 空间数据库的组织和管理

(a) 文件与关系数据库混合管理；

(b) 对象-关系数据库管理；(c) 面向对象空间数据库管理

2. 全关系型数据库管理

这种组织方法也是基于关系型数据库的，只是将一些定长字段扩充为变长字段（现代通用的 RDBMS 已允许变长记录，允许过程直接调用 SQL 语言），用来存储图形数据，从而将图形数据这一变长数据完全交给关系型数据库管理。其特点是查询方便、效率高、数据完整性好，但实现较麻烦、嵌套关系不易表示。

3. 对象—关系数据库管理

这种组织方法是在关系型数据库中引用面向对象机制，形成对象—关系数据库来组织管理图形数据和属性数据，主要有以下两种实现方式：

(1) 在现有关系型数据库的面向记录的数据模型顶层加一层面向对象的数据模型，在此将地理信息查询语言转化为标准的 SQL 查询语言，借助索引数据的辅助关系实现对图形数据的操作。

(2) 修改现有关系型数据库系统，使之支持面向对象，以提供通用的数据库环境，形成对象—关系数据库系统。如图 1-12 (b) 所示。

4. 面向对象空间数据库管理

这种组织方法使用全新的面向对象模型，直接操作图形数据和属性数据，实现图形数据和属性数据的完全统一管理。如图 1-12 (c) 所示。有以下三种实现方式：

(1) 扩充面向对象程序设计语言，增加 DBMS 的特性。其优点是充分利用语言本身功能，易结合 C++ 软件，使用面很广；缺点是没有充分利用 DBMS 功能，仅是将语言进行扩展。

(2) 扩充 DBMS，在 RDBMS 中增加面向对象特性，其优点是利用 RDBMS 功能，可采用面向对象程序设计语言；缺点是数据库的输入/输出 (I/O) 检查费时，完成的操作比单纯的面向对象数据库管理系统 (OODBMS) 慢。

(3) 建立全新的支持面向对象数据模型的 OODBMS。其优点是能开发纯粹的 OODBMS，操作效率高，重视计算完整性，非过程查询；缺点是开发量大、复杂。

1.8 地理信息系统与其他系统的集成

1.8.1 GIS 的系统集成与技术集成

从学科和技术的角度看，地理信息系统是一门集多学科和技术于一体的新型交叉边缘学科，同时也是多学科集成并应用于多领域的新兴技术。目前，GIS 集成很大一部分体现在与其他系统的集成上。

系统集成是采用一定结构形式。通过某种技术并利用其内在联系将多个系统有机结合

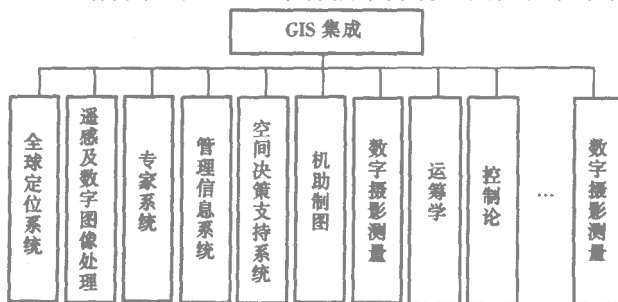


图 1-13 GIS 集成