资源信息系统教程

主　编：吴冲龙

副主编：汪新庆　刘　刚

地质出版社

·北　京·

内 容 提 要

本教材经地质工科资源普查与勘探课程指导委员会审查，适用于资源勘查工程专业学生使用。全书共分六章，包括地矿资源勘查系统的基本理论、开发方法、系统分析，地矿资源属性数据库子系统的设计与应用，空间信息子系统的设计与应用，图件机助编绘子系统的设计与应用。

本书可供资源勘查工程专业、地质工程专业、理科地矿类专业本科生使用。

图书在版编目（CIP）数据

资源信息系统教程/吴冲龙主编. —北京：地质出版社，2004.7
ISBN 7-116-04077-3

Ⅰ.资… Ⅱ.吴… Ⅲ.矿产资源—地质勘探—信息系统—高等学校—教材 Ⅳ.P624

中国版本图书馆 CIP 数据核字（2004）第 043898 号

责任编辑：李 燕 祁向雷
责任校对：郭惠兰
出版发行：地质出版社
社址邮编：北京海淀区学院路 31 号，100083
电 话：（010）82324508（邮购部）；（010）82324577（编辑室）
网 址：http://www.gph.com.cn
电子邮箱：zbs@gph.com.cn
传 真：（010）82310759
印 刷：北京京科印刷有限公司
开 本：787mm×1092mm 1/16
印 张：15.5
字 数：377 千字
印 数：1—2000 册
版 次：2004 年 7 月北京第一版·第一次印刷
定 价：20.00 元
ISBN 7-116-04077-3/P·2474

（凡购买地质出版社的图书，如有缺页、倒页、脱页者，本社出版处负责调换）

前　言

本教材经地质工科资源普查与勘探课程指导委员会审查立项，适用于资源勘查工程专业四年制本科生，也可作为地质工程专业和理科地矿类专业本科生的教学参考书。本教材的内容适合 60 ~80 学时的课堂教学，与教材相匹配有上机实习软件和指导书。

地矿资源勘查工作每日每时都在采集信息、加工信息、存贮信息、处理信息和利用信息。目前，信息技术已经渗透到地矿工作的全过程，成为新技术新方法应用的主流，大大加快了资源勘查工作信息化和现代化进程。地质矿产资源勘查工作的信息化水平，也已经成为衡量地质矿产资源工作现代化程度的主要标志。地质矿产资源工作信息化工程的核心是信息系统建设，而信息系统建设的主要目标，是充分发挥资料数据的利用价值。计算机技术使数据的存贮、检索和处理实现自动化、高速化，既提高工作效率，又保证工作精度；信息高速公路的迅速开通，则使全国范围内的信息联机检索和交互处理变为可能。

建设数字国土，实现地矿资源勘查与管理信息化可以有不同的解决方案，但基础都是基层勘查单位（数据采集点）点源信息系统的建设与应用。点源信息系统是实现地矿资源勘查信息综合利用、综合处理和综合解释，从而提高地质资料利用价值、推动地质科学向定量化方向发展的重要条件，同时也是为各类、各级综合数据库系统提供数据流的源泉。对于基层勘查和研究单位，它是功能强劲的工作站；对于全国地矿信息系统而言，它是数据齐备的网络结点。本教材着重介绍这方面的理论、方法与技术，其中包括资源信息科学与资源信息系统的基本概念，信息系统工程的基本原理与方法以及资源勘查的数据库子系统、空间信息子系统、图件机助编绘子系统的设计与应用等方面的内容。

本教材在编写过程中力求体现如下特色：①实用性。内容针对地矿工作各个环节的信息采集、加工、存贮、管理、处理与应用，与我国各地矿资源勘查（察）单位正在实施的“地矿工作信息化”目标和任务相适应。②先进性。所阐述的理论、方法和技术，一部分是取自国内外文献、专著和相近教材中的新成果；另一部分是取自本书编者的近年来研究成果。③系统性。系统地介绍了地矿资源信息系统的基本知识、基本原理和基本方法，学生们结合上机实习，可以从中学习到系统的、完整的实用知识。④实践性。注重理论与实践相结合，传授地矿资源信息系统的应用设计方法和操作方法。教材配备有大量实习软件，学生们通过实习可以很快地掌握实际应用的方法和技巧。教材内容涵盖较宽，各专业在教学中应当有所取舍。

本教材编写分工如下：第一、二章，吴冲龙、毛小平、田宜平；第三章，吴冲龙、李绍虎、张夏林；第四章，汪新庆、何珍文、韩志军；第五章，李伟忠、刘志军、袁艳斌；第六章，刘刚、罗映娟、孔春芳、翁正平。全书初稿于 1999 年由吴冲龙汇总编纂。赵鹏大院士和王世称、侯景儒、谭照华教授等审阅了书稿，并提出了宝贵的修改意见。作者根据这些意见并结合 4 年来在中国地质大学（武汉）9 个班的试用情况，对书稿做了进一步修改。值此教材出版之际，谨向所有帮助过我们的专家和同行们致以衷心的谢意。

编者　2003 年 10 月 26 日

目　　录

前　言
第一章　绪　论 …… (1)
　第一节　地矿资源勘查数据采集与管理技术 …… (1)
　第二节　地矿资源数据处理与资源预测评价技术 …… (6)
　第三节　地理信息系统技术在资源勘查中的应用 …… (11)
第二章　地矿资源勘查信息系统的开发方法 …… (13)
　第一节　地矿资源勘查信息系统的基本概念 …… (13)
　第二节　地矿资源勘查信息系统的开发方法 …… (22)
　第三节　地矿资源勘查信息系统的开发过程 …… (32)
第三章　地矿资源勘查系统分析 …… (38)
　第一节　地矿资源勘查系统需求与工作环境分析 …… (38)
　第二节　地矿资源勘查业务现状与数据现状分析 …… (44)
　第三节　地矿资源勘查实体模型研究 …… (51)
第四章　地矿资源属性数据库子系统的设计与应用 …… (72)
　第一节　数据库系统原理与方法概述 …… (72)
　第二节　地矿资源属性数据库结构设计 …… (85)
　第三节　地矿资源属性数据库数据字典设计 …… (92)
　第四节　地矿资源属性数据库系统的应用 …… (104)
第五章　地矿资源空间信息子系统设计与应用 …… (121)
　第一节　地矿资源空间信息子系统的结构与功能 …… (121)
　第二节　地矿资源空间数据库的应用设计 …… (131)
　第三节　地矿资源空间数据编辑、处理与空间分析 …… (156)
　第四节　地矿资源空间信息子系统应用举例 …… (175)
第六章　地矿资源图件机助编绘子系统的建立和应用 …… (190)
　第一节　地矿资源图件机助编绘子系统概述 …… (190)
　第二节　地矿资源图件机助编绘子系统的开发基础 …… (195)
　第三节　专题图件机助编绘的程序设计 …… (213)
　第四节　地矿资源图件机助编绘子系统的应用 …… (226)
参考文献 …… (238)

第一章　绪　论

地质调查、矿产勘查与工程勘察各专业的工作过程，本质上都是信息的获取、整理、处理、解释和应用过程。从野外数据采集到室内数据综合整理、数据管理、数据处理、图件编绘、成果分析与解释、勘查对象的预测与评价，再到工作成果的保存、管理使用和出版印刷，甚至地矿勘查工作的科学管理与决策等等，无一不与信息技术紧密相连。地矿点源信息系统就是信息技术在地矿勘查与地球科学领域的一种综合性应用系统。戈尔(1998) 提出的“数字地球”概念，作为一项整体性、导向性的国家战略目标，在一些国家引起了积极反响，也受到我国政府和科技界的关注。“数字地球”包括3个重要的组成部分：信息获取、信息处理、信息应用。我国在新一轮国土资源大调查中开展的地矿勘查工作信息化工作——“数字国土”工程的规划、计划和实践，显然可以归入这一范畴。地质矿产勘查工作信息化涉及地质学、矿产勘查学、工程地质学、计算机应用和信息系统领域的诸多分支学科。这里仅就所涉及的几个主要方面作些简单介绍。

第一节　地矿资源勘查数据采集与管理技术

地矿勘查数据采集与管理技术主要包括数据采集技术、关系数据库技术、空间数据库(图形库) 技术和信息发送服务技术。这些技术既是开发和建造地矿信息系统的基础，也是地矿勘查与地球科学领域中一切计算机应用的基础。

一、地矿资源勘查数据采集技术

数据采集工作的自动化程度，是地矿勘查领域计算机应用的瓶颈问题。由于地矿勘查数据具有多源、多量、多类、多维和多主题特征（吴冲龙，1998)，其采集和输入方式不可能统一，需要具体问题具体解决。根据数据来源，地矿勘查数据主要有7类：地球物理勘探与遥感数据、地球化学勘探数据、野外观测数据、室内化验测试数据、地形地物的三角测量数据、综合整理数据和图件数据。

地球物理勘探与遥感数据基本上都是用仪器进行测量和记录的，但大致可分为人工读数记录、模拟自动记录和数字自动记录3种方式。目前，西方先进国家的地球物理勘探几乎全部实现了数字化自动记录，并且直接采用微机控制和接收；各种航空、航天遥感数据则一开始就是采用数字化自动记录的。我国各地矿勘查部门在航天航空遥感、航空物探和地震勘探数据的采集方面，也已全部实现了数字化自动采集，但地面电法、磁法和重力数据的采集不少仍沿用传统的手工方法。室内化验测试数据主要来自化学分析、同位素分析、岩矿鉴定和各种物理性质测试，其中包括大量的地球化学勘探数据。目前西方先进国家已经对其中的多数实现了模拟自动记录或数字自动记录，少数仍停留在人工记录的水平上。我国除科研院所、高等学校和省局一级实验室或测试中心引进的国外大型仪器设备具有自动记录功能外，多数实验室仍保持手工记录方式和手工入机方式。具有自动记录的仪

器设备，由于软件功能不同，入机的方便程度差别也很悬殊。有的已经实现全自动记录，例如英国近期开发制造的激光拉曼光谱仪，将数据采集、结构分析和成分分析结合为一体，能同时完成各种探测参数记录和分析结果记录，甚至实现了测试过程的全程微机自动控制；有的则只是实现人机交互的半自动记录，或者只能进行简单的数据自动拾取。

野外观测数据的采集包括区域地质调查和矿产、水文、工程等地矿勘查中所涉及的露头、坑探、槽探和钻探的肉眼观测编录数据。国外 20 世纪 80 年代在沿用传统笔记本记录方式时，对采集内容进行了标准化、定量化和代码化，并制定出标准化表格，在野外按表格填写，回室内由人工键入计算机。对于那些不需要计算机处理的描述性信息，在野外仍然用文字记入笔记本，回室内也不输入机内，只作为技术档案保留，以备查询。到了 20 世纪 90 年代初期，人们开始广泛使用掌上机采集数据，其采集内容也在 20 世纪 80 年代标准化、定量化和代码化的基础上作进一步改进。掌上机具有体积小、质量轻、电池寿命长、携带方便的特点，可以使用 Windows-CE 开发的专用软件来采集野外属性数据，还能够接受全球定位系统（GPS: Global Position System）导航仪提供的空间数据。其缺点是功能低、容量小，目前还不能装载数据库系统（DBS: Data Base System）和地理信息系统（GIS: Geographic Information System），难以采用“多 S”结合与集成技术。从近年来计算机技术的发展情况看，功能强劲的便携机（又称为膝上机或笔记本电脑），在工作条件较好的地区有可能取代掌上机而成为野外数据采集的主要工具。

与此同时，野外数字录像和数字照相技术在发达的西方国家也开始得到广泛应用，所录的影像信息可以直接输入计算机内。显然，采用“多 S”结合与集成化装备来补充传统的“老三件——铁锤、罗盘、放大镜”，改变野外数据采集方式的落后面貌，是地球信息科学与技术发展的必然趋势。相比之下，我国目前的地矿勘查野外数据采集现代化水平比较低，基本上保持传统方式，目前已取得了许多重要进展。

地形地物的三角测量数据是地矿勘查的空间定位数据。地矿勘查所使用的地形图，通常是测绘部门提供的。而矿区、勘查区、工程施工区的大比例尺地形图，以及某些地质点、矿化点、采样点、探槽、探硐、钻孔和坑道的位置和高程数据和图形，还是需要自己测量和绘制的。以往测量数据的获取，在地面上主要靠肉眼仪器观测、读数和手工记录，在空中主要靠感光摄影和人工选点、转点、量测像点像片坐标或模型坐标。目前由于全球定位系统（GPS）的引进，已经有可能实现全部测绘数据采集的自动化。GPS 是美国国防部为了满足其军事部门对海、陆、空高精度导航、定位和定时的需要而建立的一种卫星定位与导航系统。用 GPS 同时测定三维坐标的方法，将测绘定位技术从陆地和近海扩展到整个海洋和外层空间，从静态扩展到动态，从事后处理扩展到实时（准实时）定位与导航，从而大大地拓宽了它的应用范围和在各行各业中的作用（李德仁等，1998）。利用机载 GPS 和 RS、GIS、DPS（Digital Photograph System——数字摄影测量系统）及 ES（Expert System——专家系统）的集成化技术，可以实现全自动的空中三角测量；利用手持 GPS 导航仪和误差校正仪（差分仪）配合，也可以自动获取高精度的位置（水平误差 cm 级）和高程数据（垂直误差 dm 级）。如果采用单纯手持 GPS 导航仪单机游走方式，水平误差和垂直误差都在 30m 左右，而且漂移方向带有随机性，只能满足小比例尺的资源调查工作需要。

上述各种勘查资料的综合整理，也是获取数据的重要渠道。综合整理包括原始资料的

分类归并、分析套合、数据融合、数据转换、计算处理和图件编绘等等。综合整理也有手工方式、人机交互方式和全自动方式之分。一般地说，除了图件之外，采用各种方式整理得到的数据，通常仍用同种方式记录保存或输入计算机内保存。图件是地矿勘查与研究成果的主要载体，其数据形式较为特殊。采用机助编绘或计算机全自动编绘的图件，其数据可直接转存到图形库中，不存在采集和入库问题；凡采用手工编绘的图件，特别是以前的旧图件，数据输入计算机较为麻烦。能否准确、迅速而方便地采集旧有地矿勘查图件及公用地理底图的图形数据，是实现旧地矿图件计算机管理和新地矿图件计算机辅助编绘的关键环节。西方发达国家均采用手扶跟踪矢量数字化方式或自动扫描栅格数字化方式，或自动扫描栅格数字化加屏幕矢量化方式，来实现这些图形数据的采集。就目前的技术情况而言，第一种方式效率太低；第二种方式仅适合于遥感影像图、照片和分版单色地图；第三种方式可用于处理其他复杂图件。例如区域地质图和某些专项地质图，其输入后的处理工作很繁杂，需要提高其自动化程度。

二、地矿资源勘查数据库与图形库技术

数据库（Data-Base）是实现有组织地、动态地存储大量关联数据、方便多用户访问的计算机软硬件资源所组成的系统。它具有数据充分共享、支持交叉访问以及应用程序高度独立的特点，不仅能够有效地存贮、管理和检索数据，而且能够辅助各种实际系统的建模和模拟。由于在地矿勘查过程中所获取的数据资料具有反复使用、长期使用的价值，因而有长期保存的必要；同时，由于获取时的代价昂贵而又适用于不同勘查对象、不同勘查目的和不同勘查阶段，它又有共享的必要。这两种必要性的存在，使得地矿勘查数据和图件资料成为国家的宝贵财富，需要采用计算机技术加以妥善管理。因此，地矿勘查数据库和图形库的建设不但成为地矿信息系统的核心问题，而且成为地矿勘查工作信息化工程的核心问题，受到各国政府部门和勘查（察）单位的极大重视（Johnson et al.，1991）。

美国地质调查所（USGS）从20世纪60年代起投入了巨大的人力、物力和财力，经过30多年的努力，先后建立了许多重要的数据库。例如美国资源库、全国煤炭数据库、全国水文数据存储检索系统、海洋地质数据库、地球化学与岩石学分析数据库等等。在这些数据库中存放了美国数万个矿床和矿点信息、数十万处钻孔和野外露头的观测数据（姜作勤，1991）。USGS还建立了全国数字制图数据库（NDCDB），存入了全美国数万幅简化地理底图。美国各大石油公司都建设了自己的钻井数据库，完整地存储了钻井工程信息，并且在不同公司之间设置接口，允许共享对方的数据资源。他们还基于计算机系统和卫星通讯技术，建立了钻井实时分析和指挥中心，实现了数据的实时采集、实时入库和实时使用，为钻井工程的规划、分析、设计、协调、决策及控制提供了可靠的依据。英国和澳大利亚也先后建立起许多重要的数据库，实现了地矿勘查数据社会化服务。

我国资源数据库与图形库建设起步较晚，目前仍处于初级阶段。原地质矿产部曾于1986年对全国资源信息系统的建设做出了规划，将目标定为“到2000年基本上实现主要地矿成果资料的存储、管理和使用的计算机化，大幅度降低地质报告和地质成果图存放、修复费用，提高资料的利用率和可信度；利用计算机系统功能，节省查找和翻阅资料的时间，更有效地发挥以往地矿成果的作用”。根据总体方案，该系统共包括12个子系统、77个国家级数据库和若干个模型库、方法库。到目前为止，已经建成各种综合数据库和基础数据库，内容几乎涉及地矿勘查工作及其管理的各个方面。中国地质调查局于1999年重

新进行国家基础地质数据库建设全面部署，并着手开展国家基础性、公益性地质图件的数字化及系列数据库建设。通过几年的努力，国家基础地学数据库的建设取得重大进展。关键生产与处理技术接近发达国家水平。至 2003 年 6 月，已建立多个专业的总数据量大于 3TB 的上万个数据库（其和日格等，2003），主要包括：①区域地质图空间数据库与小比例尺地质图空间数据库；②区域水文地质图数据库；③矿产地质数据库与钻孔地质数据库；④专题地质数据库。此外，区域海洋地质数据库的总体设计也已经完成。

中国石油天然气总公司、煤炭部、冶金部、核工业部地质局、有色金属总公司和水利部等部门，也都分别于 20 世纪 80 年代中后期做出了相应的近期和长远的发展规划，并且分别开发了大量专业数据库和点源数据库。

随着“数字地球"概念的提出和“数字中国"、“数字区域”和“数字城市”工程的实施，国家信息中心在“九五”期间开始了基于网络的分布式空间信息系统研究，在信息共享标准、共享政策、模式、软件平台等方面做了大量工作，初步完成空间对象数据库管理系统的模型设计。还开发了支持空间元数据网络系统的空间元数据编辑器、空间元数据服务器管理系统软件，提出了空间信息共享平台系统模型，设计了系统的总体结构框架，研制了支持 SDE 的空间信息共享平台。迄今为止，已经完成了空间信息共享网络技术系统的技术原型，形成了试验运行的空间信息交换网络应用系统，初步建成国家公用基础地理信息平台，建立了包括全国土地、水、森林、海洋、矿产资源和地区经济等 6 个全国资源综合数据库，1 个省级资源环境与社会经济示范数据库（海南），全国基础地理数据库（1：400 万、1：100 万、1：25 万基础地理数据库和 1：50 万全国地理底图数据库、1：100 万和 1：50 万海洋基础地理信息系统）、全国遥感资源环境综合数据库和 NREDIS 综合数据库在内的 10 个可供分布式网络共享的空间信息系统及其空间元数据库群。

总之，我国的“数字国土”工程正在蓬勃地向前发展。但总的看来我国的点源数据库建设仍然比较薄弱，一些基层单位为着各自应用目的而建立的信息系统，数据环境基本上属于应用数据库范围，大多数是简单地利用现成的商业化软件来装载数据，较少做高层次的二次开发。在资源图件库方面的工作也缺乏统一的部署和管理，不但未严格遵循统一的规格和标准，重复采集数据的现象也十分严重。

为了解决上述问题，有关政府部门和标准化机构一方面组织技术攻关和软件开发；一方面抓紧法规和标准的制定。全国地质矿产标准委员会与各地矿部门经过共同努力，至“九五”末共制定了国家标准 1 种，即《地质矿产术语分类代码》；行业标准 8 种，其中包括：《地质钻孔（井）基本数据文件格式》、《石油钻井地质数据文件格式》、《水文地质数据文件格式》、《煤田地质钻孔数据文件格式》、《固体矿产钻孔地质数据文件格式》、《固体矿产矿点（床）地质数据文件格式》、《地下水资源数据文件格式》、《物探化探异常数据文件格式》。另外还有一些行业标准也正在加紧制定之中。这些国家标准和行业标准，将成为规范我国开展地矿勘查与基础地质调查数据库建设的重要依据。

从规模与分布特征上看，目前国际上的地矿勘查数据库和图形库有两个并行的发展方向，一个是大型集中式方向，一个是微型分布式方向。西方诸国早期所建立的全国规模信息系统，以及我国已建立的各种全国规模的数据库，基本上都是大型集中式的。现行的大型集中式数据库基本上都建立在巨型和大、中型机上，其优点是数据专业化程度高、分类清楚、内容简洁，便于集中管理，有利于支持上级机构和政府部门的决策；缺点是专业划

分过细，不便于各地使用，难于组织、容纳繁多的数据类别和复杂的数据结构，更难于应付勘查和科研中日益增多的信息处理需求。近年来，随着高容量、高速度、低价格的微机和工作站大量涌现，分布式数据库和图形库系统受到普遍的重视，特别是网络技术的发展和信息高速公路的建成，使分散于各地的计算机资源和信息资源的管理、交叉访问及远距离传输成为可能，分布式数据库和图形库系统的发展前景得到更普遍承认。

数据库的另一个重要发展方向是与信息处理相结合。当前评价一个资源数据库系统的功能，不仅看其数据存取效率和方便程度，而且看其能否有效地支持各种复杂的数据处理、图件编绘、过程模拟和评价决策系统的运行。大型综合性数据库系统、数据银行（Data Bank）、数据仓库（Data Warehouse）和数据集市（Data Mart）的出现和在石油天然气勘探领域的应用，代表了这一发展方向的一个侧面；而地理信息系统（GIS）的引进和在各种地矿勘查领域的广泛应用，则代表了这一发展方向的另一个侧面。

数据仓库技术是一种分析型数据处理新技术，最初应用于企业经营管理的决策支持。数据仓库有四个基本特征：数据是面向主题的；数据是集成的；数据是不可更改的；数据是随时间不断变化的（Inmon，1992）。

GIS 不仅将属性数据库与空间数据库结合起来，而且将数据库技术与 CAD（Computer Aided Design——计算机辅助设计）技术结合起来，改变了以往的图形库以单纯存贮图形（位图）为主的状况，使图形要素与其空间位置、拓扑关系和属性数据并行存贮与管理，在实现地矿勘查图件信息的空间分析方面迈开了重要的一步。

三、数字通讯网络技术与信息传输

实现地矿勘查信息的远距离传输和社会化服务，是地矿勘查工作信息化的重要目标之一。欧美发达国家的地矿勘查资料传输主要通过通讯卫星、微波和数字通讯网络，其中数字通讯网络的应用最为普遍。数字通讯网络是由计算机、光缆和中继装置组成的现代化信息传输网络，通常有局域网络和广域网络之分，广域网络又可再分为地区网络（短程网络）和国家网络（远程网络）之分。网络之间的互联，便构成全国性和全球性互联网络——Internet。它包含有通过一组通用协议（TCP/IP）和工具连接在一起的成千上万个网络，因而可以看成是全世界最大的计算机网络。Internet 把全世界的企业、个人、研究机构、学校、医院、新闻单位、政府部门，甚至一些竞争性商用网络连接在一起，实际上已经成为个人、政府、企业、研究和教育机构极其重要的数据通信工具。Internet 允许用户通过 FTP（文件传送协议）交换文件和信息，存取大量的数据、程序、图形、图像和其他资源，可以大大地增强各地人员之间的研讨与协作，为地矿勘查信息的远距离传输和社会化服务提供了便利条件。

美国地质调查所（USGS）于 1985 年建立了全国范围的远程通讯网络 Geonet，通过这个网络将其分布在各地的多个局部网络和多个计算机系统联接起来，并实现了声音与数据同时传输的功能。各个局部网络允许使用不同的网络技术与传输速率以满足不同用户的需要，并在本地区或单位内部共享资源和信息。在网络上运行的应用程序包括计算机图形系统、数据管理系统、统计分析系统和数据输入系统等。Geonet 还负责整个美国内务部的远程通讯，它将 USGS 的数据库系统与渔业、野生动物保护单位、国家公园管理局、矿产资源局、矿业局和土地管理办公室等的数据库系统相联接。该网络的使用率每月达数千人次，发生数十万次计算机对话，传输数十亿个字符的数据。从 1988 年起，USGS 的 Geonet

与全美网络和国际网络联接，进一步实现了与 1500 所美国及其他国家的研究中心通讯，同时可以方便地使用联络于网络上的国家超级计算机资源。澳大利亚地质调查所（AGSO）根据不同的机型和不同需求采用不同的通讯网络，其中，PC 机用 Novell 网，苹果机用 Applelk 网，VAX 小型机用 Decnet 网，设置多路转换器、中继器和桥，并通过以太网和 TCP/IP 协议将这些网络联接起来，共享数据资源和软件资源。澳大利亚地质调查所也安装了既能通话又能传输数据的专用自动交换机，还与国际互联网络 Internet 联接。大量的数据库通过网络向广大用户提供信息服务，不但推动了地矿勘查工作的发展，还促成了一个崭新的资源信息产业的形成。在西方经济发达的国家中，各种地矿勘查数据和图件资料均已商业化，走向了市场，并通过社会化服务取得了显著的经济效益。

目前，我国各行业、各部门、各单位都在加紧建立自己的局域网络、地区网络、国家网络以及国际互联网络（Internet）。至 2003 年 12 月份，全国各省、区已有 600 多个城市的数万家单位和数千万家庭参与国际互联网络，近 3，000 万台微型计算机入网运行，7，800 万人利用 Internet 进行信息交流，国际出入口带宽已达 20G。我国各级地矿勘查与管理机构，也都纷纷利用 Internet 的电子邮件、远程协作、信息采集和提供等三个方面功能，来传输地矿勘查数据、发布有关信息、监测资源与环境的动态变化、研究资源管理问题和制定有关管理与开发决策方案，获得了显著的社会效益和经济效益。中国地质调查局国际互联网网站 2000 年 1 月 1 日正式开通，通过网络系统，向社会提供地质调查成果资料目录、地质文献资料和全国地层数据库等方面的信息服务。到 2002 年 12 月，地质调查局网站已运行 36 个月，已超过 180，000 人次访问或下载信息。当然，使用 Internet 进行远距离的资源信息传输，需要针对一系列国家机密采取有效的安全措施问题。总之，随着国家信息高速公路和通讯网络建设的加速进行，信息传输技术将会有一个大的发展，我国地矿勘查信息传输方式落后的面貌正在迅速改观。

第二节　地矿资源数据处理与资源预测评价技术

相关技术大致包括：勘查数据分析与资源统计预测技术、勘查图件机助编制与多维图示技术、地质-成矿过程计算机模拟技术、资源预测与勘查评价的人工智能技术等。

一、资源勘查数据分析与资源统计预测技术

勘查数据分析与资源统计预测是利用电子计算机的快速运算功能，来实现各种数学模型的解算，达到压制干扰、突出有用信息的目的，并且对有效信息进行各种统计、分析和综合。其内容包括物探方法模型的正、反演计算、化探及地质编录数据的统计分析、矿产资源的定量预测、储量的计算与统计、工程岩土力学和水力学计算等，此外，还包括大量日常工作的数据换算。地矿勘查数据分析是资源行业最早应用计算机技术的领域。随着以地质数据多元统计分析为基础的数学地质的迅速发展，以及矿床统计预测的理论和方法的提出和完善（赵鹏大等，1983，1994；侯景儒等，1993），这方面的技术日趋成熟。

由于计算机硬件和软件的迅速发展，发达国家勘查数据的计算机分析技术自 20 世纪 80 年代起，已从单纯的运算、统计扩展为地质资料的综合解释推断。国外在石油地质领域中所开发的许多软件，都属于这种类型，其中以美国的 GeoQuest 和 Landmark 最为著名。这些软件能够对二维、三维的地震勘探、钻探和测井数据进行构造、沉积和含油气性

的综合解释，能够辅助人们完成勘探研究、井位确定、成藏分析和油藏描述等项任务。国内勘查数据分析方面的大型软件研制较少，长期以来都是依靠进口，与先进国家的差距比较大。由我国地调局和信息院联合组织开发的 KPX 系统，是一个较大型的固体矿产勘查软件系统，具有很强的日常勘查数据处理及储量计算能力，在标准化、系统化和商品化方面都有大的改进，推广应用前景很好。国内用于日常地质勘查数据分析的中、小型软件主要靠自己开发，创意和技巧水平不比国外差，但通用性较差。

矿床统计预测是一项“把区域地质、成矿规律、地质数学信息和概念加以综合，并用计算机进行综合评价的工作”(Agerterbeg，1974)。根据预测的目的、范围、比例尺、内容、方式及成果的差异，矿床统计预测可分为矿产资源总量预测和潜力评价、成矿远景区定量预测、矿床及矿体定量预测。赵鹏大教授等(1981，1984)经过多年的实践和探索，总结出矿床统计预测的基本理论、准则和方法。

矿床统计预测的基本理论包括：①相似类比理论；②求异理论；③定量组合控矿理论。这 3 方面理论相互依存、相互补充、缺一不可。相似类比理论是矿床预测的基础，主张详细了解和大量占有国内外已知各类矿床的成矿条件、矿床特征和找矿标志，然后通过成矿地质环境的对比，在地壳范围内预测和选择最可能成矿环境，或者在给定的地段内，判断可能寻找和预测的矿产。求异理论是成矿预测的核心，强调在相似类比的基础上注意发现不同层次、不同尺度水平、不同类型的异常，进而根据成矿背景场和地质、物探、化探及遥感等异常的分析，在确定的有利成矿环境或地段内进行预测靶区的选择。定量组合控矿理论是成矿预测的依据，要求把握一切与矿床有成因联系的地质、化学和生物作用，掌握一切与成矿有关的因素及其表征，并通过成矿概率大小和成矿优劣程度的分析，在圈定的成矿远景区中评价和优选最可能成矿地段或可能成矿的最佳地段。

目前，矿床统计预测方法很多(李德威，1995)。用于矿产资源总量估计和潜力评价的方法有：单元区域价值估计法、丰度估计法、体积估计法、矿床模型估计法、德尔非法、特征分析法、成因地质模型法、“镶嵌模型”估计矿床产出概率法、矿床值概率分布法、综合法，等等。用于成矿远景区定量预测的方法有：找矿信息量分析法、回归分析法、判别分析法、聚类分析法、逻辑分析法、秩相关分析法、欧氏距离分析法、条件概率分析法，等等。用于矿床和矿体定量预测常用的数学地质方法，除了成矿远景区定量预测常用的几种外，还有趋势面分析法、最优分割法、阶梯函数分析法、因子分析法、对应分析法、对应聚类分析法、马尔科夫过程分析法、特征分析法、矿床模型法、不确定模型法，特别是我国学者所提出的基于求异理论的无模型预测法(朱章森等，1991)和基于地质、物探、化探、遥感资料的综合信息法(Wang Shicheng 等，1992)，等等。这方面的软件品种多样且数量巨大，许多程序已经在国内外公开发表，基本上可满足广大用户使用。

地质统计学方法是一种十分重要的空间数学地质方法。这种方法在充分考虑观测点空间关系的基础上，应用统计学的各种基本原理，提出了一些用于对空间对象(地质特征)及其变化进行描述和估计的统计工具，如变差函数、体积-方差关系、克里格法等。地质统计学在矿产勘探、储量计算、资源预测、矿山生产、地下水勘查、工程地质勘察、地球化学调查、地球物理勘探以及环境保护领域中，得到了广泛的应用(侯景儒等，1990，1993)。当前地质统计学的一个重要发展方向，是与空间信息的存贮、管理和提取技

术——地理信息系统（GIS）结合起来。这两种空间信息技术的结合，将使空间信息在图形和数据之间方便地来回转换，可能促使地质与矿产资源空间信息的分析和区域评价技术产生质的飞跃。

近些年来，灰色系统理论（邓聚龙，1982）、模糊数学、分形几何学和模式识别方法等也被逐步引进，成矿过程的非线性理论的研究也已经起步（赵鹏大等，1998；於崇文等，1998），同时，数据库系统（DBS）、地理信息系统（GIS）、遥感（RS）技术、专家系统（ES）和人工神经网络系统（ANNS）技术也得到了广泛的应用。这一切大大地丰富了矿产资源的定量预测的理论和方法。

二、资源勘查图件机助编制与多维图示技术

应用计算机图形技术来编制和显示资源图件，既能保证质量、减少编图、制图和修编的工序和时间，还可以实现图形数据共享，方便图形的存贮、保管和使用。西方发达国家经多年的探索，到20世纪80年代中期，都在不同程度上实现计算机辅助编制和出版正式资源图件，并且从对旧图的修编与再版，发展到在实际填（编）图项目中的一次性成图。

我国地矿资源勘查图件计算机辅助编制技术的研究开始于20世纪80年代初期，到20世纪80年代末期已取得可喜成绩。中国地质大学开发的GeoCAD采用统一（行业标准和部颁标准）的图式图例和标准（国家标准）的术语代码，而且与数据库系统相结合，基本上实现了各种勘查图件的计算机辅助编绘。至于根据野外地质调查数据，一次性生成彩色区域地质图的计算机辅助设计，国内外尚无成熟的软件系统进入市场。中国地质大学所开发的计算机辅助地质调查系统（GeoSurvey），是在这方面的成功尝试。

地矿资源勘查图件机助编制技术的发展方向，一是与三维图示技术结合，实现地矿资源数据资料的立体可视化表达；二是把人机交互方式和人工智能方式结合起来；三是使系统走向集成化。

可视化就是通过对研究目标内部机理的观察和建模，将检测或获得的数据转化成几何量，然后利用计算机形成可视的图形或图像，让研究者直接观察到计算/仿真的过程和结果，以便直观地认识和揭示深层信息的一种高新技术。可视化技术已经涉及的领域包括分子模型、医疗图像、空间开发、天文观察、计算流体动力学、高能物理、有限元分析和各种虚拟现实等。

计算机多维图示技术能将庞杂的地质数据形象化、立体化、动态化，大大地提高了它们的可视性和推断解释的准确性（Mark，1992；Bruce，1996）。目前国外三维数字地理模型静态图示软件已经商品化，三维的多边形域柱状图编绘技术趋于成熟，任意形状的复杂地质模型立体图示技术也已经取得了重大进展。多维图示技术与数据库、GIS、地矿勘查数据分析及地质过程模拟技术相结合，产生了计算机地质动画技术。只要操作人员所占有的资料数据充分、分析模型可靠，便可以跨越时空，按照各种地质现象发生的“时序”，形象地再现地质体和地质作用的发生、发展历程。由此，地质工作者能够“透视”地下地质体，直观地理解地质作用、成矿作用和地质灾害的发生发展过程，动态地监视滑坡、塌陷等地质灾害和地下水的变化。这种地质模型的多维动态图示技术，也称为“图形诊断技术”，已经开始用于油气藏描述、盆地模拟和工程基础岩体描述。

三、地质过程计算机模拟与资源评价技术

地质（包括成矿）过程计算机模拟也称为地质过程数学模拟，它是近20年来在计算

机地质应用领域里迅速发展着的一种仿真技术。地质过程所涉及的空间之大、时间之长、影响因素之多、相互作用之复杂，是难以用实物的物理模拟方法和化学模拟方法来再现的。这个问题长期以来困惑着地质学家，既阻碍了人类深入、全面地认识各种地质过程，也阻碍了地质科学的定量化进程。自从计算机技术引进地学领域后，地质学家们就开始了地质过程的计算机模拟研究，并迅速地取得了进展。

地质过程计算机模拟的一般方法是：首先通过地质研究来建立对象的地质过程和地质特征的概念模型，再选择适当的数学模型来描述其中的主要过程，然后让计算机按一定的时序和法则来执行数值运算和逻辑推理（Harbaugh et al.，1980）。所采用的数学模型可以是准确的数学函数表达式，也可以是概率性的、经验性的甚至逻辑启发性的关系表达式，地质过程数学模拟随之可分为静态确定型（如构造应力场模拟）、动态确定型（如地下水动力学模拟）、动态随机型（如某些沉积作用模拟）和动态混合（确定＋随机）型，如油气成藏动力学模拟）和人工智能型。其中动态模型可以看作一系列静态型的有序组合，而确定性模型可以看作随机模型的特例。这些模型通过演绎和推理，可以随着输入数据的增加而提高仿真程度，又可以按离散的时间增量前进而实现动态效果，还可以及时地将结果同实际数据进行对比，不断地修改模型以适应实际情况。当所归纳的地质过程在计算机内持续的“时间”足够“长”，便可以产生与实际地质过程相似的数字化“结果”。地质工作者可以将这种从研究对象中抽取的概念模型及其相应的数学模型看作实验工具，通过改变各种条件和参数来观察它的反应，从而定量地揭示各种地质事件中主要影响因素的相互关系，以及变化趋势和可能结果，达到理解和掌握地质过程的规律性和特殊性，进而实现工程地质条件和矿产资源的预测、评价，以及地质灾害的预报目的。

计算机模拟最早是从与地下水有关的研究开始的。美、日、法、英等国在十几年前就已经广泛地应用计算机模拟，来解决地下水资源管理问题，进行储量评价和预测。然而，计算机模拟在近年来发展得最为迅猛的领域要算石油天然气勘查领域。这就是人们常说的盆地模拟（Basin Simulation）。盆地模拟从研究盆地的构造演化、沉积演化和地热演化入手，分析岩层分散有机质的降解和烃类的生成、排放、运移、聚集过程的影响因素及控制条件，建立其概念模型和数学模型；然后利用计算机来再现这一过程，达到预测、评价盆地和富集带油气潜力的目的。由于这一方法在预测盆地油气潜力、提高勘探命中率、降低投资风险方面有显著效果，从20世纪80年代初期以来得到广泛的重视（Leonard，1989）。

与此同时，在地学各个领域，各种地质作用计算机模拟的研究也在蓬勃地开展着，大至板块相互作用、造山运动、造盆运动、古气候演化和古生物演化的模拟，小至岩层褶皱作用、断裂作用、成矿作用、地质灾害的酝酿、水库和渠道的渗漏和淤填、坝基和边坡的稳定性等的模拟，都取得了显著进展。多媒体和虚拟现实技术的采用，使地质过程数学模拟有了良好的声像表现形式。1989年美国能源部曾利用Cray-1巨型机，完成了几百秒钟的火山喷发过程模拟，十分逼真、生动，使人如同亲临其境，可以深化人们的科学认识。

地质过程计算机模拟的发展方向是：加强研究对象的实际地质过程分析，提高概念模型的相似性；注重地质作用过程各影响因素的相互作用分析，尤其注意揭示其中的反馈控制关系；考虑实际地质过程的非线性特征和地质体的分形特征，开展非线性动力学和分形模拟研究；与数据库及GIS密切结合，充分利用地矿勘查中所获取的庞大空间和属性信

息；与人工智能模拟相结合，吸取专家的知识与经验，体现勘查人员的分析思路与方法；与三维动态图示技术结合，将动力学模拟与非动力学模拟结合起来，向整体的三维化、动态化发展。

四、人工智能技术在资源预测与勘查评价中的应用

人工智能技术包括专家系统（ES: Expert System）与人工神经网络系统（ANNS: Artificial Nerve Network System）两种类型。资源的预测评价的专家系统，是一种计算机程序，它由“知识库”和“推理机”两部分组成。在知识库内凝聚有相当数量的权威性知识，并能根据用户提供的信息，运用所存储的专家知识，通过推理机以专家水平或接近专家的水平来解决特定领域中的实际问题。地质学的很多问题如大陆漂移、岩石的破裂、矿床的形成等等，都不是单纯的数学推理能够解决的，定性分析一直是认识地质作用、预测评价资源的主要方式。地质学又是实践性很强的科学，到自然界去调查研究、在实践的基础上进行分析、对比、归纳、分类，将今论古，根据事物的已知规律去推断未知，是地质学的最基本的研究方法。许多复杂地质问题的解释和处理，在很大程度上依赖专家的知识和经验。

世界上第一个地学领域的专家系统 prospector，是 1974～1983 年期间由斯坦福国际研究所与美国地调所协作开发与完善的，主要用于固体矿产资源评价。应用该系统对华盛顿州的一个老的钼矿进行研究，发现了新的矿体，增加了储量。国内地学领域专家系统的研究是从 20 世纪 80 年代初期开始的，先后涌现出物探资料推断解释、油气资源评价、古沉积环境分析、金属矿产评价、地震资料解释、水资源寻找等专家系统。较为著名的有胜利油田的石油测井解释专家系统、海洋石油勘探开发研究中心的油气资源评价专家系统（王伟元等，1993），吉林大学的 AMRB4 航空物探专家系统、浙江大学的钨矿预测评价专家系统、中国地质大学（武汉）的大比例尺矿床统计预测专家系统（MILASP）和长春地质学院的金矿综合信息预测专家系统。MILASP 系统是根据地质异常理论和矿床统计预测理论、方法设计的，它集多源信息输入、数据处理、图形和图像处理、表格处理和专家系统功能于一体，主要用于大、中比例尺矿床统计预测（李新中等，1995）。该系统在川陕交界处的多金属矿床预测中，取得了显著效果。

目前，专家系统有两个明显的发展趋势：其一是与数据库、GIS 及图示技术相结合，尽量避免仅依靠少量数据进行简单逻辑推理、判断，应通过大量占有资料、分析资料和综合资料来提高解释、推断能力；其二是采用人工神经网络（ANN: Artificial Neural Network）技术，研制新型专家系统——神经网络专家系统。人工神经网络（ANN）仿照人脑的结构和机理，由大量的简单元件（神经元）按一定的模式和规则（算法）广泛相互连接而成。在信息处理过程中，可以模拟人脑按神经元之间的连接强度，对信息作放大或缩小处理，因而能够模仿人的神经活动，进行信息传输和从研究实例中自动获取知识。其神经元之间的连接强度系数，通常由 ANN 的学习规则（算法）决定。

利用 ANN 进行资源预测评价，不要求有假设条件，也不要求事先搞清楚控矿条件和找矿标志与资源体及资源量之间的定量关系，而只要求提供足够的学习样本（实例）。ANN 通过学习，自己会找出标志和条件与资源之间的定量关系（且通常是非线性的），并且以隐含方式存贮于网络权阵中。经训练后的 ANN 可用于资源远景区的识别和资源量预测，但目前这方面的报道还很少。Morgan（1990）曾利用 BP 网络对美国怀俄明州香农盆

地的油气资源作了预测评价试验，所得的结果与实际情况相符合。克拉里西等（1995）也曾利用 BP 模型来进行矿床的可采储量计算试验，结果与采用普通克立格法和条件模拟法可以对比，并且估算精度随着训练次数的增加而提高。ANN 还可以用于解决复杂的组合优化和多目标决策问题，例如，地矿勘查、开发和管理中的勘查技术方法和手段的组合优化决策问题、最优勘探方案的选择、资源配置与合理利用问题、勘查投资结构的优化及投资风险的评估决策等。

第三节　地理信息系统技术在资源勘查中的应用

地理信息系统（GIS）是一种特定而又十分重要的空间信息系统，它是采集、存贮、管理、分析和描述整个或部分地球表面（包括大气层在内）与空间和地理分布有关的空间信息系统（李德仁等，1995）。

地理信息系统的核心是空间图形数据库，而关键技术是建立图形元素的空间拓扑关系和图形与属性的联动关系（李裕伟，1998）。这些关键技术保证了对空间图形信息能够进行有效管理和操作，以及在对图形单元进行任何复杂操作的情况下，都能得到与该图形单元对应的属性表。空间图形对象与对象属性的完美结合，是地理信息系统实现空间分析的奥妙所在。地理信息系统的数据库可以存贮大量的图层，而地理信息系统的空间分析技术可以实现图层的分解与合并。在每个图层中存放着逻辑同一、类型一致的空间对象。

例如，一些具有共同边界的线条（断层、地质界线等）和区域（地层、岩体、变质带等）信息，可以存放在同一图层中；而一张区域地质图可以根据需要分解出断层图层、沉积岩图层、岩浆岩图层、变质带图层、产状符号图层、矿产图层，等等。每个图层的空间对象（如每条断层、每个地质单元、每个矿床或矿点等）都可存入相应的属性数据。当人们对空间图形对象进行操作时，属性数据也会随之联动。于是，一张复杂的地质图变成了若干相对简单、有序，而又相互紧密联系的图层，这就相当于把一张地质图拆分成了许多构件。利用这些构件，既可以恢复原图，又可以按照地质学家的愿望组合成其他专题图件；既可以利用地质图上的空间对象来检索它们的属性，又可以根据属性来检索相应的空间对象。这样一来，对地质图上的空间信息的检索、修改、拆分、合并、分析、综合就变得非常容易了，地质图不再是一张复杂的、内容拥挤不堪的、固定不变的平面纸图，而是一批由计算机管理的、逻辑上、物理上有序组织的、能按照地质学家的愿望迅速组合的空间地质信息。

GIS 能够使研究人员以一种全新的、综合的方法，来解决地质领域中的各种空间信息的分析问题，近年来被广泛地用于基础地质、矿产地质和环境地质的勘查与研究中（Reddy R. K. T et al.，1990；Moore，1991；Franz and Coy，1996），取得了一系列令人瞩目的成果。

随着 GIS 本身的技术发展，其在地矿资源勘查中的应用已深入到基础地质分析、地学断面研究、勘查图件辅助编绘、三维沉积盆地分析、三维冰川沉积分析、沙漠表面特征研究、地下水动力学模拟、石油化探数据分析、矿产资源综合预测、环境管理决策、地震灾害评价和风灾水灾监测等多个地学领域。地质学家们可以利用数字技术来检验、完善已知的模型，并且建立新模型。在太平洋沿岸各城市和地区，人们应用 GIS 将底图数据（水

文、交通、行政界限、地质数据)、地质技术数据（垂直速度，孔隙比，土壤类型)、地面响应数据、坡度数据、地震数据（震源、强度）以及用卫星和航测图像生成的有关图件信息进行综合分析，生成各种图件用以指导减轻地震灾害。弗吉尼亚州地质调查局用GIS评价了威廉普县三叠纪盆地中活性有机化合物的污染。根据该三叠纪盆地的水文地质资料、有毒物质清单（有毒废料的产生者和使用者)、三叠纪盆地地貌等资料，结合从州的水资源利用系统和全国水质采样点数据系统检索的数据，研究了污染的性质和蔓延情况，查明了影响地下水移动的水文地质条件，提供该州进行城市供水井、工业用地、有害废料处理的科学选址。

GIS在我国地矿资源勘查与预测评价中的应用，至今仍然处于探索阶段。国土资源部岩溶所最早于1984年应用于研究和评价区域岩溶分布特征，随后建立了岩溶塌陷评估预测信息系统。中国地质大学（武汉）等利用GIS进行金属成矿预测、水文地质条件评价与水资源系统分析，中科院地质所利用GIS研究花岗岩区新构造，中国矿业大学利用GIS进行煤矿突水预测（张大顺等，1994)，成都理工学院利用GIS评价地面岩体的稳定性，国家地震局利用GIS进行首都圈震害环境质量评价，河北地矿局利用Mapinfo进行榆关镇幅1∶5万填图（董国臣等，1997）等等，都是这方面的探索性成果。

在基础软件开发方面，到20世纪90年代中期，我国已经开发出了一批具国际先进水平的工具软件，如中国地质大学的MapGIS和武汉测绘科技大学的GeoStar等。

由于GIS至今尚缺乏三维及多维数据的管理和处理能力，缺乏强大的关系数据库管理系统和适合于地矿勘查数据分析的各种方法库与模型库，面对地矿勘查的多源、多类、多量、多维和多主题数据，显得能力不足。国外GIS的发展趋势，是逐步走向集成化、智能化（Miller，1991；李德仁，1995)、三维化（Hughes，1996）和可视化。目前在地矿勘查和预测评价中应用GIS，通常是把重点放在空间数据管理与处理功能的二次开发方面，同时在可能的条件下开展“多S”(GPS、RS、DBS、GIS、CADS、MIS、DSS和ES等等)的结合与集成化技术研究。

思　考　题

1. 如何从地矿勘查信息化的角度来理解“数字地球”的概念?
2. 信息技术在地矿勘查与地学研究中的作用是什么?
3. 地矿勘查数据管理技术主要体现在哪几个方面?
4. 从地矿勘查数据处理技术的发展现状和趋势中可以得到什么启示?
5. 地质过程数学模拟的本质是什么？如何才能使模拟结果符合实际?
6. 试比较专家系统方法与人工神经网络方法的差异。
7. 什么叫地理信息系统？它在地矿勘查与地学研究领域的应用价值如何?

第二章　地矿资源勘查信息系统的开发方法

地矿资源勘查信息系统的建设是信息技术在地矿勘查、开发与管理领域应用的基础，因而也是地矿勘查与管理工作信息化的核心问题。为了使大家对地矿资源勘查信息系统有一个全面的认识，在学习地矿资源勘查信息系统的开发方法之前，有必要首先了解地矿资源勘查信息系统及其相关学科的基本概念。

第一节　地矿资源勘查信息系统的基本概念

在地质矿产资源的勘查过程中，人们每日每时都在获取着大量的资料和数据。现代地矿勘查工作所面临的将是一个“ 数据海洋 ”。怎样管理好这些数据并迅速、有效地经常利用这些数据去解决各种复杂的地质问题，开展有效的地质评价和资源预测，是摆在我们面前的一个重要任务。根据国内外的成功经验，完成这一任务的最优途径，是采用“ 多 S ”结合与集成化技术来建立完善的地矿勘查区点源信息系统，简称为地矿资源勘查信息系统(RIS) 。为了便于介绍和理解，我们将这个概念分为信息、信息系统、地矿资源勘查信息系统三个层次。对于任何概念，各个领域的学者通常是从各自学科的角度出发来下定义的，本书也将从地矿勘查学和地矿资源信息科学的角度来定义这些概念。

一、信息的概念

信息是一个复杂的综合体。由于专业领域的差异和理解上的差异，人们在许多场合中常把数据、知识与信息等同看待，结果使初学者无所适从。为着更好地进行信息的采集、管理、处理和应用，有必要对信息、数据、知识等概念作明确的区分（李之棠等，1996) 。

1. 数据

数据是客观事物（包括概念）的数量、时空位置及其相互关系的抽象表示。它可以是单个的符号、数字、字母、文字和词语，也可以是它们以某种形式和规则的集合，例如一个数组、一段文字、一句话、一篇文章或者是一幅图。总之，一切能为人感知的抽象表示都可以是数据。前面所说的在国土地矿调查过程中获得的 7 类数据，即：地球物理勘探与遥感数据、地球化学勘探数据、野外地质观测数据、室内化验测试数据、地形地物的三角测量数据、综合整理与图件数据，就是地矿资源勘查信息系统将要存储、管理和处理的数据。数据和数据载体是两种不同的概念，数据是逻辑概念，而载体是物质概念。载体有时又称为媒体、媒介或介质。一批数据可以记录在多种媒体上，同样，一种媒体也可以记录多种不同的数据。

2. 信息

信息是数据的含义或约定，表示事物运动状态和存在方式。数据是信息的载体，信息寓于数据之中。只有准确地表达了数据的真正含义的信息，才是完整的和有价值的信息。

例如，在一个地区出现的重力异常，可能是岩石圈结构异常特征的反映，也可能是地壳深部结构异常特征的反映，还可能是在当地的地壳浅部存在某种矿床的反映。如果我们无法用另外的方法或从另外的地方进一步查清它们的真实含义，那么我们在实际上并没有得到完整的信息，其价值就很有限。从这个意义上讲，信息需要通过对数据的分析和解译来获取。而各种实物媒体是数据的物质载体，就像多波段遥感数据是地貌、植被、水体和某些地下地质信息的逻辑载体，而磁带（或卫星照片）是多波段遥感数据的物质载体。正因为三者密不可分，在特定情况下，人们也常将数据和信息甚至数据载体当作同义词看待。

一般地说，数据结构越复杂，所表达的信息量越大。为了便于计算机对信息的存贮、管理和处理，可以将数据分解成为一组属性及其属性值，即（属性 1：值 1，属性 2：值 2，……属性 n：值 n）。这种数组形式可以完整地描述一件事情、一个物体、一种现象、一条信息或一个概念（统称为对象）的存在状态和行为方式（统称为属性，例如时间、地点、程度等）。例如“辽宁抚顺盆地是早第三纪的断陷盆地，其含煤岩系出露面积 60 km^2”，这一信息可表达为“地点：辽宁省；盆地名称：抚顺盆地；形成时代：早第三纪；盆地性质：断陷盆地；盖层类型：含煤岩系；出露面积：60 km^2”，也可以表达为表 2-1 所示的二维数表。

表 2-1　辽宁省抚顺盆地信息二维数表

地点	盆地名称	形成时代	盆地性质	盖层类型	出露面积
辽宁省	抚顺盆地	早第三纪	断陷盆地	含煤岩系	60 km^2

3. 知识

知识是信息的集合，是通过多个信息的关联和组合而表达的认识和经验，它来自于人类改造客观世界的实践中。例如，“上盘下降下盘上升”，“正断层”，“SiO_2 含量高于 40%”，“岩浆岩”，“花岗岩”等分别是一些孤立的信息或原子事实。如果我们用表示因果关系的关联词“如果……则……”把其中两个或两个以上的孤立信息关联起来，就构成了一条知识。例如，“如果断层的上盘下降而下盘上升，则为正断层”；“如果岩浆岩的 SiO_2 含量高于 40%，则为花岗岩”。这两条都是简单的地质学知识。

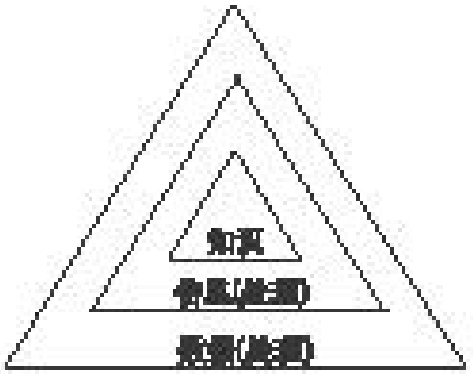


图 2-1　信息处理的层次
（据李之棠等，1996；有修改）

根据以上定义和论述可知，数据、信息、知识之间存在着明显的层次关系（图 2-1）。如果要分别对它们加以处理，则对应的处理也应当构成这种包含的层次关系。知识（处理）依赖于信息（处理），而信息（处理）依赖于数据（处理）。随着由下往上层次的上升，需要存贮和处理的对象越来越多，也越来越复杂。

地矿信息是自然过程和人类在地矿勘查、研究、开发、利用和管理过程中各种状态的客观显示，也是人和自然资源在相互作用的过程中所交换的内容。它们有时表现为物质形态，有时表现为非物质形态，既反映了这些事物在运动中的各种差异及规律，又反映了这些事物之间的相互作用和相互联系。可以认为，信息在把地质体性质、特征及其形成、分布、演化规律转化为人类意识的过程中，

甚至在人类社会与大自然的相互联系、相互作用和协调发展过程中，始终起着中介作用。可靠而且健全的地矿信息，可以消除人类在地矿开发利用方面对社会可持续发展问题认识的不确定性——导致由人类和自然界所组成的人-地系统的有序性增加，即负熵增加。当然，失真而且残缺的地矿信息，必然增加人类对社会可持续发展问题认识的不确定性——导致由人类和自然界所组成的人-地系统的有序性减少，即熵增加。现代信息资源管理的主要标志之一，就是以计算机为基础的各种信息系统的建立。信息资源计算机管理技术的发展和应用，不仅大大地改善了信息工作的条件，有力地推动信息工作向产业化方向发展，而且能够促进数据→信息→知识的快速转化，提高知识创新和技术创新能力。

二、信息系统的概念

信息系统是计算机技术和信息资源管理相结合的产物，由计算机硬件、软件、信息、方法和人组成，是对信息进行采集、整理、存贮、管理并提供查询、检索和处理功能的一种综合性技术系统。信息系统包括办公自动化系统、数据处理系统、信息检索系统、管理信息系统、决策支持系统（含专家系统）、通讯网络，等等。所谓信息资源管理，就是运用管理科学的一般原理和方法，对信息活动中的各种要素（信息、人、技术、设备、机构等）进行科学的规划、组织、协调和控制，以确保信息资源的充分开发和合理利用，从而有效地满足社会的信息需求的过程（胡继武，1995）。

1. 信息系统工程学的发展

研究、开发和建设信息系统的学科，称为信息系统工程学。这是20世纪70年代末、80年代初在国外伴随着计算机技术的发展而兴起的信息科学分支。信息资源管理学、信息系统和信息系统工程学的兴起，有其深刻的背景。首先是信息经济的迅速崛起。信息作为一种资源，使生产力的要素更加丰富，而作为一种经济商品，又使信息交换与信息服务行业蓬勃兴起。其次是信息技术的高速发展。现代信息技术的高速发展，使信息存贮介质的密度、信息处理和利用的效率和信息传递速度越来越高，有力地提高了信息资源管理的效能。计算机技术和远程通讯技术的结合，又推动了信息网络技术的发展，特别是全国性和全球性计算机互联网络的建成，使信息的远距离大规模传输和共享成为可能。

总之，科学和技术的发展及其所带来的各种问题，要求对信息采取全新的管理思想和管理方式，从而促进了信息资源管理学、信息系统和信息系统工程学的产生和发展，进而推动信息资源管理学的形成。

信息系统结构功能的发展水平取决于内在的数据环境。一般信息系统的数据环境可分为数据文件、应用数据库、主题数据库和信息检索系统（James，1977）等4种。4种数据环境与各行业计算机应用及地矿资源勘查信息系统建设的几个阶段相对应（表2-2），反映了有关信息技术水平和理论方法体系的发展过程。从目前的情况看，我国各行业多数信息系统的数据环境，基本上属于应用数据库范围。这类数据库的建设方式大多数是简单地利用现成的商业化软件来装载数据，其显著特点是：①以功能处理为核心，功能软件为基础，设计依据是某个企业、业务管理部门或研究单位的当前需求——为了编制某些专用图表、解决某些专门问题、实现某些功能处理、分析某些市场规律或编制某些工作设计与研究报告等；②只从主管部门的决策需要出发，较少顾及基层生产单位和研究单位的当前和未来需求，或者只是为了快速组建信息网络而人为地规定入库内容，数据项和数据模式不是经过深入的系统分析来选择和制定的，导致许多有用信息丢失；③缺乏统一的概念模

型、数据模型、数据标准和数据代码，缺乏统一的硬、软件平台和接口。

表 2-2　信息系统建设的发展阶段划分

Norain 模　型		初级阶段		中级阶段		高级阶段	
		起　步	蔓　延	控　制	集　成	优　化	成　熟
Jemes　环境		数据文件	应用数据库		主题数据库		信息检索
技　术 特　征	核心问题	功能处理			数据管理		
	追求目标	数据处理的功能			数据资源的利用价值		
通讯方式		无		邮寄磁介质	局部网络	远程网络	

发展时序 20 世纪							
	北　美	50 年代	60 年代	70 年代	80 年代	90 年代	以后
	中　国	70 年代	80 年代	80 年代	90 年代	90 年代	以后

应用数据库和图形库优点是系统简洁，易于掌握和应用；缺点是无法实现多用户交叉访问，构成了一个个“信息孤岛”，信息既不完整又有冗余，许多数据和图件资料被重复存贮、重复加工，无法实现交叉访问，不能支持未来的再开发、再提高，难以满足迅速增长的信息处理要求，也难以被接纳到地矿资源勘查信息系统网络中去。西方发达国家在 20 世纪 70～80 年代曾采用“系统集成”的方式来将“信息孤岛”联接成“信息大陆”，但修改管理软件、重新组织数据和编制接口软件所耗费的成本，大大超过重建系统（张颖，1993）。

今后的地矿工作所面对的，将是数量极大、极复杂的数据集合。这些数据是随着地矿勘查工作的进行而逐渐积累的，经常被目的不同的用户同时使用。用户对数据检索、查询、显示、处理等总是要求尽可能方便、迅速、准确。不难想像，没有一个结构合理，信息齐备的点源信息系统，既难以支持日常的检索、查询和处理功能，也无法建立区域性或全国性的综合信息系统，即使建立了也无法实现综合数据的动态管理并保证信息的准确性和完整性。解决这些问题的途径是采用主题数据库（Subject Data Bases）的设计思路与方法，即不是以功能处理为核心，而是以数据管理为核心；实行数据模式标准化、代码标准化与图式图例标准化，兼顾各自行业的当前需求与未来需求；通过严格的数据分析和模型构筑来形成与各种业务主题相关联的点源数据库、图形库，进而通过网络联结发展成为完善的信息检索系统。主题式地矿点源数据库是地矿资源勘查信息系统的基础。

2. 信息资源管理的基本原则

信息资源管理是一个新的管理领域，是关系到国家和社会组织在争夺信息资源的国际竞争中能否处于有利地位的关键。由于信息资源涉及信息、设备、人、政策、法律、技术、经济等多种因素，并且广泛渗透到社会政治、经济、军事、文化及科技等领域，所以它是一项十分复杂的管理活动。搞好信息资源管理必须遵循一定的原则，才能维护主体的最高利益并符合信息运动的客观规律。一般说来，信息管理必须遵循以下基本原则，这也是信息系统建设所要遵循的基本原则。

(1) 共享原则

这是信息资源管理的目标和归宿。国土资源信息是国家的宝贵财富，理应为社会各阶层、各单位和各人民团体所共同利用。实践结果表明，信息利用越广泛，其“资源”作

用发挥就越充分。实现资源共享，是时代和科学发展的必然。同时，信息交流活动本身的社会化和集约化，也要求使全社会的信息资源实现共享。根据共享原则，必须开展信息系统建设，建立完备的社会化的信息资源保障体系和高效率的信息流通、传播及利用体系，并且通过有效的管理，保证信息资源能为全社会众多用户最大限度地利用。当然，共享性不等于效益性，推动信息资源共享的机制是市场。只有当信息及其产品被作为信息商品推向市场之后，信息共享原则才能得到完全的实现。

(2) 系统原则

按照系统论的整体效应观点，整体大于部分之和。信息资源管理系统的整体功能不同于各个组成部分功能的简单相加。系统的规模越大，结构越复杂、越合理，它所具有的超过个体性能之和的性能就越多。换句话说，应当使国土资源的信息资源，包括各个相关行业、各个方面、各种类型以及以各种渠道所获得的信息资源，按照系统科学的要求形成一个相互联系、相互作用的系统，才能真正发挥出“资源”的作用。要做到这一点，一方面必须重视系统分析，有的放矢地进行信息系统设计和信息采集；另一方面必须打破部门之间和单位之间的相互封锁，打破条块分割、各自为政的局面。只有这样才能使信息资源管理真正做到“整体大于部分之和”。系统观点是信息资源管理不同于以往信息管理的最大特点，它将使信息管理获得新的生命和新的活力。随着社会信息化水平的提高，信息环境将会更加复杂，影响因素也会随之增多，在信息资源管理中坚持系统原则也就更加重要。

(3) 科学原则

科学原则是指信息资源管理应当遵循信息运动的客观规律，体现信息资源管理的特殊性。信息的效能和生命，就在于它能准确和真实地反映客观事物及其运动变化的特征和差异。信息资源管理要使信息服务于社会，发挥其“资源”作用，就必须要求整个信息运动过程，即从信源选择到信息的收集、处理、贮存、管理、传递、利用乃至反馈，都是严谨的和高保真的，尤其是在“信息污染”日益严重的情况下，保持信息的准确性和可靠性显得更加重要。信息又具有很强的时效性，过了一定的期限，其效用就会减少、丧失乃至出现负值。加紧研制最新的现代信息技术并且迅速地将其应用于信息管理，是解决这些问题的重要途径。这些就是信息管理的基本科学原则。

(4) 安全原则

随着信息资源共享，尤其是软资源（数据、程序）共享的推进，信息的安全问题日趋严重，成为信息资源管理所面临的重要问题。形形色色的电子犯罪触目惊心，各种信息资源都面临着被侵害、被窃取的危险，还有计算机病毒的蔓延等等，使人们日益感到不安和恐怖。信息安全问题涉及领域广泛、因素众多，单从技术上以传统的密码学为基础的计算机专门防护措施已显得力不从心，人们需要寻求新的途径、运用新的手段、从全新的角度进行综合防范和治理。信息资源管理作为信息管理的新发展，从诞生之日起就面临严重的安全问题，在信息系统建设中，应当将信息安全原则作为重要的管理原则之一，高度地关注并进行综合治理。

3. 信息系统的分类

目前各行各业所使用的信息资源管理系统，都是利用计算机技术建立的，通常称为计算机信息系统，或简称为信息系统。它们是计算机技术和信息资源管理学相结合的产物，

是由计算机硬件、管理软件及其中所包容的各种数据所组成的一种综合性技术系统。计算机硬件包括主机及各种输入输出装置（外围设备）；管理软件包括各种数据库、图形库、模型库和方法库的管理系统。由于服务对象、性质、内容以及数据本身的特点差异，随着应用领域的扩大，信息系统的类型逐步增多。目前还未见到有关信息系统分类的完善定义和标准，我们暂时根据系统服务对象、性质、内容以及数据本身的特点，大致地把信息系统分为面向数据资料、面向业务过程和面向金融经济三大类。

（1）面向数据资料的信息系统

这是以第一线工作人员和科研人员为服务对象，主要着眼于自然数据资料的收集、存贮、管理和处理的信息系统。由于面向数据资料的信息系统所存贮的是人类活动的某种“产品”本身——描述客体的特征和性质的数据资料集合，它所作的处理则是对这种“产品”进行再加工，因此从某种意义上说，面向数据资料的信息系统本身既是“原材料”与“产品”的管理工序，也是“产品”的生产工序。勘查数据处理系统（EDPS: Exploration Data Processing System）、地理信息系统（GIS）、地矿资源勘查信息系统（GMIS: Geological and Mineral Information System）和文献资料检索系统（LIS: Literature Information System）等，都属于这一类。

（2）面向业务过程的信息系统

这类信息系统以各级业务管理和决策人员为服务对象，主要着眼于“（物质或精神）产品生产过程”和“产品流通过程”的信息收集、存贮、管理和处理，为产品结构调整、产量和质量控制、品种更新、新产品设计以及日常生产管理，提供自动化工具和决策依据。目前用于业务管理的、基于计算机的信息系统有 4 类：事务处理系统（TPS: Transaction Processing Systems）、企业管理信息系统（MIS: Management Information Systems）、决策支持系统（DSS: Decision Support Systems）和办公自动化系统（OA: Office Automation）。从 TPS 到 MIS，再到 DSS，代表了面向业务过程的信息系统发展的 3 个阶段。

（3）面向金融经济的信息系统

这类信息系统通常由各国政府或集团公司按地区、按行业建立，主要用于收集、处理和汇总各地区、各行业的经济信息。金融信息系统、股市信息系统、税收信息系统和国家经济信息系统都属于此类。国家计划委员会和经济贸易委员会在 20 世纪 90 年代所建立起来的国家经济信息系统，是我国目前最大的信息系统。该系统是一种由国家级主系统与若干部门分系统组成的全国性多级信息系统，可为国民经济的统计、预测、规划、计划、调度及指挥等提供依据。从纵向上看，它分为 5 个层次：中央级、省级、中心城市级、县级和企业级。其中包含了 31 个省直辖市、几百个中心城市、几千个县和几千个骨干企业。从横向上看，除了由国家经济部门组成的主系统外，还有若干个与国民经济关系密切的部门系统，如金融、财政税收、物资、商业、外贸、能源、交通运输、农业、工业和国土资源。

三、地矿资源勘查信息系统的概念

地矿资源勘查信息系统是指建立于基层勘查单位、矿山、水利水电工程、社区、城镇（数据采集点）的基础信息系统。由于其信息源具有点源性质，因此也称地矿点源信息系统。一个完整的地矿点源信息系统，由勘查数据管理、勘查数据处理和地矿资源预测评价等 3 大部分组成。它既是基层地矿资源勘查单位对资料数据进行收集、存贮、管理、处理

和使用的综合性技术系统，也是上级主管部门（公司、省局和部委）组建信息网络系统、开展地矿信息服务的基础。对于信息源所在处或基层地矿勘查单位而言，它们是功能强劲的微型工作站；而对于国家或省局级地矿资源勘查信息系统而言，它们是数据齐备的网络结点。地矿点源信息系统的建立既可以使数据的存贮、检索和处理实现自动化、高速化，也可以为支持政府决策和进行地学研究提供必要的保证。

1．地矿资源勘查信息系统的特点

地矿资源勘查信息系统与一般企业信息系统相比较，在服务对象、服务性质和服务内容方面存在着本质差别。一般企业信息系统的服务对象是企业内各级管理人员和上级机关管理人员，而地矿资源勘查信息系统的服务对象除此之外，更主要的是第一线的技术人员。从服务性质和服务内容上讲，一般企业信息系统主要着眼于“物质产品生产过程”和“物质产品流通过程”的信息收集、存贮、管理和处理，为产品结构调整、产量和质量控制、品种更新、新产品设计以及日常生产管理，提供自动化工具和决策依据，因此也被称为企业管理信息系统；地矿资源勘查信息系统则主要着眼于“地矿（类似于企业的原材料）的自然属性和时空特征”的信息收集、存贮、管理和处理，不但为地矿（原材料）价值的综合评价及其利用方式、方向的决策提供依据，同时也为揭示地矿（原材料）形成的自然规律提供依据。而且，由于地矿勘查的“产品”是勘查报告（或称调查报告、勘探报告）——描述被勘查对象的特征和性质的数据资料集合，因此，从某种意义上说，地矿资源勘查信息系统本身既是“原材料”与“产品”的管理工序，也是“产品”的生产工序。

与一般的地矿资源勘查数据处理系统相比，地矿资源勘查信息系统在服务对象、服务性质和服务内容方面的差别也很显著。一般地矿资源勘查数据处理系统的服务对象，只是基层勘查单位内部或科研机构内部的技术人员，通常不存在单位之间数据共享问题；而地矿资源勘查信息系统的服务对象除此之外，还包括企业内部各级管理人员和上级机关管理人员，以及不同单位的科技人员，因此必须考虑单位之间数据共享问题。一般地矿数据处理系统都以处理功能为核心，其软件开发的重点是各种数据处理功能（数据分析、统计、计算和编图），为了数据输入方便，通常利用数据文件或简单的应用数据库来组织数据，很少花大力气对数据库管理系统平台作二次开发。目前，我国地矿数据处理系统的开发以二维、静态和人机交互形式为主，但向集成化、系统化和实用化方向发展。地矿资源勘查信息系统与一般数据处理系统的最大区别，就在于它以主题数据库为核心，数据处理功能的开发围绕数据库展开。为此，要求对所选用的数据库管理系统平台进行大量的二次开发。

地理信息系统是数据处理系统的优化与改进特例。现行地理信息系统具有强大的二维空间信息管理、分析、提取、转换、叠加、图示和评价等功能，能够很方便地处理各种二维空间数据，并且从中提取大量的信息。然而，其内置的关系型应用数据库，也仅为满足空间信息处理的需要而设，难以存储和管理复杂的地矿属性数据。建立于基层的地矿资源勘查信息系统，不仅应当具备功能强劲的空间信息处理子系统，还应当具备结构合理、功能强劲的主题数据库子系统。地矿资源勘查信息系统与地理信息系统的主要差别，就在于它必须集强大的空间信息系统与强大的点源主题数据库系统于一身，不仅要管理和处理具有二维空间分布特征的地矿数据，而且要管理和处理多源、多类、多量、多维和多主题的

地矿属性数据，应当是一种多库联合、“多 S”集成的综合技术系统。

2. 地矿资源勘查信息系统的结构与功能

勘查数据管理、勘查数据处理和地矿资源预测评价等 3 大功能，是一个完整的地矿资源勘查信息系统所必备的。勘查数据管理功能由勘查数据库子系统来实现；勘查数据处理功能由勘查数据综合整理子系统和勘查图件机助编绘子系统来实现；地矿资源预测评价功能则由决策支持子系统来实现，其中包括地质成矿过程数学模拟、地矿资源预测和地矿资源评价等工作模块。根据结构-功能一致性准则，系统的功能要求是系统结构组织的依据，而系统结构组织是系统功能要求的体现。这种一致性准则将通过系统设计来体现。

地矿资源勘查信息系统的结构从技术方法体系角度可分为内、中、外 3 层（图 2-2）。内层为数据管理层，是整个系统的核心，它贯穿于系统运作的全过程，也贯穿于系统功能应用的各个层次。数据管理层的职能是实现数据组织、存贮、检索、转换、传输和交叉访问。外层是技术方法层，是整个系统的手足，可为功能应用层的各个层次及数据管理层服务。它包括各种高功能的硬、软件平台和 GIS、多媒体技术、人工智能和人工神经网络技术。中层是功能应用层，是整个系统的躯干。它由下而上分为数据采集、点源数据库、综合整理、图件编绘和决策支持等 5 个层次。功能应用层的职能是实施系统的全部功能处理，其底部的主题数据库子系统与内层的数据管理连为一体，成为点源数据库子系统。

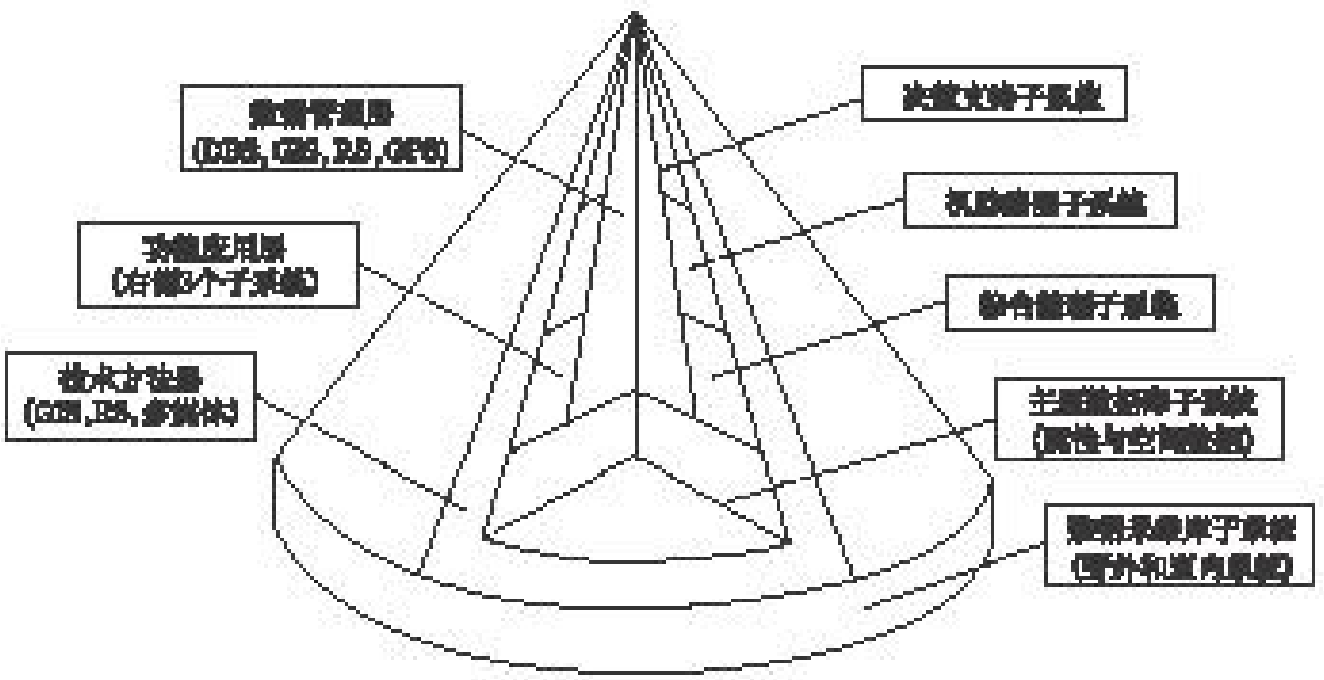


图 2-2　以主题数据库为核心的地矿资源勘查信息系统结构

（据吴冲龙等，1998）

数据采集子系统处于地矿点源信息系统的底层，是整个系统的基础，可实现野外与室内各种属性数据和空间数据的手工和机助入库。

主题数据库子系统处于数据采集子系统之上，是整个地矿资源勘查信息系统的核心。在本系统中，各种功能处理软件的开发都有共同的数据库基础——主题数据库。该数据库子系统以数据管理为根本，采用专题关联的先进设计技术，实现数据模型与代码标准化，并且有强大的数据存贮、管理和操作功能；有信息齐备、功能齐备、安全高效、应用方便的特点，在为各个功能应用层提供原始数据支持的同时，能为区域性和全国性地矿信息网

络提供点源数据服务。

地矿勘查的综合整理子系统处于点源信息系统的第 3 层，包括野外资料整理、专题资料汇总、日常数据处理、储量计算、多元统计分析和地质规律分析等次级子系统。这个子系统既可直接为地勘单位的日常生产管理和报告编写服务，又为地质规律、成矿规律及找矿勘探方法的研究服务。地质规律分析主要包括勘查区（研究区）的构造作用、岩浆作用、沉积作用、变质作用和成矿作用的规律分析，由于涉及参数多、关系复杂，通常采用人工专家系统和人工神经网络技术来实现。

地矿勘查图件机助编绘子系统处于本系统的第 4 层，包括地面地质、钻探、物探、化探和遥感等技术手段所形成的各种日常生产图件、解释图件和综合图件。地矿勘查和科研工作的成果大多以图件的方式来表达，这些图件的种类之多、数量之大、结构之复杂，也是一般企业管理信息系统所不能比拟的。完整的图件编绘子系统的设计和研究包括两大部分，其一是图形的计算机辅助设计（Map CAD)，其二是图件的计算机辅助出版（Map CAP)，二者既有区别又有联系。CAD（Computer-aided Design）所追求的目标是利用计算机的辅助功能，高效率地将各种技术手段所获取的、经过数值计算子系统处理过的原始数据资料转化为高质量的图形信息；而 CAP（Computer-aided Publish）所追求的目标是利用计算机的辅助功能，准确地将 CAD 或手工方式所设计的图形读入并编辑成合乎要求的印刷版型。二者的根本区别在于前者着眼于地矿信息的图形表达方式设计，后者着眼于已经形成的图形的修饰和排版设计。此外，图件的多媒体动态、彩色、立体编制技术和显示技术也是一个重要的部分，目前正在迅速地向地矿勘查信息处理全过程渗透。

地矿勘查开发决策支持子系统处于地矿资源勘查信息系统的最高层，包括勘查区本身及其外围和深部的矿产、水文、工程、环境等资源的数量、质量及赋存、分布状况预测的次级子系统，以及地矿勘查开发的地质技术经济条件综合评价、决策支持次级子系统。决策支持子系统的职能，是为领导机构及地矿勘查单位立项或承接任务提供决策依据。它们可以采用静态的方式，通过经典数学、随机数学、模糊数学和灰色系统方式，以及人工专家系统和人工神经网络方式来实现；也可以采用动态的方式，通过地质成矿过程数学模拟(对于含油气盆地而言，有盆地模拟、油气成藏动力学模拟、油气勘探目标模拟等）和矿产资源勘查开发过程模拟的方法来实现。此外，还应当开展地矿资源勘查开发过程——企业生产工艺过程、产品流通过程和企业发展过程模拟，即“企业过程”的模拟。

总之，地矿资源勘查信息系统是一种信息齐备、功能完善的地质与矿产勘查（察）区点源信息系统，是一种在计算机硬、软件的支持下，高效率地对地矿原始数据进行收集、评价、存贮、维护、检索、统计、分析、综合、显示、输出和发送、应用的综合性技术系统。

3. 地矿资源勘查信息系统应用软件的设计思路与策略

一个功能强大的地矿点源信息系统，不仅需要开发大量的应用软件和高效率的机助软件工程（CASE）技术，建立完善的、适合于不同服务环境和不同服务对象的系统工作平台及信息工程方法论体系，还需要采用一般系统工程的基本思想和适合于我国实际情况的系统设计思路与策略（吴冲龙等，1996，1998)。

在我国一系列大型地矿资源勘查信息系统的开发应用中（吴冲龙，1998)，所采取的设计思路与策略可大致概括为以下 6 点:

1）地矿资源勘查信息系统的设计应当建立在结构-功能一致性的准则之上，成为一种以点源主题数据库为核心的综合技术系统，其结构是技术方法与应用模型的层叠式复合结构，各种功能处理软件的开发都应当有共同的主题数据库基础。

2）将结构模块法与原型化法结合起来，严格地遵循以系统分析为系统设计基础的原则，根据地矿勘查的工作特点和流程，勘查点源数据库的设计应采用由下而上和由上而下相结合的设计方式，即通过用户需求调查和数据现状调查，由下而上建立实体模型、数据模型和概念模型，然后转化为物理模式，再根据地矿勘查工作的当前模型和未来模型确定总体模型，由上而下分解实体集及其属性，建立数据模式。

3）为了保证系统的各组成部分之间相互协调以及整体目标的顺利实现，在地矿资源勘查信息系统的开发和构建的过程中，对 DBS、CADS 和 GIS 的系统分析和系统设计，必须运用系统工程的理论和方法统一进行，不是刻意追求个别软件功能的最优化，而是努力追求系统整体功能的最优化，使之成为从野外采集数据到数据存储、管理、数据分析、图件编绘，再到地矿预测、评价的综合技术系统。

4）为了使该系统具有数据共享和软件共享的双重性能，应当采用行业或部门统一的数据模型、标准的代码体系、规范的图式图例、约定的处理方式和通用的软件接口，应当着力开发友好的用户界面，并且对系统进行集成化和商品化包装。

5）尽量利用现有的基础软件和应用软件，在计算机技术与地矿勘查的结合点上下工夫，注重二次开发，并且使所开发的应用软件同时成为新的开发平台，能支持用户按需要进行补充再开发，还能方便地移植到更新更强的基础软件平台上。

6）应当借鉴地理信息科学的研究成果，根据地矿勘查工作的特点和实际需要，采用面向对象技术、数据仓库技术、网络技术、ODBC 和 OLE 技术，进行“多 S”（RS、GPS、DBS、GIS、CADS 和 ES 等）结合与集成的探索，不断提高和强化其数据管理和处理功能。

这种具有完善的硬、软件配置的分布式地矿资源勘查信息系统，能够实现从多源、多类、多量、多维、多主题数据采集到建库、数值处理、图件编绘，再到综合解释、预测、评价和决策的完整工作流程；该系统又是一个开放式的应用软件开发平台，不但提供各种接口以便能随时接纳国内外的新功能软件成果，还能支持各用户按各自的特殊需要方便地进行补充再开发。这种综合性技术系统是提高地矿勘查资料利用价值，推动地球科学与地矿科学向定量化方向发展的重要条件，也是为各种综合数据库系统提供数据流的源泉。

第二节　地矿资源勘查信息系统的开发方法

地矿资源勘查信息系统的“开发”与“建造”是两个既有区别又有联系的概念。“地矿资源勘查信息系统的开发”是指软件开发者根据某一专业（或行业）领域的用户一般需求，通过对信息系统基础软件（DBS 和 GIS 等）进行再开发（或称为二次开发），从而建立一个具有一定通用性的高功能 RIS 专用软件。“地矿资源勘查信息系统的建造”则是指用户根据本单位地矿信息管理和处理的实际需求，选择并利用适宜的专用 RIS 软件，来具体地建立一个适合于本单位使用的点源信息系统。以往这两方面的工作都由专业软件人员承担，今后随着计算机应用与信息系统的普及，地矿资源勘查信息系统的“开发”工

作将由软件人员和用户共同承担，而地矿资源勘查信息系统的“建造”工作则将主要由用户——新一代地质与地矿勘查技术人员自身来承担。

正确的开发策略和方法是提高信息系统开发质量和成功率、降低开发费用的基本保证。地矿资源勘查信息系统的开发与建造是一项复杂的系统工程，必须从系统的观点出发，按照系统发展的客观规律进行作业。从数据库诞生以来的30多年中，人们通过大量的实践，探索和发展了许多指导管理信息系统开发的理论和方法，如结构化生命周期法、企业系统规划法、战略数据规划法、原型法和面向对象的分析及设计法等。其中，最常用而且有效的方法是结构化生命周期法、快速原型法和面向对象法（刘鲁，1995；李之棠等，1997）。这些方法在原则上也适合于地矿资源勘查信息系统的开发。

一、结构化生命周期法

结构化生命周期法是结构化分析和设计方法的简称，出现于20世纪70年代，属于目前使用较多的成熟方法。这种方法是对20世纪60年代发展起来的各种信息系统开发方法的总结和优化。其分析过程是一个自顶而下的功能分解过程，而设计过程是一个自下而上的功能合成过程。每一个过程都被严格地划分为若干个阶段，并且预先规定了各个阶段的任务，要求依照准则按部就班地完成。下面对其开发工作的特点和阶段划分作一些简要介绍。

1. 结构化生命周期法的特点

与20世纪60年代的其他信息系统开发方法相比较，结构化生命周期法有如下特点：

1）预先明确用户的要求，根据需求来设计系统　强调直接为用户服务，要求以用户需求为系统设计的出发点，而不是以设计人员的主观想像为依据；在未明确用户需求之前，不得进行下一阶段的工作。需求的预先定义，使系统开发减少了盲目性。这既是结构化生命周期法的主要特点，也是相对于在它之前的各种开发方法的主要优点。

2）自顶向下来设计或规划信息系统　结构化生命周期法着眼全局，从维护系统总体效益的角度来设计或规划系统，保证了系统内数据和信息的完整性、一致性；注意系统内部或子系统之间的有机联系和信息交流，并且努力防止系统内部数据的重复存贮和处理，从而大大地减少了数据的冗余量，保证了系统运行的有效性。

3）严格按阶段进行开发　该方法将系统开发的生命周期划分为若干个阶段，每个阶段都规定了明确的任务和目标，进而将各个阶段划分为若干项工作和步骤。这种有序的安排，不仅条理清楚，便于计划管理和控制，而且有利于后续工作的展开，基础扎实，不易出现返工现象。

4）工作文档标准化和规范化　文档既是本阶段软件开发工作的重要成果，也是下一阶段软件开发工作和用户了解并使用软件的依据。为了保证对各子系统和模块之间通讯内容的正确理解，该法要求文档采用标准化和规范化的格式、术语、图形和图表，使开发人员及用户有共同语言。

5）运用系统的分解和综合技术，使复杂的系统简单化　把信息系统看成是功能模块的集合，自顶而下地将一个复杂的大系统分解为一系列相互联系而又相对独立的子系统直至模块，可以使对象简单化，便于建模、设计、软件开发和系统实施，并且可以使系统的设计和实施过程，成为一个自下而上的功能合成过程。利用这种技术，能够方便地将子系统及其功能模块综合成完整的系统以体现总体功能。

6）强调阶段成果的审定和检验　为了增强系统开发的有序性，要求加强阶段成果的审定和检验，及时发现并清除系统中的隐患，弥补各阶段工作的不足与过失。根据结构化生命周期法的建模原则，只有得到用户、管理人员和专家认可的阶段成果，才能作为下一阶段工作的依据。

2．采用结构化生命周期法进行系统开发的过程

结构化生命周期法的基本思想是将系统开发看作工程项目，有计划、有步骤地进行工作。通常认为，各行各业的具体信息系统处理内容尽管不同，但信息系统的开发过程都可以划分为 5 个主要阶段（图 2-3），每个阶段可以再细分为若干个工作步骤。

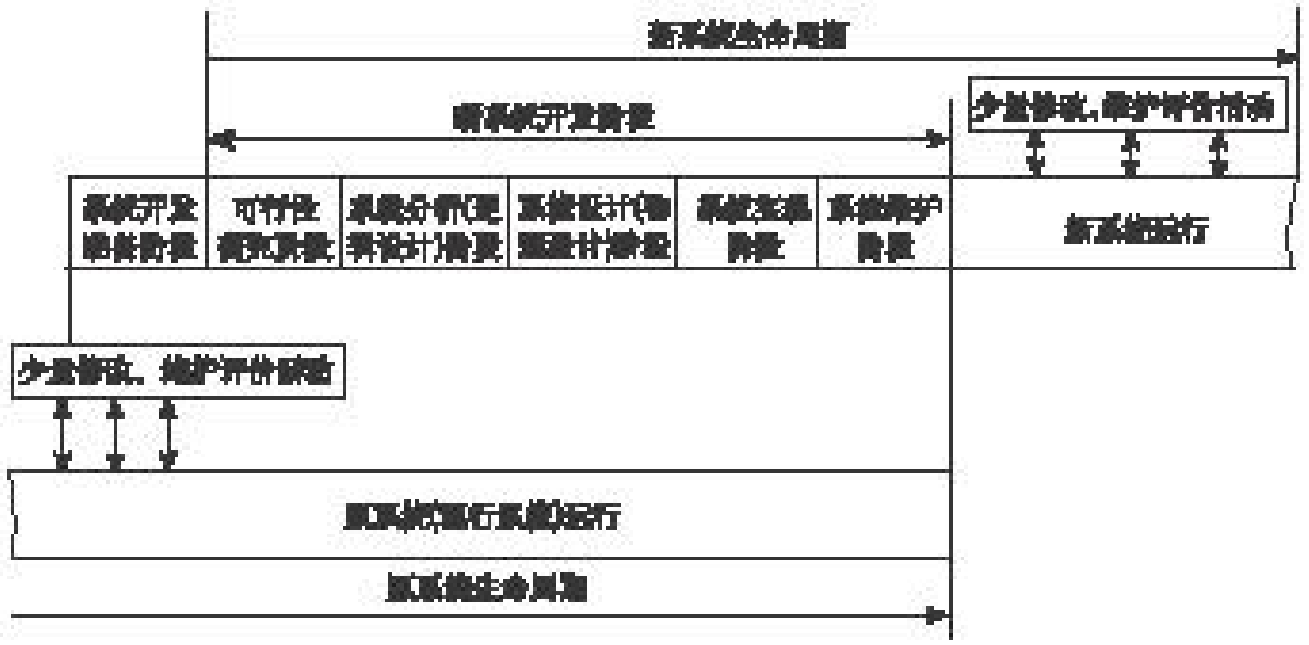


图 2-3　信息系统的生命周期

（据刘鲁，1995）

1）系统开发准备阶段　在筹划建立地矿资源勘查信息系统或者现行系统不能适应新形势要求而着手开发新系统之前，用户单位首先应当选派有关人员进行初步调查，然后组成系统开发领导小组，规定时间进度并制定开发计划。如果系统开发采用外包方式，本阶段还应包括投标过程。

2）可行性研究阶段　系统分析员采用各种方式，调查用户的业务工作现状和数据现状，搞清系统的界限、组织分工、业务流程、资源及薄弱环节等，并用一系列相关图表加以表示。在此基础上，进行需求分析，同时结合系统开发的经费来源、社会效益与经济效益、政策因素与法律约束、用户的组织机构和隶属关系等方面进行可行性综合分析。然后与用户协商讨论，提出对系统目标的初步设想和论证，编写并提交可行性报告。在许多专著与教科书中，可行性研究阶段与系统分析阶段也被合并为系统分析阶段。

3）系统分析阶段　系统分析是指在建立新的信息系统之前，从系统观念出发，应用各种技术方法和手段，对系统的目的、目标、结构、功能和信息工作流程等进行全面而深入的分析和评价，找出存在的问题及其产生的原因，并根据现时的目标需求和所处环境条件，拟定出系统模型及解决问题的可行性方案。系统分析阶段也称为系统的逻辑设计阶段，其具体步骤是：①进行目标分析，具体规定系统目的和目标，明确所要解决的问题；②编制数据流程图，确定系统的输入、输出以及数据存贮、管理与处理功能，做好各种数

据和资料的收集和分析工作；③划分子系统及功能模块，建立系统的逻辑模型，确定最优方案并提交系统逻辑设计说明书。

4）系统设计阶段　系统设计是信息系统开发的核心，其基本任务是把系统分析阶段提出的逻辑模型变成系统的物理模型，即按照系统分析阶段所确定的目标和逻辑模型，具体地设计出运行效率高、适应性强、可靠性高且经济实用的系统实施方案和应用软件。因此，系统设计又称为系统的物理设计。系统设计阶段可成总体设计和详细设计两个亚阶段。其步骤是：①总体设计：设计系统的工作平台，设计软件的总体结构并定义物理模块，设计输入输出的格式、内容及使用频度；确定安全保密和操作控制规范；②详细设计：定义输入、输出介质，完成人-机过程、代码和通讯网络设计，逐一制定每个功能模块的具体算法与数据结构。系统设计的关键是系统总体的结构厘定和模块化分解。

5）系统实现阶段　这是把系统的物理模型付诸运行的实践过程。其任务是：①制定实施计划、组织实施队伍、购置硬软件及配套设施；②重新组织信息流程、修订业务规程；③程序编制与调试；④系统平台（硬、软件）的安装与调试；⑤信息系统的整体调试；⑥培训上岗人员、整理数据；⑦数据库与数据文件的实际加载；⑧试运行、评价与验收。以上任务中，④、⑤、⑥及⑦等四项工作可以并行开展。系统实现阶段的输出文档包括程序文档和系统实施报告两大类。程序文档是今后系统维护、修改、扩充的主要详细技术依据，由程序设计报告、源程序清单及程序调试报告等组成。程序文档的编制应该按照国家颁布的软件设计规范来建立。系统实施报告是系统验收、审计、评价以及运行维护的依据，主要有：①系统实现计划；②设备采购及安装验收报告；③业务规程及有关制度；④系统调试及试运行情况报告；⑤系统转换及验收报告；⑥系统的操作使用手册等。

6）系统维护阶段　信息系统不仅结构复杂，而且时限性显著。任何一个企业和单位的业务活动与管理活动总是随着系统的目标、环境及自身条件的变化而不断变化和发展的。为了适应这种变化，一方面信息系统开发时必须留下发展的余地；另一方面也必须不断地加以改进和更新。这就需要进行系统维护。系统维护的内容是多方面的，如系统的处理过程、程序、文件、数据库的修改，甚至某些设备和组织的变动。一旦旧的系统在总体上不能适应发展的需要，就要适时地组织和开展新一轮系统调查、系统分析、设计、实现与维护。新旧系统的每一轮交替，就是一个新的信息系统生命周期的开始。

3. 结构化生命周期法的局限性

结构化生命周期法的理论基础较为严密，它要求系统开发人员和用户在系统开发初期对系统的功能就有全面的认识，并制定出每一个阶段的计划和说明书，来规范随后的各项工作。其思想基础是认为任何系统都能在建立之前被充分理解。在 20 世纪 60 ～70 年代，信息系统的范围比较狭窄（大多偏重于企业的财务、仓库、设备等方面的管理）、规模有限，使用环境也相对稳定，用户对这些系统的工作方式大都比较了解，而且易于和系统开发人员通过各种通信方式和渠道进行沟通。因此，在系统开发初期就有可能对系统的功能进行解剖、分析和设计，进而开发出满足用户要求的系统。这样的系统一旦完成，往往不要求做很大的变动。近年来，随着时代的变迁和计算机技术的发展，出现了许多新情况，对这种传统的系统开发方法提出了挑战。一方面，信息系统的开发已经成为一个行业，许多软件开发人员不具备实际工程——用户专业领域的知识；另一方面，用户常因忙于自己的业务工作而无暇与系统开发人员仔细讨论。在这种情况下，传统的结构化生命周期法便

显得无能为力了，需要寻找一种易于和用户沟通的开发方法。

结构化方法的着眼点在于一个信息系统需要什么样的加工方法和过程，并把数据结构和处理方式截然分开，用过程抽象来应付系统的要求。对一个系统来说，结构是相对稳定的，而行为则是相对不稳定的。结构化方法把基点放在相对不稳定的行为（即功能）上，难以适应系统的变化。这是又一个很大的缺点。另外，结构化方法主要由数据流程图（DFD）及控制结构图（CSD）等图表工具来描述，一般只能表达顺序流程和平面结构，且不太精确，不能全面描述信息系统模型。从某种意义上来说，结构化方法是强调人们从计算机软件开发人员的角度，而不是从用户的角度来思考要实现的信息系统。显然，这种方法是与计算机应用的初级阶段相适应的。

二、快速原型法

快速原型法是为了加快同用户的交流，尽早明确用户的要求，缩短系统的开发周期，提高软件开发效率而提出来的。这种方法使用户在大规模的系统开发展开之前，能够尽快地看到未来系统的全貌，了解系统功能及效果；开发人员可以及时对模型修改补充，为用户展开新的模型，直到用户满意为止，形成最终产品。

1．快速原型法的基本思想

快速原型法是从人本身灵活、多变、依经验行事的特点出发而提出来的一种新的信息系统开发方法。这种方法的基本思想是:

1）并非所有的需求都能预先定义　需求的预先定义虽然在某些情况下是可能的，但在更多的情况下却难以实现。例如，当系统面对的数据具有多源、多类、多维、多量和多主题特点，而且需要做各种复杂处理时，其实体模型、概念模型和数据模型一时难以用常规方式来建立；在信息系统开发的初期，用户的要求是经常变动的，有时甚至是模糊的；最终用户对计算机能为他们做些什么并不完全了解，而软件人员又常常不熟悉用户的业务；开发人员与用户之间存在着通讯上的障碍，不能充分地传递信息和交换意见；用户只有在看到一个具体的系统时，才能清楚自己的需要和系统的缺点；每个信息系统开发项目的参加者都有自己的观点，他们对系统的理解和看法不会完全吻合等等。

2）反复修改是必要的、不可避免的，应该加以鼓励　用户需求的多变是预先定义法实施中的最大困难。为了方便，该法在系统开发的早期就把用户需求加以“冻结”，使得所开发的信息系统带有很大的惰性，用户无法扩充其数据存储和处理功能，更难以随着社会环境和客观条件的发展而发展。为了使未来系统所提供的信息能够真正满足管理和决策的需要，系统的开发方法应当接受和鼓励用户不断提出更多、更高的新要求，并且要能够根据这些新的要求对系统进行反复修改。基于这种思想，快速原型法将系统的可修改性作为一项重要的功能指标。

3）需要一个系统模型来作为系统开发的雏形　在系统开发初期，开发人员和用户都需要有一个能够说明系统结构和功能的模型。采用结构化生命周期法进行系统设计时，通常是用结构化的语言及图表来描述预先定义的系统模型——概念模型和数据模型。当用户的需求模糊或不能预先定义时，用这种结构化的语言及图表所描述的系统模型很难被用户理解和接受。为此，需要建立一个直观的、动态的系统模型。这个模型尽管不是完善的和最终的系统模型，但却是可以实际运行并且易于为用户理解和接受的实验模型。更重要的是，这个模型必须是能够不断地根据实际情况和用户的需求进行修改和完善的，因而可以

作为系统开发的雏形——系统初始原型。

4）只要有合适的工具，就能够快速建造和修改原型　系统原型要为用户所理解和接受，并且作为系统开发的雏形，就必须在立项初期甚至前期就快速地建造出来，随后又能够根据用户的要求方便地进行修改。这需要有软件开发工具的支持，特别是要有高水平和高效率的计算机辅助软件工程（CASE：Computer Assist Software Engineering）技术的支持。近年来随着计算机硬件和软件技术的发展，支持原型化的工具软件和 CASE 技术发展十分迅速，一些集成化的软件工具也已经进入市场。对于属性数据库开发而言，第四代语言（Fourth Generation Languages，简称 4GLs）是支持原型开发的代表性工具软件。该系统不仅带有大量通用的模块，而且有一个应用生成器（Application Generator），可以根据有关描述快速地生成应用程序。快速原型法大量采用并发展这些技术来强化自己，使开发人员摆脱了繁重的手工作业，实现了系统原型的快速建造和修改。

2．快速原型法系统的生命周期

一般认为，利用快速原型法开发信息系统要经过如下几个阶段：

1）确定用户的基本需求阶段　在这一阶段中，用户应当从系统输出的角度清楚地表达自己的基本要求——系统必须具备的基本功能和人-机界面的基本形式等。系统开发人员据此来确定哪些要求是现实的，所需数据的来源、获取方式、数据类型以及存贮、管理和处理方式，同时对开发成本做出估计。

2）开发系统原型阶段　系统开发人员根据用户的基本需求，使用高效的软件工具和高层次的开发语言，快速地开发出一个具有人-机交互功能的系统原型。该原型仅用于反映用户基本需求，只需易于演示和理解，并不要求完善。研发者也可以直接在先期开发的类似系统的基础上，按用户的基本需求改造出新的原型。

3）提炼用户需求阶段　用户通过原型演示或亲自使用，了解其信息需求得到满足的程度以及存在的问题。开发人员一方面从用户口中获知系统的缺点与不足；另一方面借助原型来引导和启发用户表达对系统的最终要求。在用户和开发人员的反复讨论过程中，进一步提炼用户需求，确定需要修改和变动之处。

4）改进系统原型阶段　系统开发人员根据用户所提出的修改意见或存在问题，对系统原型进行修改、扩充和完善。这个阶段与第三阶段应当重复进行，开发人员反复与用户接触并充分交换意见，一直到用户满意为止。整个过程如图 2-4 所示。如果用户对修改过的原型表示满意，则该原型便成为一个新的运行原型和应用系统开发的基础，或者可以成为新的应用系统。

3．系统原型的种类和构造方法

（1）系统原型的种类

从系统发展的观念上看，原型可以分为丢弃型和进化型两类；从系统应用的观念上看，原型可以分为研究型、试验型和进化型三类。

研究型原型也称为需求原型，在没有任何常规需求分析能够满意地识别和确认用户真正需求的时候，可用于强化需求定义和功能分析阶段的工作。研究型原型不仅能有效地帮助用户加深对系统功能的理解，而且能够激发用户的需求欲望和创造性，有利于将初始解迅速转化为信息系统的最终解。

试验型原型是最终系统描述的一种补充或强化方式，可分为有系统整体试验型原型和

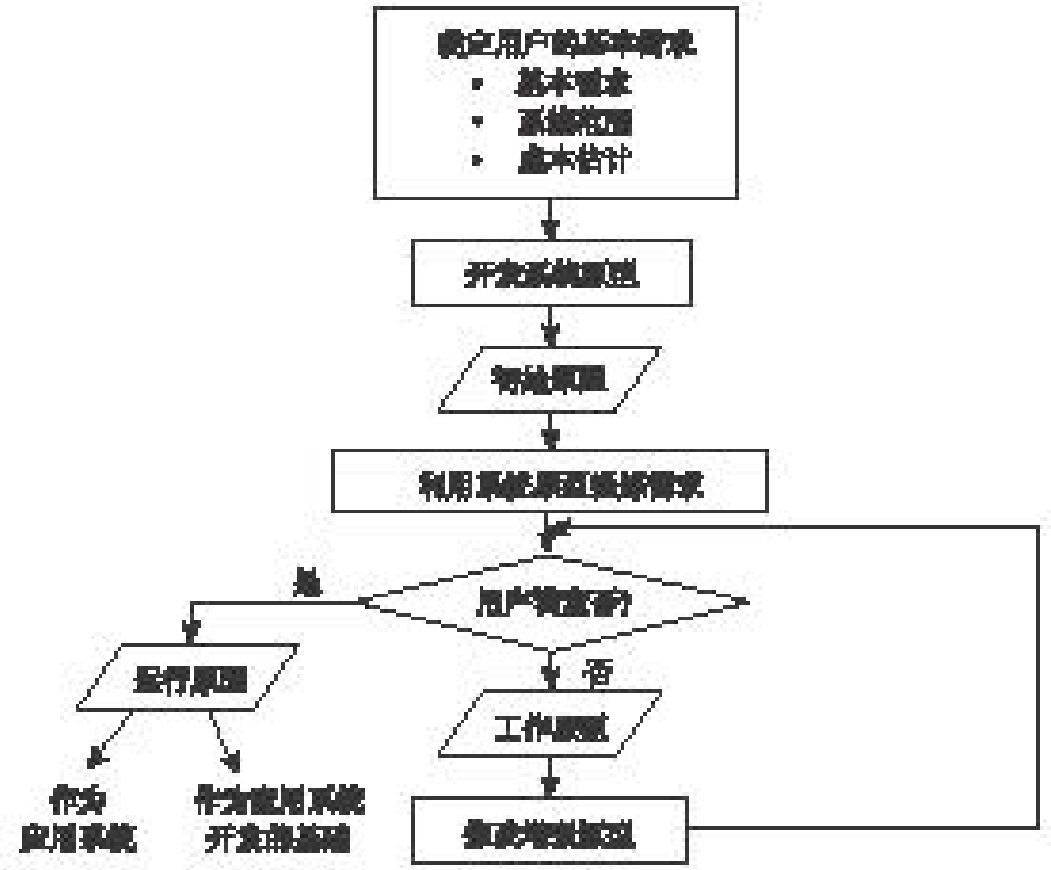


图 2-4　利用快速原型法开发信息系统的过程

局部试验型原型。这种原型在系统的各个开发阶段，用于对系统整体和局部的信息管理、处理功能以及系统的可行性、透明性、方便性、可靠性等进行仿真试验。

研究型原型和试验型原型只在系统开发初期或开发过程中起作用，一旦完成系统设计便将被丢弃，因而都属于丢弃型原型。

进化型原型是快速原型法的核心，它将随着系统开发的深入而被不断修正和改进，并且成为应用系统的开发基础，甚至进化成为最终的应用系统。该原型的提出基于如下事实：①围绕交互式应用系统的环境连续不断地变化，新的需求不断地出现；②交互应用系统改变着周围环境，也引起了新的需求。因此，通常认为进化型原型的建立和修改、完善的过程，也就是快速原型法的实施过程。

(2) 原型的构造方法

不同原型有不同的用途，其构造的思路和方法也不同。

1) 研究型原型　这类原型的构造没有规范的形式和规律可循，其构造过程具有非正规化的不确定性。正是这种不确定性，使得该方法更具有创造力。由于用户对系统开发的期望将受到原型的影响，设计人员可以从用户需求的某一特定功能着手，设计出使用户易于理解的初始原型。让用户运行这种功能的初始原型，评价所提出的部分解决方案的可行性。开发人员对反馈的信息反复进行分析和提炼，直到用户满意为止。这时，开发人员及时将资料整理出来，作为系统设计的参考，然后丢弃该研究型原型。通过用户的密切协作，这个过程有可能在 1 ～3 个月内完成。

2) 试验型原型　试验型原型的构造方法在很大程度上取决于试验目的和所选择的试验策略。具体方法类型有如下几种：①人机交互界面仿真原型；②轮廓仿真原型；③局部功能仿真原型；④全局功能仿真原型。一个好的试验型原型的构造方法，应当成为最终系

统的一种强化描述工具。它所起的作用是：补充系统用户需求描述；从已有的描述或部分描述中提炼出主要的描述信息；作为介于描述和实现的一个中间阶段，为下一阶段服务。

3）进化型原型　这种原型的构造方法是采用一种进化中的近似问题解来表征原型系统，从而使初始原型成为进化中系统的核心。在进化型原型的每一个发展阶段，都会产生出更多的系统特征。这就使得原型的构造过程成为整个开发过程的一种代替，而原型最终也必然成为一种运行的系统。其开发形式有递增式和进化式两种。

A. 递增式原型　适合于解决需要集成的复杂系统设计问题，其开发过程是一种缓慢的累积扩展过程，大致可分为总体设计和反复进行的功能模块实现两个阶段。在递增式原型的开发初期，先给出系统的总体框架，各模块的功能及结构也要清楚地表达出来，并且用一些表演数据来说明。在以后的开发过程中，通常采取类似计算机硬件的“插接策略”，逐个地完善这些模块。递增式原型开发的前提是：在开发过程中，系统的组织机构、模块的外部功能均不发生变化。

B. 进化式原型　适合于解决用户需求不清或一时难以搞清的系统设计问题。其开发过程是从设计到实现，再到评估，反复进行的周期过程。所开发的产品被看成是各阶段的版本序列。在进化式原型的开发初期，不仅系统的总体框架无法给出，连各模块的功能-结构也难以全部清楚地表达出来。为满足用户变化的需求，要求有一种既能严格计划与组织，又能根据实际情况随时改变的信息系统开发方法。研究型原型和试验型原型在进化式原型的系统开发早期混合使用比较有效。一旦弄清了用户需求，便可以逐步地将研究型原型和试验型原型丢弃，只留下改进了的进化式原型继续开发。

4. 快速原型法的适用性

快速原型法采用设假求真的策略，比结构化生命周期法更为明智。其优点主要有：①开发周期短、费用相对较少；②用户的参与更及时、更直接、更富有建设性；③易于改进，因而生命力较强；④易学易用，用户培训时间短。但是，快速原型法也有一些明显的缺陷，例如：难以直接面对大系统和复杂系统；开发过程管理困难，如果用户配合不好会拖延开发进程；软件人员易于用原型取代系统分析。从目前情况看，快速原型法比较适合于用户需求不清、管理及业务处理不稳定、需求经常变化、系统规模较小且较为简单、信息可以相对分散处理的用户。

三、面向对象法

面向对象（Object-Oriented，OO）方法起源于面向对象的程序设计语言。在 20 世纪 60 年代末期，Simula 67 语言的设计者为了便于在计算机上进行仿真建模，在 Algol 语言的基础上引入了对象概念。随后，Smalltalk 的设计者引入了对象、类、继承和动态束定等机制，率先以面向对象的程序设计为核心，集各种开发工具为一体，建立了具有很强的图形功能和多窗口用户界面的计算机环境。面向对象法的基本思想是：对问题领域进行自然分割，以更接近人类通常的思维方式建立问题领域的模型，便于对客观的信息实体进行结构模拟和行为模拟，从而使设计出的软件尽可能直接地表现问题求解的过程。由于这种方法符合人的思维习惯，便于实现问题空间与解空间的自然映射，可以应用于复杂问题的软件分析和设计，因此当 20 世纪 70 年代末出现软件生产落后于硬件生产的危机之时，其研究和应用便成了一股强劲的浪潮。经过十几年的发展，产生了以 C + + 为代表的新一代面向对象程序语言，并且形成了从软件的分析、设计到编程实现的一整套面向对象方法论体系。利用

这些面向对象的程序设计语言和方法，可以实现 GIS 和 CAD 所需要的各种性能，甚至可以描述和操作可扩充的复杂数据模型，因而适用于 GIS 和图像、拓扑数据的管理和处理。目前，国内外普遍运用面向对象的 DBMS，来建立 GIS 的属性数据库系统。

1. 面向对象的概念和术语

面向对象开发方法涉及一系列概念和术语，其中最重要的是对象、类和继承。

1) 对象　对象是所研究和描述的事物，可以是具体的也可以是抽象的。例如，一套地层、一片土地、一个区域是对象，一个工厂、一个车间、一个生产组也是对象；同样，一个地矿勘查过程、一个产品的设计过程、一个生态环境的发展过程、一种数据结构和一种系统行为等等也都是对象。面向对象方法以对象作为最基本的元素，将软件系统看成是离散的对象的集合。面向对象方法分析问题解决问题的核心就是对象的分析和描述。

2) 类　类是具有一致数据结构和行为的对象的归纳和抽象，例如，“地质”、“矿产”、“土地”、“资源”和“环境”等都是类。每个类都是个体对象的可能的无限集合，而每个对象都是其相应类的一个实例。类反映了与对象的应用有关的重要性质，而忽略了其他一些与应用无关的内容。任何类的划分都应该与具体的应用有关。类中的每一个实例均有各自的属性值，它们的属性名和操作是相同的。

3) 继承　继承是对具有层次关系的类的属性和操作进行共享的一种机制。在面向对象的方法中，定义一个类，可以在一个已存在的类的基础上进行。其要领是先把这个已存在的类所定义的内容（行为）作为自己的内容，然后加入若干与新类有关的内容。这个已存在的类被称为父类，通过继承而由父类所定义的新类被称为子类。利用这种继承机制，可以最大限度地减少系统设计和程序实现中的重复性。

2. 面向对象的分析和设计

面向对象的分析（OOA: Object-Oriented Analysis）就是研究和理解系统中的对象及其之间相互作用时出现的事件，并以此来把握系统结构和系统行为的一项工作。其任务是理解现实的问题（系统）并用便于软件开发人员理解的术语和方式表达出来，即建立系统的逻辑模型；其内容包括根据用户的需求，确定系统应由哪些类（对象）组成，明确各个类（对象）之间的相互关系（从结构上），明确各个类包含哪些属性，分别提供哪些服务，以及明确各类（对象）之间是怎样通过消息的传递和响应来实现系统所需各项功能的。

面向对象的设计（OOD: Object-Oriented Design）是根据面向对象的分析所提供的系统逻辑模型，建立能用现有开发工具实现的、具有较高的性能/价格比的物理模型的一项工作。从实质上看，面向对象的设计过程是将面向对象的分析结果映射到某种实现工具的结构上的过程。其内容包括系统设计和对象设计。系统设计的任务是确定实现系统的策略目标系统的层次结构，以及所优化的性能和初步的资源配置。对象设计则是在系统设计的基础上加入一些实现细节。例如确定解空间中的类、数据结构、关联、接口形式及实现服务的算法。面向对象设计所采用的实现工具可以是面向过程的（例如 COBOL、C 等），也可以是面向对象的（例如 C^{++}、Smalltalk 等）。当实现工具是面向对象时，由于分析者、设计者、程序员和最终用户都使用相同的概念模型，映射过程有比较直接的一一对应。

3. 面向对象分析和设计的实现方法

据不完全统计，自 1988 年以来，国内外公开发表的用于实现面向对象的分析和设计

的方法有 50 多种。这些方法大致可划分为两种类型：以分析作为开发重点的方法，例如 OMT（Object Modeling Technique）方法、OOSE（Object-Oriented System Engineering）方法、ER 模型（Entity-Relationship model）、RDD 方法、OOA 方法（Object-Oriented Analysis Method）、OSA（Object-Oriented System Analysis）方法和 VMT（Visual Modeling Technigue）等；以设计作为开发重点的方法，例如 BOOCH93 方法等。下面仅对常用的 OMT 方法和 BOOCH93 方法作些简单的介绍。

OMT 方法（Rumbaugh，1991）是在实体主关系模型（ER 模型）的基础上扩展了类、继承和行为得到的。这是一种以分析作为开发重点的代表性方法。其特点是：强调对系统和相关领域的理解，并在此基础上建立模型。它引入了三个模型来描述问题——对象模型（Object Model）、动态模型（Dynamic Model）和功能模型（Function Model）。对象模型用信息结构图（Information Structure Diagram）来描述系统的静态结构、问题域中的对象和对象之间的结构，并通过对象图来描述系统内部的对象组成、对象间的关系以及对象的属性和操作。动态模型是对象模型的参照，通过事件图和状态转换图（State Transition Diagram）描述对象的动态行为，用对象的事件状态响应来刻画对象的时序性。功能模型用数据流图（Data Flow Diagram）描述了系统的所有计算方法——对象操作的含义，即如何从输入值得到输出值，而不考虑计算值的次序。这些模型从三个不同的角度（或三个不同的坐标维）描述信息系统的不同侧面，比传统、单一的数据流程图更全面、更深刻。IBM 公司 1996 年所公布的 VMT 方法，是对 OMT 方法的发展，它集合了 OMT、OOSE、RDD 等方法的优势，同时推出可视化编程工具 VisualAge 以实现 VMT 设计模型。

BOOCH93（Booch，1993）是以设计作为开发重点的代表性方法，它把工作集中在开发过程中的设计阶段。整个开发工作分成微观过程和宏观过程两部分。微观过程主要用于建立一个反复、递增的开发框架，而宏观过程则用于对微观过程进行控制。微观过程由脚本和产品的体系结构驱动，包括如下反复循环的 4 个步骤：①确定类和对象；②确定类和对象的语义；③确定类和对象的关系；④规范类和对象的接口和实现。宏观过程只关心开发过程中的管理问题，包括以下 5 个步骤：①概念化、建立核心需求；②分析、建立理想的行为模型；③设计、创建体系结构；④细化并完善和实现模型；⑤维护、管理和提交模型。这种方法的优点是能够细致地刻画类和对象的事件、触发、操作及其控制。

4. 面向对象方法在信息系统中的应用

面向对象方法在数据库技术中得到广泛的应用。目前 ODL（Object Design Language）正成为面向对象数据库系统的设计标准，以取代传统关系数据库的 E/R 模型；OQL（Object Query Language）和 SQL3 也正在成为面向对象的数据库系统的操作标准，将逐步取代 SQL。新推出的软件工具大多支持面向对象的开发，为信息系统的开发人员提供了极大的方便。其中，最受欢迎的专业化通用开发工具，当属微软公司的 Visucl C^{++}。它提供丰富的、功能强大的基础类库（MFC），完全支持面向对象的封装性、继承性和多态性，可充分发挥面向对象方法的特长，为信息系统的应用开发免除了大量的基础设计工作。前面提到的 IBM 开发的 VisualAge，还能支持面向对象开发的全过程。其他应用于数据库开发的软件工具有 Power Builder、Delphi、Visual Basic、Visualt 和 Foxpro 等。

一个较理想的面向对象数据库系统（OODBS），除了具备数据库管理系统的基本功能外，还是一个面向对象的系统，支持面向对象的数据模型。OODBS 赋予数据库设计和应

用开发人员很强的面向对象能力，从而大大扩展了数据库系统的应用领域，提高了开发人员的工作效率和应用系统的质量（冯玉才，1995）。与传统数据模型相比，面向对象数据模型的优势主要体现在以下几个方面：①具有表示和构造复杂对象的能力；②封装性和信息隐藏技术提供了程序的模块化机制；③继承和类层次技术提供了软件的重用机制；④滞后束定等技术提供了系统的扩充能力。

OO 方法与传统的过程式开发方法相比有着明显的优势。传统的结构化分析、设计技术对软件生命周期的各个阶段进行了严格划分，要求人们以过程方式来理解系统（面向过程），这些过程把给定的输入转换为所希望的输出，开发者的开发工作必须按先后顺序进行。而面向对象方法和面向对象技术则没有分析阶段与设计阶段的严格区分，要求开发者把系统定义为互相联系的一组对象集，这些对象互相作用并通过这些作用发生状态上的变化。系统分析和系统设计可以并行开展，当系统和问题复杂时也可以交叉进行。而且，由于面向对象分析、设计、编程、维护均以对象作为系统的基本构件，对象的稳定性和相对独立性使系统具有一种宏观的稳定效果。

具体地说，面向对象方法有如下优点：①直接从地质矿产工作流程和空间实体问题域出发，将数据模型和处理模型综合在一个完整的对象模型之中，使得分析模型与问题域形成良好的对应，从而避免了对研究对象问题域和系统责任的错误理解；②面向对象的组件式软件开发技术，能使文本、图形、图像和视频多媒体等多种类型的数据聚集于一个系统中，进行混合管理和操作，共同完成地矿勘查空间数据采集、管理与处理功能；③OO 方法分析者、设计者和实现者之间虽有不同的教育背景和工作特点，但可以使用共同的概念术语进行交流和沟通，使 OOA、OOD、OOP 及 OOI 可做到直接而自然的衔接；④以最稳定的成分——对象作为构成系统的基本单位，而把容易变化的成分——服务及属性封装在对象之中，这就有效地遏制了一处修改处处牵连的“波动效应”，使得面向对象技术有较好的弹性。当地质矿产勘查与开发需求发生变化时，系统的功能部分（对象的服务部分）、外部系统或设备的接口部分及描述问题域事物的数据（对象的属性）都能随之发生变化，而系统整体及其地质实体对象却可以保持稳定。

然而，由于国土资源特别是地质矿产的数据本身的多样性和复杂性，要做到属性数据与空间数据管理并重，面向对象的 DBMS 尚存在一些亟待解决的问题。首先是面向对象数据库尚缺乏完善的理论基础和对关系系统的改进，用户仍受到需要预先确定数据编录路径的限制，不便执行任意连接和特定的查询，也不便使用标准表和标准数据库；其次是关系型 DBMS 多年来在数据完整性、保密、并行访问、事务处理、破坏后的恢复、可编程界面和分布式数据库等方面所获得的成就，面向对象的 DBMS 难以一下子实现。在目前条件下，解决这个难题可采用对象-关系型数据库系统的设计方式，即在原有关系型数据库的基础上，增加描述非文本、非结构化数据的对象类和相应的对象支持功能，使数据库系统可以满足跨平台和不同媒体对象的应用；还可以运用各种先进的标准化系统连接技术，例如 Windows 环境下的动态数据交换技术（DDE）、对象的链接与嵌入技术（OLE）以及开放数据库连接技术（ODBC）等，将关系数据库（RDBS）与 GIS 集成起来，建立组合式地质数据库，并且在实践中不断地加以拓展和再开发，形成双向或多向灵活、流畅的信息流通体制，进而解决关系数据与空间数据的混合提取、转换、组合与处理等难题。

第三节　地矿资源勘查信息系统的开发过程

由于用户的专业领域和业务范围不同，所期望建立的地矿资源勘查信息系统在规模、级别、结构组成与功能需求上会有显著差别。这些差异以及用户需求的不确定性，使得各种系统开发方法都显现了固有的局限性。如何根据实际情况来选择一种合适的方法，以便保证地矿资源勘查信息系统研制的高效率、高质量和低成本，是开发人员和用户应当共同关心的问题。

一、系统开发方法的选择与综合应用

上面所介绍的结构化生命周期法、快速原型法和面向对象法等三种有代表性的开发方法，都是信息系统工程发展的历史产物，有着各自的优缺点和不同的应用领域及适应范围。从它们与使用者的关系来看，结构化方法离计算机人员近一些，快速原型法离用户近一些，而面向对象方法处在二者之间。按照系统建模的基本工作方式，信息系统开发的方法可以划分为两大类：预先严格定义法（如结构化生命周期法、企业系统规划法和战略数据规划法等）和非预先严格定义法（如快速原型法等）。面向对象法作为一种新的思维方式和通用的软件开发技术，既可以用于预先严格定义的方法中，也可以用于非预先严格定义的方法中，还可以作为一种独立的建模方式来使用。

在地矿资源勘查信息系统开发中，方法的选择应当把握及时、真实、完整地反映用户需求的准则。表 2-3 分别列出了选择预先严格定义法和非预先严格定义法的几种具体条件。

表 2-3　采用预先严格定义法和非预先严格定义法的影响因素

采用预先严格定义法的条件	采用非预先严格定义法的条件
用户要求定义明确，可以预先定义	用户需求不明确，难以预先定义
系统规模大且层次复杂	系统规模小且较为简单
要求数据管理与处理标准化	不要求数据管理与处理标准化
系统运行程序确定、结构化程度高	系统过程是非结构化的
系统的使用寿命较长	系统的使用寿命较短
开发过程要有严格的控制	系统要求在短期内实现
开发人员经验丰富且熟练程度高	开发人员缺乏该类系统的开发经验
用户环境与需求稳定	用户环境与需求易于改变
系统文档要求详细而且全面	拥有第四代语言或其他原型化工具

在实际的信息系统开发过程中，往往会遇到一些各种方法都无法单独解决的复杂情况。这时，需要把几种方法结合起来，灵活地加以运用。信息系统开发的实践表明，快速原型法和结构化生命周期法虽然在指导思想和具体做法上不同，但并非互相排斥，而是可以互相补充的。我们可以把原型的开发过程作为结构化生命周期法开发过程的需求定义阶段（图 2-5），弥补结构化生命周期法在需求定义阶段存在的或可能产生的困难。一旦需求完全清楚，就可以丢弃各种原型，采用严格的结构化方法进行开发。

地矿资源勘查信息系统是复杂的系统。为了解决这种系统的分析、设计与开发问题，

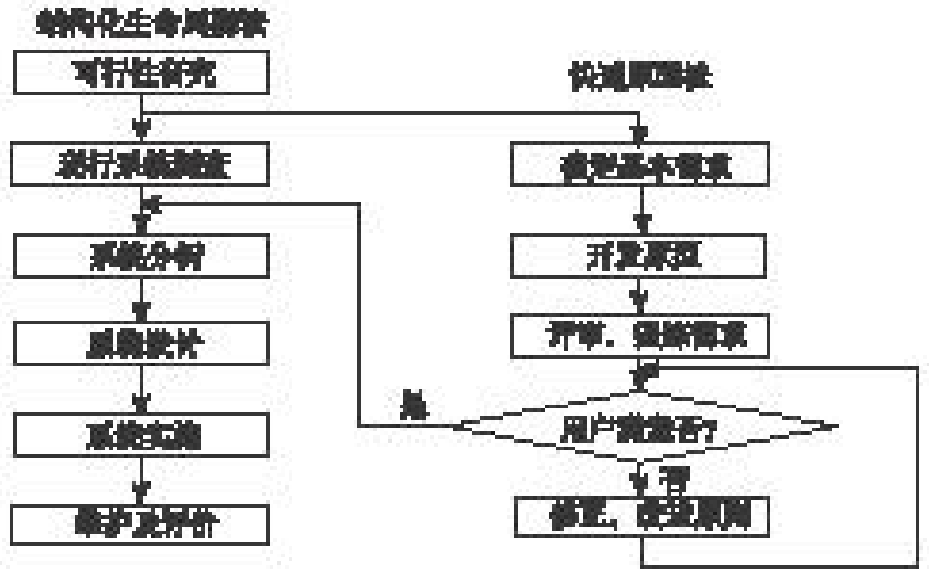


图 2-5　快速原型法与结构化生命周期法的结合方式

(据刘鲁，1995)

还可以以结构化生命周期法为基础，在需求定义阶段采用快速原型方法，在系统分析阶段采用结构化方法，而在系统设计和系统实现阶段采用面向对象方法；或者先进行面向对象的分析与设计，再用结构化生命周期法来编程实现；也可以在面向对象的系统开发和生命周期的有关阶段中，采用快速原型法进行模型求真。

二、地矿资源勘查信息系统的开发与建造过程

根据地矿勘查工作的特点与需要，本教材主要介绍以结构化生命周期法为基础、多种方法并用的系统开发与建造技术。在这里，首先简单介绍实施这种综合开发方法的基本过程及其主要工作内容。从一般系统工程的角度看，建立一个适合于本单位使用的 RIS，要经过系统调查、系统分析、系统设计、系统实现和系统维护等 5 个阶段（图 2-6）。

1．系统调查

这是地矿资源勘查信息系统开发的第一阶段。这个阶段的主要工作包括：①用户需求概略调查；②原有系统（如果有的话）概况调查；③新系统原型求真；④项目可行性论证。其中工作量较大且较重要的是原型求真和可行性论证。

进行原型求真时，开发人员要根据用户所表述的基本需求设计出一个初始原型，并向用户演示；然后倾听用户的评价意见，从中提炼出用户真正的需求，对系统原型进行修正、改进，再向用户演示和征求意见。这个过程必须循环往复进行，直到用户满意为止。所提供的原型应当符合用户所从事的地矿勘查类型，例如金属矿床勘查信息系统原型、非金属矿床勘查信息系统原型、煤田（煤产地）勘查信息系统原型、油气田勘查信息系统原型、水资源勘查信息系统原型和工程地质勘查信息系统原型等。必要时还可以提供更为细致的原型，例如脉岩型金矿勘查信息系统原型、中小型陆相盆地石油勘查信息系统原型、大型水利枢纽坝区工程地质勘查信息系统原型。

2．系统分析

系统分析在系统原型求真和项目可行性论证基础上进行，大致包括如下内容：①用户需求调查与需求分析；②地矿勘查工作的业务现状分析；③地矿勘查工作的数据现状分析；④业务发展趋势研究和系统动态分析；⑤信息系统功能目标分析；⑥地矿勘查的实体

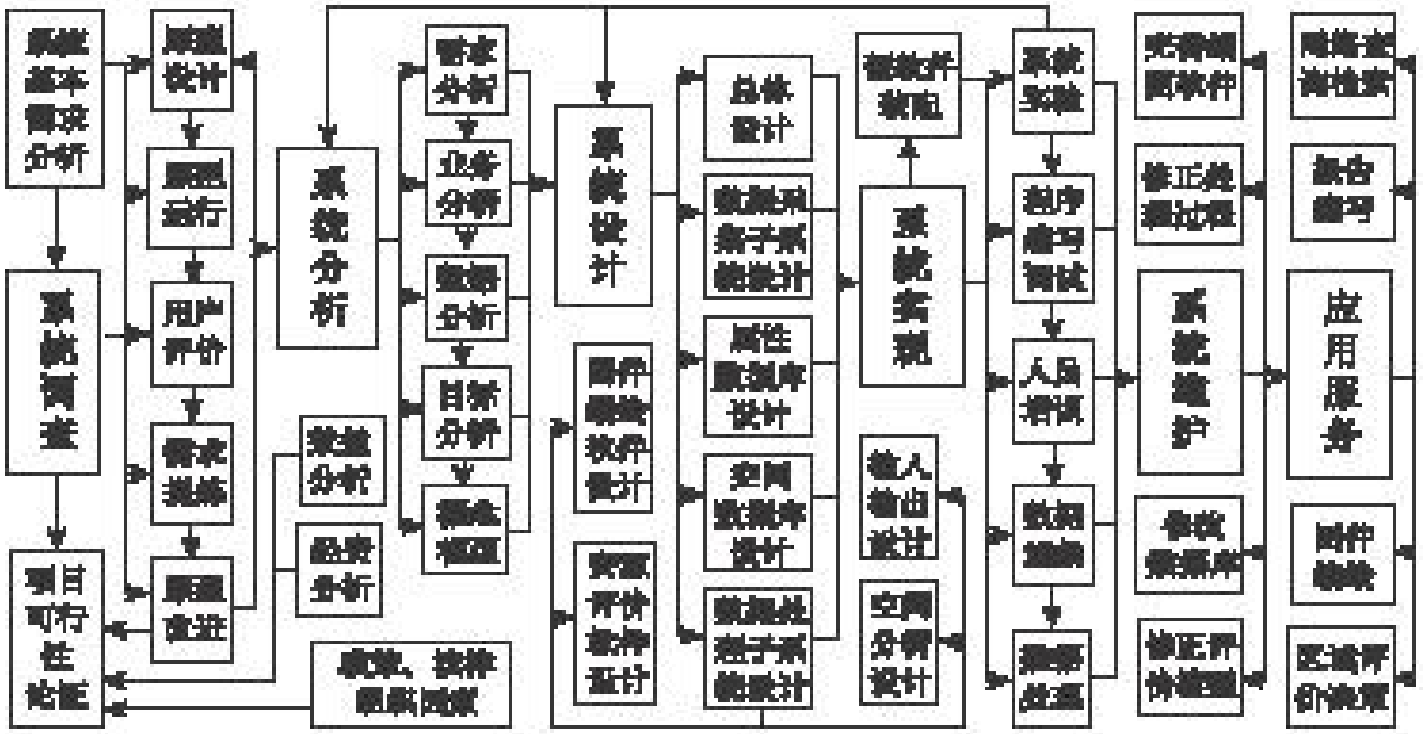


图 2-6　地矿资源勘查信息系统（RIS）的设计和建造流程

模型研究；⑦地矿勘查的概念模型（逻辑模型）研究；⑧系统安全保护策略与措施分析。系统分析的核心是需求分析、业务分析、数据分析、目标分析和概念模型研究。

与系统调查阶段的系统基本需求分析相比，本阶段的需求分析要求深入、具体、全面和周详。业务现状分析是指对具体地矿勘查单位的日常工作内容、技术手段、质量标准和工艺流程等的调查、研究和整理、归纳。数据现状分析不仅要求查明该单位在业务过程中的数据来源、数据类型与特征、数据采集方式与方法，还要求弄清在本单位使用数据以及为社会服务过程中的数据的流向。由于地矿勘查工作带有很强的研究性质，要使所建立的信息系统具有强大的生命力，在系统分析阶段进行业务发展趋势研究和系统动态分析是十分重要的。与此相应，系统的功能目标分析也应当包括当前和未来两部分内容。通过目标分析，要具体规定系统开发目的、当前功能目标和未来功能目标，要明确所要解决的问题。系统的未来功能目标不必在本次开发过程中实现，但要求所开发的系统能够支持未来的再发展。这些工作都要求有用户的参与和配合。

通过用户需求分析、业务现状分析、数据现状分析和当前系统功能目标分析，可以建立当前的地矿勘查实体模型；而通过系统动态分析以及未来的系统功能目标分析，可以建立未来的地矿勘查实体模型。因此，实体模型研究应当包括当前和未来两种模型。这两类实体模型的概括和抽象，便是系统的概念模型。概念模型的研究包括子系统及功能模块的划分和最优方案的确定。这显然是一种“由下而上”的建模过程。这种概念模型是一种逻辑模型，因此，通常也把系统分析阶段称为系统的逻辑设计阶段。系统分析阶段结束时，要求提交系统逻辑设计说明书，以便作为系统设计的依据。

3. 系统设计

按照结构化生命周期法的基本规定，系统设计的基本任务是把系统分析阶段提出的逻辑模型变成系统的物理模型，即按照系统分析阶段所确定的目标和逻辑模型，具体地设计

出运行效率高、适应性强、可靠性高且经济实用的系统实施方案和应用软件。

在系统总体设计时，开发人员要勾画系统的总体轮廓，划分并确定系统的软件模块和硬件系统的结构和组成，选定信息系统的层次结构模型。地矿资源勘查信息系统在总体上可分为数据采集子系统、属性数据库子系统、空间数据库子系统和信息处理子系统等四大子系统，每个子系统又可以包含若干个功能模块。开发单位的负责人必须根据系统分析得到的系统目标和概念模型进行规划，并且定义物理模块、设计输入输出的格式和内容、确定安全保密和操作控制规范，还要对开发人员进行具体分工。

在系统详细设计时，应当以建设属性数据库和空间数据库为重点，并且围绕数据库展开各种处理功能的研究和软件设计。每个子系统的负责人要进一步给出该子系统的层次结构模型，明确地定义输入、输出介质，完成人-机过程、代码和通讯网络设计，逐一编写每个功能模块的具体算法和数据结构；要编制实现每一个功能的说明书，特别是相应的软件模块说明书，指出每一个功能模块的功能目标、开发要求以及如何去实现它。该说明书是程序员编写程序或修改、移植现存的软件——对基础软件进行二次开发的依据。

如果说在系统分析阶段的逻辑建模过程是一种“由下而上”的过程，那么在系统设计阶段所进行的数据模式建造过程，则是一种“由上而下”的过程，即在进行勘查区数据模式研究时，先根据地矿勘查工作的现状确定总体模型，再根据地矿勘查科学的发展及勘探技术的可能改进逐级分解实体集及其属性。

4. 系统实现

在系统实现阶段也需要用户与开发人员密切配合。系统开发人员的主要任务是：①按照系统设计说明书的规定编写并调试各个子系统的功能模块；②进行子系统的整体调试和子系统之间的联合调试，③进行系统优化与集成、开发统一的用户界面；④协助用户单位重新组织信息流程、修订业务规程；⑤培训上岗人员，并指导其数据整理与数据输入；⑥制定系统维护方案和安全保护措施；⑦完成系统整体试运行，并交付评价与验收。按照一般信息系统的结构化生命周期法设计要求，开发人员在系统实现阶段必须提交高质量的程序文档和系统实施报告。与此相应，用户的主要任务是：①组织实施队伍、选派人员接受培训，筹措硬件及配套设施；②在开发人员的帮助下进行系统平台（硬、软件）的安装与调试；③整理数据、输入数据，完成属性数据库和空间数据库的实际加载；④组织并参与系统整体试运行、系统评价与验收。地矿资源勘查信息系统的各子系统的功能设计都是围绕属性数据库和空间数据库展开的，其功能只有在用户完成上述相应的任务，特别是完成属性数据库和空间数据库的实际加载之后，才能得到完全的体现。

5. 系统维护

系统维护包括系统的日常管理、安全保护以及为了适应地矿勘查业务和信息管理需求变化所进行的修改与完善工作。地矿勘查单位的业务与管理活动，总是随着社会经济的发展和科学技术的进步以及系统的目标、环境及自身条件的变化而不断变化和发展的，这就要求地矿资源勘查信息系统相应地加以改进和更新。地矿资源勘查信息系统维护所进行的修改与完善工作，主要包括系统的数据处理过程及应用程序、软件设计文档、数据库结构、编图软件、地矿评价模型和输入输出等方面的修改，有时也涉及某些基础软件、设备

和人员组织的变动。当发现旧的系统在总体上不能适应发展的需要，甚而阻碍了业务工作和信息管理活动时，系统维护人员有责任及时而慎重地提交分析报告，请求开展全面的系统评价，以便决定是否结束该信息系统的生命周期，进行新一轮系统开发。

总之，开发和建造功能强大的地矿资源勘查信息系统，涉及到地质矿产的业务知识、经济理论、计算机科学、运筹学、统计学等多学科的理论和方法，必须坚持多学科协同、跨专业协作的原则，做到软件开发人员与用户专业人员密切配合。

思　考　题

1. 说明信息、数据和数据载体的含义，比较它们之间的差别。
2. 什么是信息系统？其结构和组成如何？信息系统可分为哪些类型？
3. 信息系统建设应当遵循哪些基本原则？为什么说信息系统建设应当以数据库为核心？
4. 信息资源管理与信息管理有何联系？有何区别？
5. TPS、MIS 和 DSS 的结构和功能有何特点？它们之间有何联系？有何差别？
6. 什么是地矿资源勘查信息系统？它有些什么特点？简述其结构与功能。
7. 地矿资源勘查信息系统的开发思路和策略是什么？
8. 地矿资源勘查信息系统与地理信息系统有何联系？它们之间的主要差别是什么？
9. 试比较结构化生命周期法与快速原型法的优缺点，它们分别适合于什么环境条件？
10. 地矿资源勘查信息系统的开发需要综合采用结构化生命周期法与快速原型法吗？

第三章　地矿资源勘查系统分析

地矿资源勘查系统分析，是将勘查单位的各项业务、研究工作及其组织、机构、人员、设备和资金等等作为一个完整的系统，对其数据管理和处理的需求、业务现状、数据现状、系统建造目标和系统结构特征进行全面分析，并且逐步建立该系统的实体模型、概念模型和数据模型的过程。系统分析是地矿资源勘查信息系统设计的基础。

第一节　地矿资源勘查系统需求与工作环境分析

系统需求与工作环境分析是系统调查阶段和系统分析的重要内容。在系统调查阶段，只要求进行概略性的需求分析与工作环境分析，目的是提出合理的初始原型；在系统分析阶段，则要求在系统调查阶段所作的原型求真和需求提炼基础上，进行深入细致的需求分析与工作环境分析。同一地区、同类型地矿资源勘查单位，由于工作性质、数据采集、数据处理的内容和方式相同，通常有相同系统需求和工作环境，其分析结果可以互相借鉴；而不同地区、不同类型地矿资源勘查单位，由于勘查工作及其数据采集、处理的内容和方式有显著差别，系统需求与工作环境分析应当分别具体地进行。

一、系统需求分析

系统需求分析主要是分析和归纳各级、各方用户对地矿资源勘查信息系统的总体功能要求及对各子系统的具体要求，然后以此为依据来确定系统的基本服务对象和内容，划定系统的边界，建立系统的实体模型和概念模型，选择合理的硬软件配置。在分析用户需求时，应当同时兼顾目前和将来，以便使系统结构趋向合理，易于扩充和转换，使系统功能保持最佳状态。根据国内外的经验与教训，系统需求分析工作一定要走在信息系统设计的前面，并且应当扎扎实实地进行，使之成为系统设计的向导。

在系统调查阶段，系统需求主要靠用户访问和原型运行启发来获知；在系统分析阶段，系统需求一方面通过总结前一阶段的调查结果得到；另一方面可通过细致的用户调查来得到。地矿资源勘查信息系统的用户包括所在勘查单位的基层工程技术人员、工程技术负责人和行政领导以及上级综合信息系统。其中勘探队的基层工程技术人员和各级技术负责人，将是本信息系统最主要和最经常的服务对象，也是用户需求调查的主要对象。用户调查主要是采用直接征询（座谈或填写专用调查表格）与查阅有关勘查规范、勘查报告相结合的形式进行的。调查的主要内容是各级、各方用户对地矿资源勘查信息系统的使用要求，例如，需要从系统中查询哪些数据？各种数据以什么方式存入？要对这些数据进行哪些处理？等等。表 3-1 所示的工程地质勘察用户对系统的需求（部分），就是通过直接征询得到的。

根据国土资源部和各有关主管工业部门以及国家技术监督局历来颁布的各种规范、标准，并参考相应的各种勘查设计书和总结报告，可以完整地归纳出系统将要面对的工作内容、所涉及的数据项和字符值，还可以统计出其中常用的属性数据和空间数据的数量。

表 3-1　××××引水工程地质勘察信息系统用户需求调查结果（部分)

序号	将向数据库查询什么数据和问题	可能查询频度				查询方式		输出方式		
		规划阶段	可行性研究	初步设计	技施设计	随机	专题	文字	表格	图形
一	渠线设计施工									
1	工程简介	经常	经常	不经常	稀少	√		√		
2	渠线位号及地质背景	经常	经常	不经常	不经常			√		√
3	基本地质条件与结论		经常	经常	经常	√	√	√	√	√
4	全线渠道长、宽、高、水深、流量、边坡、挖填比					√		√	√	
5	渠线沿途地形剖面线、渠中位、渠道纵坡、水位、控制高程	经常	经常	经常	不经常	√	√			√
6	渠线沿线渠中位、渠道纵坡	经常	不经常	不经常	不经常				√	
7	渠线上的建筑物（渡槽、闸室、隧洞等)		经常	经常	不经常			√	√	
8	各类交叉建筑物的平剖面图及主要的数据		经常	经常	经常	√	√	√	√	√
二	地质观察与描述									
1	钻孔资料		经常	经常	经常	√				
2	各土层的钻孔柱状图		经常	经常		√	√	√	√	√
3	单个钻孔所反映的地层结构		经常	经常	经常			√		√
4	沿线各段地层结构及其沉积环境					√		√	√	√
三	地理条件									
1	渠线通过河流（名称、规模、流量等)									
2	渠线沿途的气象资料	经常	经常	经常	经常	√	√	√	√	
3	施工运输条件		经常	经常	经常	√	√	√		
4	地形资料及精度					√	√			
四	渠线调查概况									
1	各阶段、各项目主要工作量		经常	经常	经常	√	√		√	
2	渠线沿途既往地质勘查历史	经常	经常	经常	经常	√	√	√		
3	渠线沿途既往钻孔概况	经常	经常	经常	经常	√	√	√	√	
五	灾害与稳定性									
1	渠线沿途地震烈度区划及各渠段的地震基本烈度	不经常	不经常	经常	经常	√	√	√	√	
2	渠段未来的地震趋势与地震危险区	不经常	不经常	稀少	稀少	√	√	√		
3	沿途灾害地质及其防治措施					√	√	√	√	
4	不同土质、岩质的边坡、坡高、坡比	稀少	经常	经常	不经常		√	√	√	
5	黄土、膨胀土、砂、砂壤土的工程地质性质及分布			经常			√	√	√	√
6	土岩渗透系数、渗透变形类型及稳定性	稀少	经常	经常	不经常	√	√	√	√	
7	七度及以上地震区饱和砂土的分布与震动液化判别	稀少	不经常	不经常	稀少		√	√		√
8	渠线通过煤矿区的压煤情况及采空区	稀少	不经常	经常	稀少		√	√	√	
9	高边坡稳定性	稀少	经常	经常	不经常		√	√	√	√

续表

序号	将向数据库查询什么数据和问题	可能查询频度				查询方式		输出方式		
		规划阶段	可行性研究	初步设计	技施设计	随机	专题	文字	表格	图形
六	水文地质调查									
1	工区范围地下水位变化情况	稀少	经常	经常	不经常	√	√	√	√	√
2	地下水化学成分及各土层中水压、气压资料	稀少	经常	经常	不经常					
3	各建筑物地下水和河水的水质评价	稀少	不经常	不经常	稀少		√	√	√	
4	各渠段地下水位	稀少	经常	经常	稀少	√		√	√	√
七	采样测试									
1	特殊土试验数据原位测试数据	不经常	经常	经常	经常		√		√	
2	地层土工试验数据及静探旁压，十字板试验数据	不经常	经常	经常	经常		√	√	√	
3	粉、细砂层厚度自重及震波速、原位测试	经常	经常	经常	经常		√	√	√	
4	各过河建筑物桩基设计有关力学参数	不经常	经常	经常	经常		√	√	√	
5	历年来物理力学试验资料及主要建议指标索引	不经常	经常	经常	经常	√	√	√	√	
6	其他室内物理力学试验常规指标	稀少	经常	经常	经常	√	√	√	√	
八	综合数据及综合评价问题									
1	砂土震动液化的评价方法	稀少	经常	经常	经常	√	√	√	√	
2	施工中特殊土边坡的防治方法措施及设计	不经常	不经常	不经常	经常	√		√		
3	穿黄工程及主要建筑物工程地质剖面及主要工程地质问题					√	√			√
4	各类型交叉建筑物的工程地质条件、选定方案、基础及其处理	不经常	经常	经常	经常		√	√	√	√
5	全线选定方案（比较纵坡面）典型地区横剖面比较和选定的坡比，衬砌型式及施工优化	不经常	经常	经常	经常	√			√	√
6	工程建筑资金，各种建筑材料用量、挖填总方量	稀少	不经常	经常	不经常	√	√	√	√	
7	占有农用田面积、移民占房总数及估价	稀少	不经常	经常	不经常	√	√	√	√	
8	调水量、调水利用系数、渠道渗水量、调水价格	稀少	不经常	经常	不经常	√	√	√	√	
9	施工周期、人力、机械功率	稀少	稀少	不经常	不经常	√	√	√	√	
10	使用维修费用（万元/年），管理调度人员总数	稀少	稀少	不经常	不经常	√	√	√	√	
11	工程使用年限及对完工后备时间段效益评价	稀少	稀少	不经常	不经常	√	√	√	√	
12	投入产出比	稀少	稀少	不经常	不经常	√	√	√	√	
13	渠线的活动断裂	不经常	不经常	稀少	稀少		√	√	√	
14	渠线各类土层所占渠段长度	稀少	经常	不经常	稀少		√	√	√	
15	渠线填方段、半填半挖、挖方段所占的渠线长度	不经常	经常	经常	稀少		√	√	√	√
16	输水隧道的工程地质分段	稀少	不经常	经常	经常	√	√	√	√	√
17	检索已有的图件	不经常	经常	经常	不经常	√		√		√

二、系统工作环境与工作流程分析

地矿资源勘查信息系统的总体模型已如图 2-2 所示，其工作环境除了硬件、软件和机房设施之外，主要是指用户单位的信息系统建设和应用条件。用户单位的计算机应用水平，业务现状和发展前景，各级行政领导、工程技术负责人和基层工程技术人员对信息系统的认同感和欢迎程度，上级综合信息系统的建立与否，以及系统应用之后可能产生的效益（社会效益和经济效益），都是十分重要的环境条件。这些条件影响着该单位对信息系统建设的积极性和资金、人员的投入量，进而决定着系统建设的规模和目标，开发人员需要认真、具体地进行分析。开发人员应当从该单位的业务发展前景和系统的未来模型出发，注意发掘其潜在的需求，进而预测改善该单位系统工作环境条件的可能性。用户单位的潜在的需求和系统工作环境条件改善的可能性，都是系统目标分析和系统设计的重要依据。

从用户的角度来看，地矿资源勘查信息系统的工作大致有 6 个方面：①数据准备；②数据输入；③数据管理；④信息处理；⑤成果输出；⑥发送应用。用户的工作模式是系统目标分析和系统设计的又一重要依据，需要根据各单位的情况具体地制定，其一般模式可用图 3-1 所示的工作流程来表示。

三、系统保护的策略与方法

从经济角度看，数据资料是国土资源特别是地矿资源勘查的主要成果，是地矿资源开发利用的决策依据，取得这些成果所花费的代价极高。其中有些内容，例如与储量、资源量、矿石品位、油质、煤质、水质、土质、岩土力学等有关的数据，则属于国家机密，是国家机构进行资源、能源、水源、土地和工程总体规划的决策依据。从技术角度看，地矿资源勘查信息系统是国家经济与国土资源信息系统的数据源和网络结点，点源数据库是地矿资源勘查信息系统的核心，数据是整个信息系统的“血液”与“养分”。没有数据不成其为信息系统，而有数据却遭到损坏或窃取，信息系统也不健全。因此，应当充分重视其安全保护。

系统的保护首先是点源数据库的保护。数据库保护的目的是排除和防止各种对数据库的干扰破坏，确保数据安全、可靠。对数据库的干扰破坏主要来自硬件设备（如 CPU，磁盘等）和软件系统（如 Windows，DBMS 等）的故障，以及非法用户（未获得使用权的人）偷窃、篡改信息和正当用户的操作失误。数据库保护的主要任务是设法防止对数据库的各种干扰和破坏，以及在数据库遭受破坏以后尽快地恢复正常。同时应搞好数据库的日常维护工作，适时地进行数据库重组，防止系统因性能下降而丧失正常的工作能力。因此，数据库的保护应当包括安全性保护、完整性保护、并发控制、数据库恢复和数据库日常维护等 5 个方面的工作（刘云生等，1992；冯玉才，1993）。

1. 安全性保护

安全性保护主要是指对数据库实行存取控制，其目标是禁止未授权使用的人非法存取数据，同时防止由于程序错误或操作错误引起非法存取数据。

地矿资源勘查信息系统的服务对象，是基层地质与矿产勘查队的技术人员、各级主管部门的领导和技术负责人员、有关的科学研究人员、友邻勘查队的技术人员以及上级综合数据库。为了体现内外有别、上下有别的政策，可以采取检查用机权限和存取控制的保护策略，即采用限制数据项值和给定专用子模式进行存取控制，并且在上机操作前利用

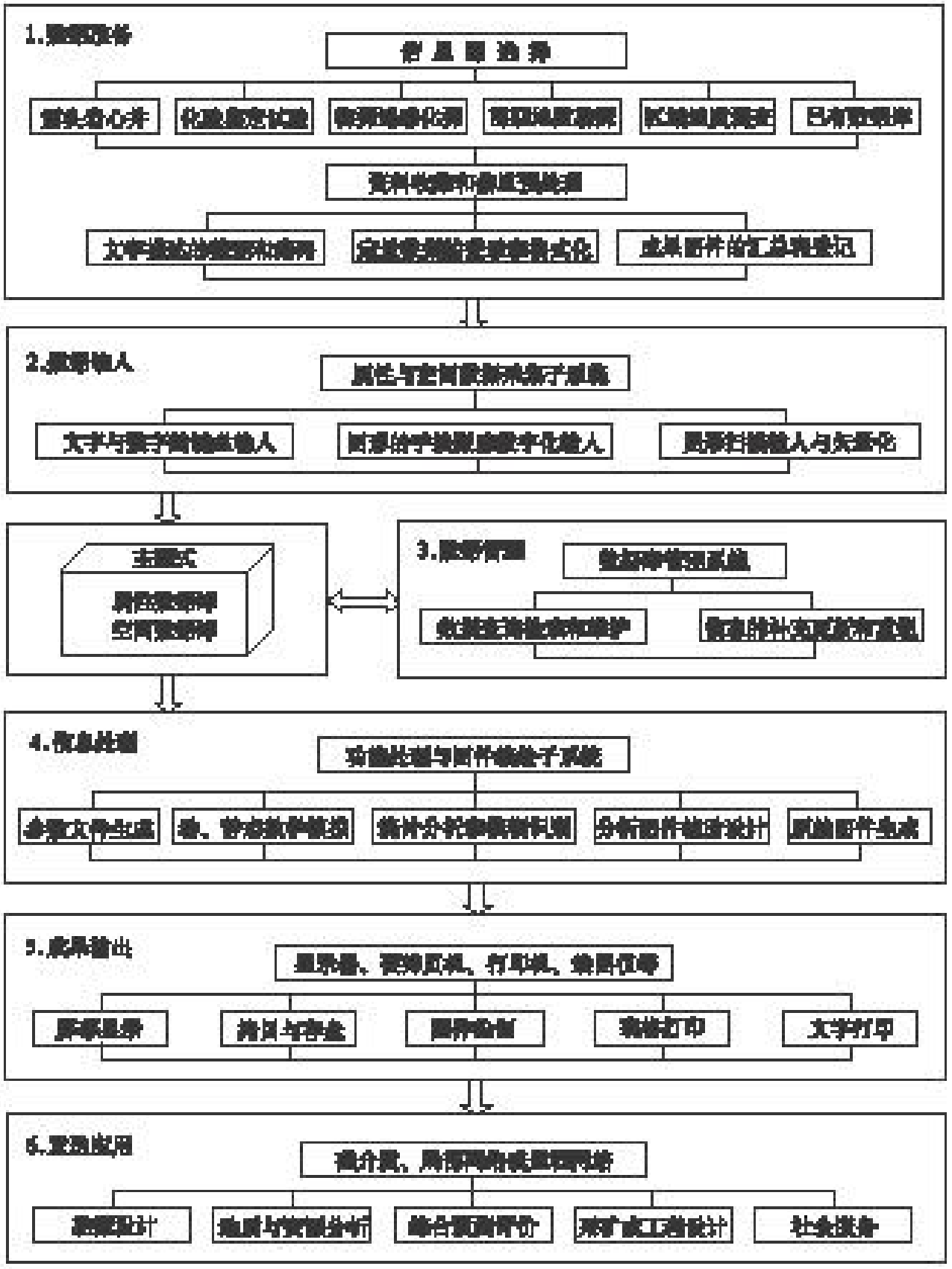


图 3-1　地矿资源勘查信息系统的用户工作流程

(据吴冲龙等，1996)

“用户标识符”，即用户代号进行用户鉴别，利用口令进行用户证实，以便审查用机人员的身份和权限。对于某些重要数据如矿体厚度、矿石品位和矿床储量等等，如有必要，可在它们入库时，实行密码变换，以密码的形式存放。

2．完整性保护

完整性保护是指预防数据的正确性、一致性遭受破坏的策略，其目标是尽量减少破坏的可能性，以及数据遭到破坏后能尽快地恢复原样。

目前常用的数据完整性保护策略有两类，一类是一般性完整保护策略，即防止一些离奇古怪的数据出现；另一类是结构完整性保护策略，即维持用户指定的函数依赖。一般完整性的保护策略是利用数据库管理系统（DBMS）来建立完整约束定义模块，用户在定义数据模式时，可利用这一模块对于所存放的每一属性都指出其域值的上限和下限，例如:

0≤岩层序号≤1000,
0 m≤岩心长度≤5.0 m,
0 °≤地层倾角≤90 °,
0≤各种岩石的矿物成分比例≤100%,
0≤各种矿物的化学成分含量≤100%。

岩层厚度、矿体厚度、煤层厚度、油层厚度、软弱夹层厚度、含水层厚度、矿石的密度和视密度、钻孔坐标等等，也都可根据当地的实标情况给定约束范围。在进行这样的完整性约束的同时，还应利用 DBMS 建立监控模块，当发现终端输入的数据，或程序计算出的数据及其他途径产生的数据与用户完整性约束不符时，让监控模块自动禁止这些数据存入数据库；当发现从数据库中取出的数据有类似的情况，则让监控模块自动提示数据的完整性已遭受破坏，并调用数据库恢复模块来恢复数据库。

结构完整性的保护策略也是由用户指定约束条件，由系统来验证实施。它要求用户和系统协调配合，在定义数据模式时，先利用定义模块指明必须满足指定的函数依赖。然后由监控模块来检查每一次输入与输出是否满足指定的函数依赖。若满足就执行存取，否则拒绝执行。例如，如果用户在定义中指定了下列函数依赖:

“钻孔号 - 层号 - 分层号 - 矿体 - 矿分层厚度”

则在数据库中就只能为每一个钻孔中的每一个矿分层保存一个厚度值。这样，凡是遇到有人要插入钻孔号、矿体号及矿分层号相同而矿分层厚度不同的元组时，DBMS 就拒绝执行这一命令。同样，如果指定了一个关系模式的关键字，则 DBMS 就应拒绝执行插入一个与该数据库文件中具有同样关键字和数据值的元组，即不允许在一个数据库文件中有两个完全相同的元组存在。如果一个元组缺少关键字，或缺少关键字的一个部分，DBMS 也不允许把它存放到数据库中。

只要做到上述各点，则数据完整性就能得到一定的保障。

3. 并发控制

为了达到信息充分共享，地矿资源勘查信息系统的数据库系统应该有多个输入、输出终端，并且还将支持短程和远程联网查询检索，因此，进行数据库设计时也应注意防止由并发操作造成的数据不一致性错误。

由并发操作造成的数据不一致性错误可能有 3 类: ①修改丢失; ②不能重复读; ③“脏”读出。为了防止这一类错误的发生，必须对并发操作实行某种限制。就目前情况看，封锁技术是行之有效且易于实现。封锁技术的要领，是在数据库系统中设置一个动态模块，用来对各个用户提出的封锁请求进行统一的管理。

一般地说，在短程的局域网络中，同时使用同一数据库的用户不是很多，数据查询与数据修改都比较少，其并发操作不是经常发生的，可以选择排它锁（Exclusive locks）的封锁方式，并且每次的封锁范围适当扩大，以选择记录作为封锁目标较为妥当，必要时甚至可

以选择整个数据库文件作为封锁目标。但考虑到今后可能与远程的国家经济信息网络联接，或进入“数字地球”、“数字国土”网络，需要采用更为先进和可靠的网络封锁技术。

4. 数据库恢复

地矿资源勘查区信息系统通常是建立在基层勘查队里，由于各方面条件都比较差，突发性事故更难避免，例如，停电、电源故障，外存设备故障和系统软件破坏等。此外还有计算机病毒的侵扰，等等。这些故障可能导致计算机寄存器中的数据、甚至整个数据库所存贮的数据全部遭到破坏。一旦遭到破坏，就应当立即设法使它恢复正常。

一般地说，数据库的恢复可以通过建立副本与采用更新的日志、设定检查点、微分文件的方法及采用 COMMIT 事务原则相结合来实现。根据地矿资源勘查区的实标情况，可采用建立光盘副本与更新日志相结合的方法来实现数据库的恢复。

地矿资源勘查区信息系统的数据是随着地矿资源勘查工作的进展而不断存入的。以金属、非金属及煤炭资源勘探为例，其钻探数据、岩心编录数据、钻孔水文地质数据和测井数据，通常在每个钻孔施工结束时存入；相应的样品化验测试数据和岩矿鉴定数据，是在结果送回之时存入的，其周期大约是一个月。露头观测数据及河、湖、井、泉水文地质调查数据，一般是在野外调查结束时入库，其周期大约是一年。所收集的前人工作成果(包括文字描述、测试化验数据和图件)，通常是在勘查设计制定前或勘查工作开始的初期存入，其周期与该次勘查工作相应，为 1 ～2 年。各阶段的综合数据和图件则随着室内整理工作的进展而逐步存入，其周期相当于一条勘探线的施工周期，大约是一个季度到半年。一个单位在一次勘查中的综合研究成果及储量计算成果，通常在勘查工作结束后的勘查报告编写过程中存入，其周期也与该次勘查工作相应，为 1 ～2 年。

综上所述，为了保证数据的安全与完整，并节约复制副本的成本，勘查区信息系统的倒库周期应与各种数据入库周期相适应，分为一月、一季度、半年、一年和两年 5 种。每次倒库的具体时间，应定在各种数据周期性入库工作结束的时刻。所建立的数据库，可提供下阶段勘探设计或矿山设计使用，也可成为新勘探阶段或矿山地质数据库的基础。

第二节　地矿资源勘查业务现状与数据现状分析

业务现状分析是在用户调查的基础上具体地分析某地矿资源勘查区都进行哪些工作？有哪些工种和技术手段？人员是如何分工和组织的？各种工作的内容和流程如何？数据现状分析则是具体地分析现有的业务工作能获得哪些数据？这些数据可以划分为几个类型？它们之间有何联系？哪些是基础数据？哪些是可以由基础数据生成的合成数据或综合数据？在进行业务现状和数据现状分析的同时，也应当对其不远将来的变化与发展做出估计。

一、业务现状与数据来源

我国的矿产资源勘查和工程地质勘察工作，通常是在区域调查、普查找矿或工程规划的基础上进行的。其工作步骤大致包括：通过区调、找矿或规划工作来圈定远景区或工程地址，然后划分成一系列的勘查区，分阶段按规定的相应精度逐步深入，直至达到查清所需要了解的情况。前一阶段工作为后一阶段工作的设计提供依据。

地矿资源勘查工作的主要技术手段是露头观测、钻探、坑探、物探（包括地震、磁

法、大地电磁法、电法等) 和实验室测试。在一般情况下，其地质数据的来源，主要是岩心描述、测井、采样化验、日常生产记录、水文地质调查、综合研究与编图，此外查阅前人成果也可获得大量数据。下面结合各种类型勘查区的实际情况作些介绍。

1. 岩心编录

岩心编录资料是各种地矿资源勘查工作最重要的基础资料之一。此种来源的数据，项目最多，种类复杂，数量也最大。不同类型勘查区的具体情况会有一些差别，但总的情况是相似的。一般地说，一个金属、非金属、煤炭或水利水电勘查区从普查阶段开始，经详查阶段到精查阶段结束，大约要施工钻孔 100 ～200 个，工程量 10，000 ～100，000 m。按照已有的描述方式，其属性数据的数据量可能接近或超过 20 MB; 如果作为空间数据描述 (野外编录图件、素描图、照片和综合柱状图) ，则其数据量可能超过 100 MB。

2. 水文地质调查

这类数据主要来源于钻孔、泉、井和河、湖的实际观测、试验，也有部分来自气象观测和民间访问。相对而言，这部分的属性数据较少，大约在 10 MB 以下，但空间数据较多，可能达到 n× 100 MB。水文地质调查数据以定量数据为主，可靠性高，使用价值也高，未来的改变不会太大，但会随着时间的推移而不断积累。

3. 样品测试、化验

这类数据主要来源于实验室对各种岩石、矿石、煤、石油、天然气、矿物、化石、地下水等的测试、化验。其中，各项工业分析、元素分析、选矿及工艺性质试验占有很大的比重。这类数据几乎全是定量的属性数据，项目齐全，可靠性最高，未来的变动也将最少。从各类勘查区的现状及未来发展的可能看，其最终数据量可达 10 MB。

4. 测井

测井数据主要是空间数据。各类型勘查区已有的钻孔 (井) 几乎都有测井数据，但 20 世纪 80 年代中期之前多数旧钻孔 (井) 都是以照相方式记录的，目前已在不同程度上采用数字测井仪。照相记录的测井数据，无法直接进入属性数据库，只能通过扫描仪或手扶跟踪数字化仪录入空间数据库中。如果用户需要，也可通过手工方法量出每个层位的峰值，再作为常规属性数据存入属性数据库中，提供用户进行各种查询检索和统计分析。至于数字测井数据，可以直接转储到空间数据库中。一个勘查区的测井数据量通常在 500 MB以上。由于数据量巨大，许多勘查单位采取另建专项数据库的方式来管理。

5. 航空、航天遥感与地面测量

随着信息技术的飞速发展，航空、航天遥感信息对地矿资源勘查的作用越来越大。其用途除了指示地质构造和地层界线之外，还在于辅助进行地矿预测与矿体圈定。航空、航天遥感数据通常以数据磁带和照片 (图像) 两种形式存在。数据磁带可以通过接口将数据直接转储到空间数据库中，而照片 (图像) 则需通过扫描方式转储到空间数据库中。地面测量数据是指各种比例尺的地形图或地理底图。航空、航天遥感和地面测量数据主要是空间数据，属性数据所占比例较小。一个勘查区所涉及的范围有限，航天遥感 (TM 和 SOPT) 和地面测量的数据量通常不会很大，估计其原始数据量有 30 ～50 MB。但是，如果将大比例尺的航空照片包括在内，则扫描数字化后的数据量将达到 300 MB 以上。

6. 综合研究

在勘查设计及勘查总结报告的编写过程中，资料整理、分析、编图、模拟和专题研

究，亦能获得大量新的数据。这类数据一般是由上述各种基础数据综合而成的，例如，勘查区的概况、自然地理条件、地质体与矿体（或煤层）的形态特征、矿体（或煤层）的稳定性、矿区构造特征、某些储量计算参数、矿区地质与成矿规律的各种分析数据和各种相应的分析图件、综合图件，等等。所编制的各种分析图件和综合图件，例如，各种勘探剖面图、平面图、综合柱状图和专项研究图件，大多具有类型繁多、结构复杂、数据量巨大的特点。这些图件都是通过人机交互方式编制的，需要在编绘成功后立即存入空间数据库中，以备本单位经常性使用或向社会提供网络检索服务。如果以矢量化的空间数据形式存贮，其数据总量可达 1 GB 以上。

7. 其他来源

包括日常钻探与测井的生产记录、勘查队的人员编制与生产设备数据。作为生产管理的必要参数，也应当存入数据库中。如果条件允许，可将勘查区的人员编制、生产设备与劳资、人事等合并在一起，另外建立办公自动化系统或企业管理信息系统，效果会更好一些。

二、地矿资源勘查数据的分类

从描述对象的本质特征看，地矿资源勘查数据可以分为属性数据和空间数据两种。属性数据是研究对象的各种特征与性质的定性与定量描述，可以是字符型的，也可以是数值型的；空间数据是研究对象的空间位置及其相互间拓扑关系的表达，可以是字符和数值形式，也可以是图形形式，在机器内可采用栅格数据结构存贮，也可采用矢量数据结构存贮。从表现形式上看，这些数据可划分为字符型数据、数值型数据、日期型数据和图形型数据等 4 类（吴冲龙等，1995）。其中，图形型数据既包括那些观测时直接以图形形式记录下来的数据（例如模拟地震及模拟测井数据），也包括用传统方法绘制的各种成品图件，它们经过栅格化或矢量化以后都可转化为定量数据。从数学性质上看，这些数据可划分为名义型数据、有序型数据、间隔型数据和比例型数据等 4 类（赵鹏大等，1983）。其中，名义型和有序型数据是定性数据，而间隔型和比例型数据是定量数据。在数据的表观形式和数学性质之间存在着一定的包容关系，例如，字符型数据是定性数据的表现形式，也可以是定量数据的概括和归纳；数值型数据首先是定量数据的表现形式，也可以是定性数据的转换形式。

1. 字符型数据

字符型数据是定性数据的表现形式，也可以是定量数据的概括和归纳。它可以用汉字、拼音字母和外文字母的形式书写、存贮和处理，在特殊情况下也可以用数字或数字与字母混合书写、存贮和处理。它包括名义型和有序型两种不同数学性质的数据。

名义型数据　这种数据没有量的概念，只是客观地表达研究对象的某些性质，而不包含相对重要性或相对幅度。例如，地层名称和代号、矿体和煤层的编号、含水层和隔水层的编号、断层名称和编号、岩石和矿物名称、矿石和矿物的用途分类、岩石与矿物的颜色、断层的力学属性、矿体和煤层的形态、沉积体的形态以及可以用是非、有无来表达的各种二态变量等等。这种数据的数据量最大，约占勘查区属性数据总量的 40%。

有序型数据　这种数据相互之间有程度上的差别，而无比例关系。例如，矿物的硬度、碎屑岩的粒度分类、水体混浊度、地下水质量级别、煤的变质程度和煤级、地质构造复杂程度、断层和褶皱的规模和级别、矿床和矿体的规模和级别、矿体和煤层的稳定性、

矿体围岩采掘性能和力学稳定性、勘查类型划分、矿产储量和资源量的类别和级别等等。这种数据量较少，约占属性数据总量的15%。

字符型数据一般是离散型的，易于存入计算机（可通过代码进行信息转换），但不便进行数值运算。只有当设法将其转化或分解为数值型数据时，才能进行数值运算。

2. 数值型数据

数值型数据首先是定量数据的表现形式，也可以是定性数据的转换形式，主要包括间隔型数据和比例型数据两种。数值型数据都是用数字来表达的。

间隔型数据　这种数据的特点是彼此之间不仅有大小和程度之别，而且其差异是相等的，并且没有自然零值。例如，地层产状、钻孔及地质点的坐标与高程、地温、气温与水温等等。这类数据量最少，约占地矿资源勘查区属性数据总量的10%。

比例型数据　这是具有绝对零值的间隔型数据，这种数据不可能有负值存在。它们所反映的数量概念最完整，意义最明确，不仅可以计算出同种数据之差，还可以算出差的倍数。矿体、煤层和地层的厚度、地球化学勘查数据、矿石和矿物的化学成分测定成果，矿石和围岩的物理性质和力学参数测定成果等等，都属于此种。目前这种数据量居第二位，约占地矿资源勘查区属性数据总量的35%。

地矿资源勘查过程中所获取的数值型数据主要是离散型的，但也有连续型的。例如，数字测井的成果就是连续型的。连续型的数据必须事先离散化后才能存入属性数据库，但可以通过扫描方式或手扶跟踪数字化方式直接存入空间数据库。

3. 日期型数据

日期型数据专指那些以三段式字符型描述和存贮的数据，如用于标识日期的年/月/日，用于标识具体时间的时/分/秒，用于标识角度和地理经纬度的度/分/秒等等，都可归入此类。这类数据量较少但很重要，存贮和处理都较为麻烦。目前，一般的数据库管理软件除年/月/日可以进行数值转换处理外，通常当作字符串来整体存贮和调出，否则要先化为十进制数值型数据。一些数据库研发厂商已经开始着手解决这个问题，某些单位和个人也利用字符串的定位截取和识别技术，开发出了相应的数值转换处理程序。

4. 图形型数据

图形型数据原意是指那些观测时直接以图形或图像形式记录下来的数据，例如，模拟地震数据、模拟测井数据、航空照片、地表露头照片和野外素描等等。有些图形，例如，用数字地震和数字测井数据形成的剖面图、曲线图和航天遥感照片，在实际工作中的应用比其原始数值数据本身还要广泛，也归入图形型数据类。另有些图形，例如，历年来地矿资源勘查报告所附的各种图件，由于是采用人工方式编绘的，除了用计算机辅助编绘系统重新编绘的之外，也只能作为图形数据看待。从几何意义上说，图形型数据是一种空间数据。图形数据通常采用手扶跟踪数字化或电子扫描数字化的方式输入，作为数值型空间数据集中存贮和管理，然后用专门的图形编辑软件进行编辑处理，形成可供实际使用的电子图件。如果所输入的每个像元都有定义，亦即经过全面矢量化，则图形处理就比较方便，可任意提取局部信息进行统计、转换、叠加，也可以进行各种空间分析，并且分别成图。

三、地矿数据特征的分析

从以上分析可知，在每一个地矿资源勘查区所获取的资料数据，除了具有来源众多、类型纷繁、数量巨大的特征外，还具有结构复杂、用途广泛的特征。

地质与矿产资源勘查数据结构的复杂性主要表现在数据本身是多元和多维的。一种地质属性要同时用多种数据来描述，这就是地矿资源信息最突出的多元特点，例如，一块岩石标本由多种化学元素、多种化合物和多种矿物组成，因此其性质、成因和含矿性都要用多种因子来描述和判定。这种多元特点对地质矿产数据而言，是一种普遍现象。多元数据在数学处理时通常使用多维空间的概念，地矿数据的多维特点不仅表现在这里，还表现在这种多元数据的多层次交叉叠复上。例如，一个钻孔，具有孔口地理坐标（x，y）和高程（z)，还有钻孔名称、性质、钻探目的、设计深度、终孔深度、孔径尺寸、技术措施、施工质量、施工单位和施工日期等参数，对于一般地理信息系统而言，已经是一种需要进行分层管理的多维空间数据了。但对于地矿资源勘查信息系统而言，这只不过是一些简单的一般性描述参数。因为更重要、更复杂的地矿信息还在钻孔岩心和露头地质描述数据中。一个钻孔有几百乃至几千层岩石，每一层岩石都有岩性、岩相、结构、构造、矿物组成等描述参数，如果是含矿层或矿体，还要采样化验其矿物共生组合、元素组成等少则几种多则几十种参数，相互之间既成层次关系又成网络关系，错综复杂。

地质矿产勘查数据的又一个显著特点是其用途广泛，这些数据既要用于矿体及矿床特征分析、控矿条件及成矿过程分析、外围及深部含矿性预测，又要用于储量计算和地质、技术、经济、环境条件的综合评价，此外，还要编制数十幅乃至数百幅各种各样的勘查图件，因此存在着多个处理主题。每一个主题都要涉及大量不同类型的数据，而每一种数据都可能有多种用途，亦即属于多个主题，数据流向复杂，数据处理过程烦琐。

地矿数据的上述特征归纳起来就是“5多”，即多源、多类、多维、多量和多主题。只有采用“多S”结合与集成策略，才能有效地进行这类数据的采集、存储、管理和处理。

四、代码的分类与选择

代码是用来代表事物某种属性的一组有序的字母-数字串（简称为字符串)。具体地说，代码可用来代替某一个名词、术语，甚至某一个特殊的描述短语。为了便于操作，提高查询、检索的效率和增强数据的共享性，必须使地矿资源勘查信息系统中的字符型数据全部代码化。代码的功能主要是鉴别、分类和排序三种。其中，鉴别功能是最基本的，也是必备的。具备这三方面的功能，信息系统就具备共享性了。共享性既是数据库设计的重要目标之一，也是整个地矿资源勘查信息系统设计的重要目标之一。这里所说的共享性，主要包括属性名的共享性、关系结构的共享性和属性域的共享性。为了实现这一共享性，在系统内，一个代码只能唯一地标识一个属性；反之，一个属性只能有一个标准化代码。

词语代码的分类一般以结构为准，所谓的代码结构是指代码内部各个符号或符号组之间的关系。在地质与资源勘查中常用的有5种，即特征组合结构、顺序结构、助记结构、分类结构和混合结构。其代码的种类也相应地分为特征组合码、顺序码、助记码、分类码和混合码5种。这5种代码体系都具有鉴别、分类和排序功能，但适用范围和应用条件有所不同。

1. 特征组合码

特征组合码由各具代表意义的几位字母或数字排列组合而成。通常是一个字母或数字代表描述对象的一种属性特征。例如，用三位数字来表示岩石和矿物的颜色:

第一位代表色调的深浅:“0”不清楚，“1”浅色，“2”正常色，“3”深色;

第二位代表配色:“0”不清楚，“1”红色，“2”黄色，“3”褐色，“4”绿色，“5”

蓝色，“6”紫色，“7”灰色，“8”白色，“9”黑色；

第三位代表主色：“0”不清楚，“1”红色，“2”黄色，“3”褐色，“4”绿色，“5”蓝色，“6”紫色，“7”灰色，“8”白色，“9”黑色。

于是浅灰绿色可表示为“174”，深褐黄色可表示为“332”，其余类推。

特征组合码通常被用于各个单位和个人所开发的应用数据库和数据、图型处理系统中。其优点是简单明了，易记易用；缺点是位数太少时容量有限，易发生重码，而位数多时又不易掌握，容易产生混乱。最大的缺点是难于进行标准化处理，不可能以此为基础实现信息共享，在主题数据库和信息检索系统（网络）中不宜独立使用。

2．顺序码

顺序码主要由按顺序排列的数字组成，有时也由按顺序排列的字母组成。每一组代码只代表描述对象属性在整个属性系列中的顺序。在地矿资源勘查信息系统中最典型的例子是岩层序号，从实测剖面起点开始由头至尾累计，或从钻孔孔口开始由上而下累计。顺序码也常用于描述某些并列属性，在这种情况下，码的顺序只有逻辑含义，而无序次的含义。作为代码的字符串，通常要求长度统一，因此在编码时应当事先估计可能出现的最大长度，然后确定代码的位数。由于顺序码最为简单明了，在各种信息系统中使用十分频繁，几乎每一种代码体系都有它的影子，但都只有局部性意义。由于本身的特点和弱点，顺序码在地矿资源勘查信息系统中不可能形成独立的代码体系。

3．助记码

助记码是描述对象属性的一种形象化代码，比较容易记忆。最常见的是采用描述对象属性的汉字拼音或外文单词的首位字母来组合。国际通用的化学元素符号和地层代号都是这样构成的。沉积岩岩相的代码也常用这种方式编制，例如，具有大型槽状交错层理的含砾砂岩，英文名称为 Large Scale Trough Cross Bedding Gravel-bearing Sandstone，首位字母可组合成“LSTCBGS”。但助记码通常规定有统一的长度，不能任意加长，凡是超过位数的都要进行删减。删减的原则是，每一个描述方面（一个定语或一个分词组）留下一个字首。如果岩相的代码只取四位，则根据上述岩相英文名的组合情况，Large Scale 可取 L，Trough Cross Bedding 可取 T，Gravel-bearing 可仍取 G，Sandstone 也可仍取 S。于是，该岩相的助记码可写成“LTGS”。

助记码的缺点是容量有限，很容易发生重码、乱码现象。对于某些单位和个人使用的微型数据库，助记码用起来很方便；但对于具有共享性质的地矿资源勘查区信息系统，特别是对于准备纳入信息系统网络的大型地矿资源勘查区信息系统，应当谨慎使用。

4．分类码

分类码是按照事物的属性类别划分数据集合的代码体系。这种代码体系主要包括面分类码和线分类码两个亚型，常为大型信息系统所采用。

面分类码　就是采用面分类法建立起来的代码系统。面分类法是将给定的母项按属性分成互不依赖的集合——面分类法。面分类没有刚性的分类结构和预先确定的最终集合，其中，对象的母项是以一些同一属性的对象的组配来表征的。这些属性的值对应着指定属性的具体描述，并且可以根据需要分解成很多互无联系的分面。

在面分类代码体系中，分类的集合由具体属性的组配对象构成。这些属性取自某个相应的面。面分类代码体系的容量同面的数目以及在面里具体属性的数量有关。面分类代码

体系的主要优点在于其结构具有柔性，因为任何一个面的改变，不会对其余的面产生本质的影响。此外，这种分类方法不仅能够从已有的面中形成新的分类集合，而且还可以不加改造地将新的面列入分类目录，或者从中除去旧的面。面分类码的柔性使对象的聚合有了可能，并且可以实现按任意组配面进行的信息检索。面分类法的不足之处在于不能充分利用代码体系自身的容量，因为许多可能组配的面实际用不上。

线分类码　属于层次分类码。其划分方法为：先在被划分的对象母项中选定若干个属性作为划分标准，组成一些大的集合，接着在每个大集合中相应选定划分标准，组成一系列小集合，然后，再分解更小的集合，如此一层层地分解下去，直到具体对象。这样，在各大、小集合之间，就形成了线性的隶属关系。线分类码的突出优点是结构清晰、易成体系、容量较大；缺点是当被描述对象的属性分类层次过多时，码的长度需要约束，而当各种属性分类层次参差不齐时，需要借用多种编码方式来构成混合码，才能完整地表达。作为线分类码的基础，属性的选择应当着眼于地矿资源勘查信息系统的整体，具体地说应当着眼于整个国土资源和行业。只有这样，所建立的分类码才具有普适性和共享性。

5. 混合码

混合码是将特征组合码、顺序码、助记码和分类码按一定的规则联结起来的代码形式。我国技术监督局所颁布的《地质矿产名词术语分类代码》国家标准（GB9649—88），就是一种以分类码为格架的混合码体系。该标准有三个分册、共 12 万多条地矿术语及编码。

该代码系统将代码分为数据项和文字值（字符值）两个层次，其中，数据项代码由 6 个英文字母（严格地说是 2 个拼音字母和 4 个英文字母）组成，文字值代码由 1 ～8 个阿拉伯数字组成。该代码体系规定，数据项代码的第一、第二位是地质学二级学科名前两个字的汉语拼音缩写，兼有面分类码和助记码性质，例如，岩石学取 YS、矿床学取 KC、煤地质学取 MD、石油地质学取 SY……。从第三位开始都是按 A ～Z 顺序排列的英文字母，其中第三和第四位通常是二级学科内的层次分类，兼有线分类码和顺序码的性质，例如 YSA 是岩石学序言、YSB 是岩石成分、YSC 是岩石结构等等；第五位代表组合数据项或数据项，也兼有线分类码和顺序码的性质，例如，YSBCA 代表碎屑颗粒、YSBCB 代表碳酸盐异化粒、YSBCD 代表碎屑岩胶结物；第六位必定是数据项，例如，YSBCBA 代表碳酸盐异化粒的内碎屑、YSBCBB 代表碳酸盐异化粒的球粒等等，通常具有顺序码特征。有些二级学科内层次分类较少且最后一层数据项的量又过大，也有采用第五、第六位顺序组合的编码方式，即由第五、第六位联合起来描述一个具体的属性。而有些二级学科内的层次分类很少，也有从第四位就采用属性的顺序码的。

字符值代码可长可短，通常只用 1 ～2 位阿拉伯数字的顺序码。但是，在某些二级学科中，由于内部层次结构复杂，文字值又太多，也有采用 4 ～8 个阿拉伯数字作为补充混合码的情况。例如，矿物学、岩石学和古生物学都是如此。在矿物学代码中，头两位码代表矿物的化学类型，例如，01 是单质、02 是碳化物-硅化物-氮化物和磷化物、06 是氧化物；第三位代表矿物的晶体格架类型，例如，011 是单质配位基型、013 是单质环状基型、015 是单质链状基型等等；第四位代表晶体结构的复杂程度，例如，0611 是简单的配位基型氧化物、0612 是复杂的配位基型氧化物；第五、第六位代表矿物的族分类，例如，061101 是简单的配位基型氧化物的方铈石族、061102 是简单的配位基型氧化物的斜锆石族。以上六位码均兼有线分类码、特征组合码和顺序码的性质。第七、第八位码代表具体

矿物种，例如，06110101 是简单的配位基型氧化物方铈石族的方铈石矿、06110201 是简单的配位基型氧化物斜锆石族的斜锆石矿。这末尾的两位码，均带有联合顺序码性质。

混合码的最大优点是容量大且唯一性好，因此使用国家标准是各种地矿资源勘查信息系统的最佳选择。在各部门组建地矿资源勘查信息系统时，可以根据实际需要进行扩充，但必须遵循这一标准编码体系，才能最大限度地保证信息的共享性。

第三节　地矿资源勘查实体模型研究

实体模型是信息系统的概念模型和数据模型的基础，用于概括地矿资源勘查工作的各项事务之间的内在联系、作业流程及其相应的信息处理内容。地矿实体模型研究的主要任务是定义模型的基本类型、确定描述这些实体模型的数据项及流程、查明各子实体模型间的联系。地矿实体模型的研究通常与用户调查、数据现状分析同时并举、分工进行。

为了满足地矿资源勘查信息系统的建设的实际需要，实体模型的研究必须面对地矿资源勘查工作的基本任务。从目前情况看，地矿资源勘查工作的基本任务有两个，其一是查明工作区的地质与矿产状况，并对其作出经济技术条件的综合评价，为开发利用服务；其二是研究工作区地质演化与资源的形成、分布规律，为进一步的地矿预测、找矿勘查或建设规划服务。每一项具体的工作任务都将形成一个信息管理和处理的主题，而每一个信息管理和处理的主题都将概括出一个实体子模型。汇总和综合多个子模型，便可以形成总实体模型。

各实体模型及其子模型的设计可以独立进行，但在设计过程中应当吸取用户调查提供的信息，并且将它们与目前生产实践的需要及将来发展的可能结合起来，不断地进行修改。子模型设计均采用由下而上的方法，即先从实际勘查工作、成矿预测、盆地分析、工程选址、地质灾害监测和矿产资源可利用性综合评价等项工作所要求解决的问题出发，研究其可能涉及的工作和资料，确定其实体（对象）和属性，然后逐级向上综合。

下面给出金属非金属、煤炭、石油天然气勘探和工程地质勘察等 4 个模型的实例。这些实例是通过若干实际系统分析而建立的，并非一种固定格式和标准模型。它们应当随着勘探（勘察）任务、对象、任务和技术的发展而变化。

一、金属非金属矿产资源勘查模型

1. 矿床地质勘查子模型

矿床地质勘查的基本任务，是利用各种技术手段查清勘查区金属或非金属矿产资源的赋存状况、储量以及开采利用的地质条件、技术条件，为下阶段的勘探或建井的设计、施工提供所需要的地质资料。矿床地质勘查工作的主要内容包括构造地质研究、地层与建造研究、岩浆岩体研究、矿石结构构造研究、矿体（层）稳定性研究、水文地质条件研究、围岩或顶底板物理力学特征研究、地理条件研究、经济价值和环境污染研究等等。显然，矿床地质勘查子模型应当准确而完整地概括这些基本工作任务和基本工作内容。

(1) 整体结构

图 3-2 表示矿床地质勘查子模型的整体结构。

其中勘查技术手段与地质特征均包含着许多内容，从数据管理系统的角度看来，都是描述对象或实体的集合，简称为对象集（或称为实体集），相互之间为多对多关系（记为 M-M），并且通过原始资料采集发生联系。原始资料采集也包含着许多方法和内容描述对象，

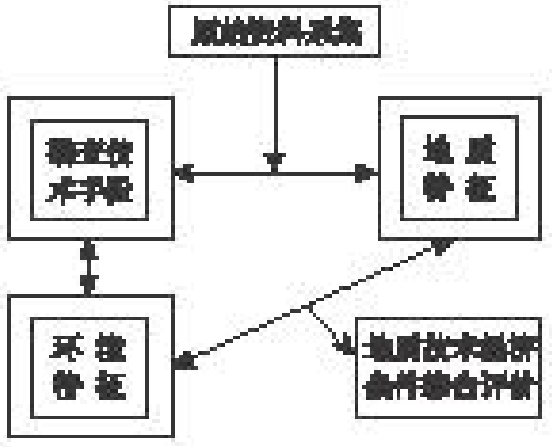


图 3-2　矿床地质勘查子模型的整体结构

同样是一个描述对象集（实体集)。地质特征与环境特征之间也具有多对多关系，二者通过地质技术经济条件综合评价联系起来，而综合评价本身也是一个描述对象集（实体集)。

(2) 描述对象集

矿床地质勘查子模型大致可分为下列 5 类子描述对象集（或称为子实体集、属性集)：

勘查区技术手段属性集　主要包括：①技术手段的种类和性能（地质填图、坑探、钻探、测井、物探、化探、大地测量)；②不同勘查技术手段的作用和应用范围；③影响勘查技术手段选择的因素（勘查阶段，勘查地区，矿床类型)；④勘查技术手段的组合方案；⑤综合技术手段的效果评价。

原始资料采集属性集　主要包括：①原始资料的类型与信息源（地面观测点、钻孔、坑道、物探、化探)；②各种原始资料采集的方式与方法；③各种原始资料采集的规范化与标准化；④各种原始资料的筛选、优化方法；⑤各种原始资料的预处理与定量化。

勘查区地质特征属性集　主要包括：①矿床形成的地质条件（区域地质、矿带地质和矿区地质)；②矿床地质和地球化学特征；③矿床、矿体的自身特征（类型、形态、规模、分布、组合、产状，等等)；④成矿系列（矿床系列）与成矿体系；⑤矿石和矿化的特征（成分、品位、结构、构造、不均一性与分带性，等等)；⑥主矿种储量与伴生矿种储量。

勘查区环境特征属性集　主要包括：①地形地貌特征；②侵蚀切割程度和第四系覆盖厚度；③气候条件与气象特征、生物-气候分带特征；④水系发育情况、流域面积、水文地质复杂程度和水动力特征；⑤构造破碎、断裂带、溶洞的发育程度和充填程度；⑥矿体（层）顶底板岩稳固程度；⑦崩塌、滑坡、泥石流和历史地震活动情况；⑧环境污染现状与发展趋势。

地质技术经济条件综合评价属性集　主要包括：①矿床地质参数（矿床类型、规模、成矿序列、成矿体系、矿床与矿石特征、品位、储量、矿石可选性可冶性、伴生有益有害组分，等等)；②政治经济和自然地理参数（国家需求状况、资源政策、环境政策，矿床内外部建设条件，等等)；③经济参数（矿产品价格、投资、成本以及利息率，矿山服务年限，等等)；④矿山主导参数（矿山生产规模、边界品位、金属回收率和开采损失率、贫化率等)；⑤开采技术参数（包括构造复杂程度、构造破碎、断裂带、溶洞的发育程度和充填程度、矿体（层）顶底板岩石稳固程度、矿体（层）的埋藏深度、矿体（层）的倾斜程度、涌水量、地温等等)。

2. 成矿预测子模型

成矿预测是一项在矿区范围内通过对控矿地质因素、矿化信息、成矿规律进行全面分析研究，从而对矿床及其深部和外围的金属矿产资源作出综合预测和评价的系统方法。成矿预测子模型是这项工作的归纳和概括。

(1) 整体结构

成矿预测的基本结构参数可以大致地概括为 6 大类，即构造参数、岩浆岩参数、沉积

参数、变质（包括成岩）参数、古地热参数和成矿参数。这6类基本参数均由许多具体的参数组成，从数据管理系统的角度看来，即为6个实体集。每个实体集之间基本上是多对多关系（图3-3）。每两个实体集之间分别由综合分析工作关联起来。这些综合分析本身，亦为描述对象集（实体集），即为更高层次综合分析的子实体集。

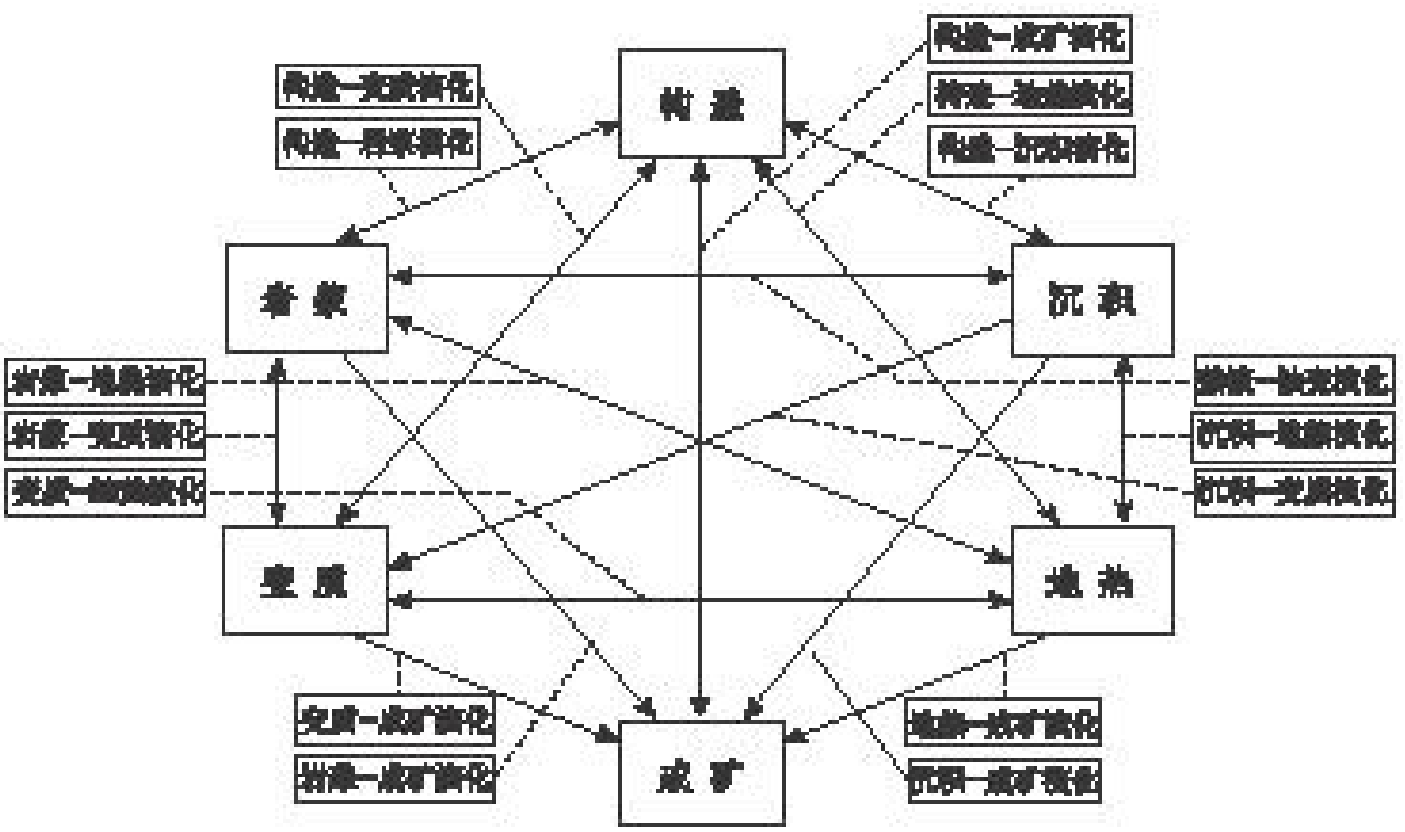


图3-3　矿区成矿预测子模型的整体结构

(2) 描述对象集

成矿预测涉及的参数众多，因此其描述对象集的构成十分丰富。根据目前国内外成矿预测实践，描述对象集常见的参数子集大致有21种。每一种参数子集都由一系列属性组成，而多数属性都可再分解为一些次级属性。

构造参数属性集　主要包括：①基础参数：区域构造特征、矿区构造特征；②分析参数：构造格架、构造应力场、构造演化。

岩浆岩参数属性集　主要包括：①基础参数：区域岩浆岩特征、矿区岩浆岩特征；②分析参数：岩浆岩格架、岩浆旋回、岩浆演化。

沉积参数属性集　主要包括：①基础参数：区域沉积特征、矿区沉积特征；②分析参数：地层格架、沉积相、沉积演化。

变质参数属性集　主要包括：①基础参数：区域变质特征、矿区变质特征；②分析参数：变质岩格架、变质相、变质演化。

地热参数属性集　主要包括：①基础参数：区域地热特征、矿区地热特征；②分析参数：地热梯度、地热场、地热演化。

成矿参数属性集　主要包括：①基础参数：区域成矿特征、矿区成矿特征；②分析参数：成矿条件、矿源岩（层）、成矿演化。

构造-岩浆演化属性集　主要包括：①不同构造发育区的岩浆岩分布和演化特征；②不同岩浆岩分布区的构造演化特征；③岩浆岩分布对构造演化的影响；④构造演化对岩浆

作用的控制或影响；⑤岩浆作用的演化对构造演化的影响。

构造-沉积演化属性集　主要包括：①不同构造发育区沉积岩分布和演化特征；②不同沉积岩分布区构造演化特征；③沉积岩分布对构造演化的影响；④构造演化对沉积作用的控制或影响；⑤同沉积构造特征及分布；⑥同沉积构造运动性质与特征；⑦同沉积构造演化。

构造-变质演化属性集　主要包括：①不同构造发育区的变质岩分布与演化特征；②不同变质岩分布区的构造演化特征；③变质岩分布对构造演化的影响；④构造演化对变质作用的控制或影响；⑤变质作用的演化对构造演化的影响。

构造-地热演化属性集　主要包括：①不同地质时代的构造热事件特征；②不同构造发育区的地热场特征及演化；③不同地热演化区的构造特征及深部背景；④构造演化对地热场的控制或影响；⑤地热场特征及演化对构造演化的影响。

构造-成矿演化属性集　主要包括：①不同构造发育区的矿体或矿床特征；②不同矿体或矿床构造特征；③构造演化对成矿作用的控制或影响；④成矿作用及演化对构造演化的影响。

岩浆-变质演化属性集　主要包括：①不同岩浆岩发育区的变质岩分布与演化特征；②不同变质岩分布区的岩浆岩分布与演化特征；③岩浆活动对变质作用的控制或影响；④变质作用的演化对岩浆活动的影响；⑤岩浆接触交代作用特征。

岩浆-地热演化属性集　主要包括：①不同地质时代的岩浆热事件特征；②不同岩浆岩发育区的地热场特征及演化；③不同地热演化区的岩浆作用特征及深部背景；④岩浆热液的物理化学特征；⑤岩浆作用对地热场的控制或影响；⑥地热场特征及演化对岩浆作用的影响。

岩浆-成矿演化属性集　主要包括：①不同岩浆岩发育区的矿体或矿床特征；②不同矿体或矿床的岩浆岩的围岩特征；③岩浆作用演化对成矿作用的控制或影响。

沉积-变质演化属性集　主要包括：①副变质岩的原岩特征；②沉积物埋藏史；③沉积物成岩作用史。

沉积-地热演化属性集　主要包括：①不同地质时代的岩浆热事件特征；②不同沉积岩层位或分布区的热导率及地热场演化；③不同地热演化区的沉积作用及沉积岩特征；④沉积作用对地热场的控制或影响；⑤沉积物热演化史与岩层有机质成熟史；⑥沉积水(沉积成因的热流体)物理化学特征；⑦地热场特征及演化对成岩作用的影响。

沉积-成矿演化属性集　主要包括：①不同沉积岩层位或发育区的矿体、矿床特征；②不同矿体或矿床的沉积岩的围岩特征；③沉积型矿源层及其成矿条件分析；④沉积作用对成矿作用的控制或影响；⑤沉积水(沉积成因的热流体)的成矿作用演化。

沉积-岩浆演化属性集　主要包括：①不同时代火山-沉积岩分布与演化；②不同沉积岩层位或发育区的岩浆岩(包括火山岩)分布和演化；③同沉积期岩浆(火山)作用及演化；④海底喷流(气)与含矿热液演化；⑤沉积环境及其演化对火山作用的影响；⑥火山作用及其演化对沉积作用的影响。

变质-地热演化属性集　主要包括：①不同地热演化区的变质作用及变质岩特征；②不同变质岩层位或分布区的热导率及地热场演化；③变质(热)流体的物理化学特征；④地热场特征及演化对变质作用的影响。

变质-成矿演化属性集　主要包括：①不同变质岩层位或发育区的矿体、矿床特征；②不同矿体或矿床的变质岩的围岩特征；③矿源层及其成矿条件分析；④变质作用对成矿作用的控制或影响；⑤变质（热）流体的成矿作用演化。

地热-成矿演化属性集　主要包括：①不同地热演化区的矿体、矿床特征；②不同矿体或矿床的地热场演化特征；③岩层有机质的热演化史；④各种类型热流体（地质流体）的复合成矿演化史；⑤地热场对成矿作用的控制或影响；⑥成矿作用对岩层热导率和地热场的影响。

3. 金属、非金属矿产资源勘查评价工作流程

综合上述各个实体模型，便可以得到金属、非金属矿产资源勘查评价的总实体模型。

在金属、非金属矿产资源勘查实践中，资料的分析与综合通常分为若干个层次。各种基础参数在分析过程中逐步引入，经分析与综合后逐步形成各层次的综合参数和结论。从某种意义上说，基础参数的数据采集、引用和分析、综合过程，就是金属、非金属矿产资源勘查、预测和评价过程。这一过程可以用下列流程图（图 3-4）来表示。

二、煤炭资源勘查模型

1. 煤炭资源勘查子模型

煤炭资源勘查的基本任务，是利用各种技术手段查清勘查区煤炭资源的赋存状况、储量以及开采利用的地质条件、技术条件，为下阶段的勘探或建井的设计、施工提供所需要的地质资料。煤炭资源勘查工作的主要内容包括构造地质研究、煤质煤化研究、煤层稳定性研究、地层古生物研究、水文地质条件研究和煤层顶底板物理力学特性研究等等。煤炭资源勘查子模型应当准确地概括这些基本工作任务和工作内容。

(1) 整体结构

由于固体矿产的勘查方法及技术手段在总体上极为相似，煤炭资源勘查子模型的整体结构类似矿床勘查（图 3-2）。

(2) 描述对象集

煤炭资源勘查子模型的描述对象集也可分为下列 5 类属性集。

勘查区技术手段属性集　主要包括：①技术手段的种类和性能（区域地质测量、坑探工程、钻探工程、测井工程、地面物探工程、大地测量工程）；②不同勘查技术手段的作用和应用范围；③影响勘查技术手段选择的因素（不同勘查阶段，不同勘查区，不同矿床类型）；④勘查技术手段的最佳组合方案；⑤综合技术手段勘查矿床的效果评价。

原始资料采集属性集　主要包括：①原始资料的类型与信息源（地面观测点、钻孔、坑道、物探）；②各种原始资料采集的方式与方法；③各种原始资料采集的规范化与标准化；④各种原始资料的筛选、优化方法；⑤各种原始资料的预处理与定量化。

勘查区地质特征属性集　主要包括：①地质构造特征；②含煤岩系特征；③水文地质特征；④煤层特征；⑤煤质特征；⑥共生伴生矿产特征；⑦开采技术条件；⑧储量特征。

勘查区环境-地理特征属性集　主要包括：①勘探区地理位置；②勘探区自然地理特征（地形地貌、侵蚀切割程度和第四系覆盖厚度、气候条件与气象特征、水系发育情况、流域面积，等等）；③勘探区经济地理特征（工农业、交通、电力、生物-气候分带、植被、土壤）；④地质灾害状况（崩塌、滑坡、泥石流和历史地震活动情况）；⑤环境（水质、大气、土壤）污染现状与发展趋势。

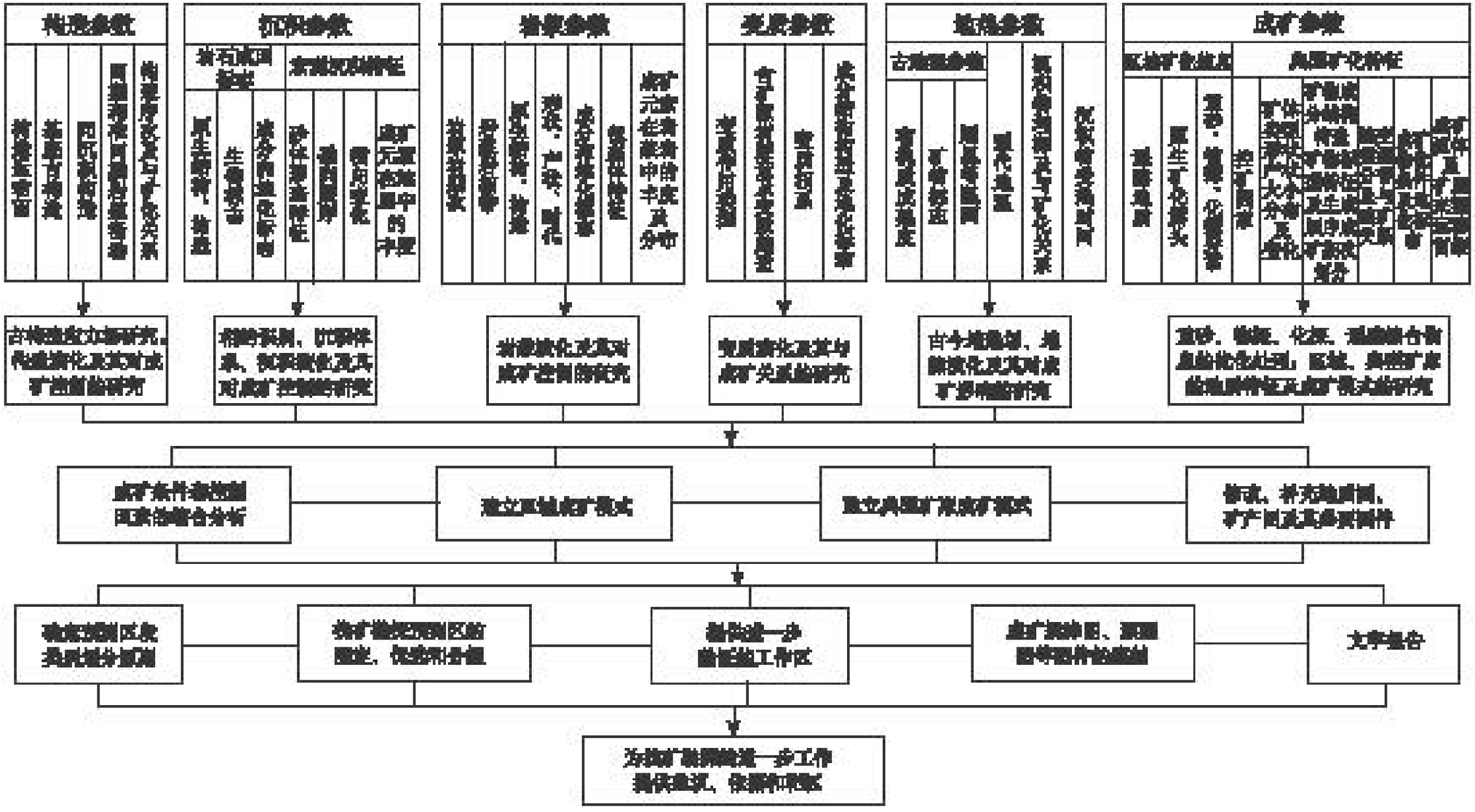


图 3-4　金属、非金属矿产资源勘查评价工作流程

（据吴冲龙等，1996）

地质技术经济条件综合评价属性集　主要包括：①煤田地质参数（煤种、储量、灰分产率、含硫量、发热量和挥发分含量，煤的可选性、伴生有益有害组分）；②政治经济和自然地理参数（国家需求状况、能源政策、环境政策，煤矿内外部建设条件，与经济区和中心城市的距离）；③经济参数（煤产品价格、投资、成本以及利息率，矿山生产规模、回采率和服务年限）；④开采技术参数（构造复杂程度、断裂带、溶洞的发育程度和充填程度、矿体（层）顶底板岩石稳固程度、煤层的埋藏深度、煤层的倾斜程度、瓦斯、涌水量、地温）。

2. 聚煤盆地分析子模型

聚煤盆地分析是将原生聚煤盆地看作一个整体，对其中的构造、沉积物和煤及其他伴生沉积矿产的形成、分布、演化进行全面研究，从而对煤炭资源作出综合预测和评价的系统方法。其基本原理可以概括为整体分析、关联分析、动态分析和反馈控制分析。

（1）整体结构

聚煤盆地分析的基本结构参数包括4类，即沉积参数、构造参数、聚煤参数和古地热参数。这4项基本参数即为4个实体集。每个实体集之间为多对多关系（图3-5）。每两个实体集之间分别由综合分析工作联系起来。

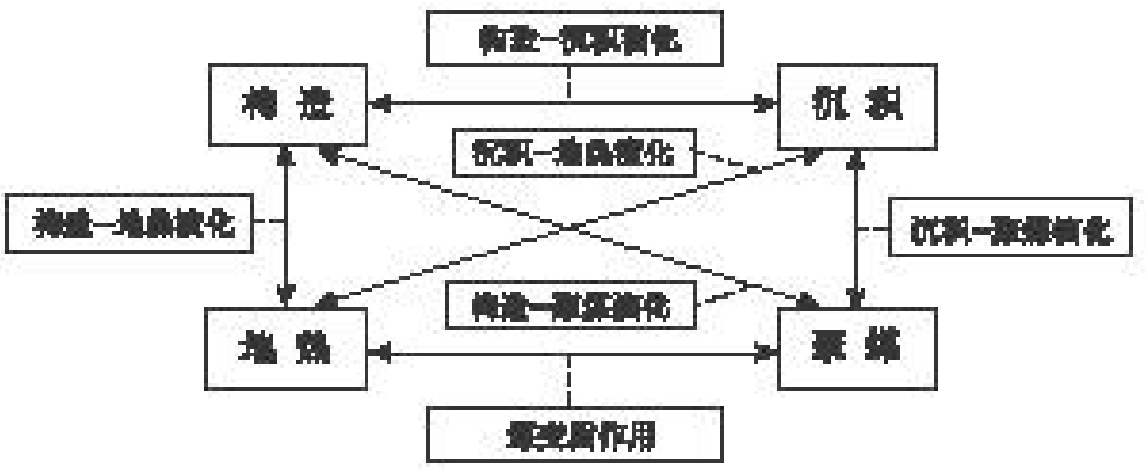


图3-5　聚煤盆地分析子模型的整体结构

（据吴冲龙等，1996）

（2）描述对象集

聚煤盆地分析子模型的描述对象集大致由10类可分属性组成。

构造参数属性集　主要包括：①基础参数：区域构造特征、盆地构造特征；②分析参数：构造格架、构造应力场、构造演化。

沉积参数属性集　主要包括：①基础参数：区域沉积特征、盆地沉积特征；②分析参数：地层格架、沉积相、沉积演化。

聚煤参数属性集　主要包括：①基础参数：区域聚煤特征、盆地聚煤特征；②分析参数：聚煤条件、富煤带、聚煤演化。

地热参数属性集　主要包括：①基础参数：区域地热特征、盆地地热特征；②分析参数：地热梯度、地热场、地热演化。

构造-沉积演化属性集　主要包括：①同沉积构造与同沉积构造运动；②沉积作用对同沉积构造运动的响应；③构造作用对沉积作用的控制。

沉积-聚煤演化属性集　主要包括：①煤层及其顶底板的沉积特征；②煤层及其顶底板的沉积环境；③沉积条件对煤层聚集的控制或影响。

煤变质作用属性集　主要包括：①煤质特征与古地热场；②煤变质作用成因分析；③煤变质作用历史分析。

构造-地热演化属性集　主要包括：①不同构造分区的地热场特征；②不同构造演化阶段的古地热场特征；③构造热事件演化史；④地幔热动态演化史。

构造-聚煤演化属性集　主要包括：①不同构造分区的聚煤特征；②不同构造演化阶段的聚煤特征；③煤层及其顶底板的同沉积构造特征；④煤层及其顶底板的古构造应力场；⑤同沉积构造对聚煤作用的控制或影响。

沉积-地热演化属性集　主要包括：①不同沉积分区的热导率与地热场特征；②不同沉积演化阶段的古地热场特征；③岩层有机质的热演化史。

(3) 子实体模型和工作流程

聚煤盆地分析的子实体模型是聚煤盆地分析子模型整体结构的细化。

将聚煤盆地分析的上述描述对象集与盆地分析整体结构组合起来，便可以得到充实了的聚煤盆地分析子实体模型（图 3-6）。从这个子实体模型中，可以得到对聚煤盆地分析工作的一个较为完整的认识。由于图 3-6 中所表示的内容是从多个聚煤盆地分析工作中总结出来的，因此实际上应当称为聚煤盆地分析概念模型。在对具体盆地作分析时，应当根据实际情况具体地制定实体模型，以便从中归纳和抽提出概念模型和数据模型。

与金属，非金属矿产资源勘查评价工作相似，聚煤盆地分析也涉及大量煤炭资源勘查资料的分析与综合。这种分析与综合也可分为若干个层次。各种基础参数在分析过程中逐步引入，逐步形成各层次的综合参数和结论。因此，聚煤盆地分析过程在实践中可以看作区域性（盆地范围内）煤炭资源预测、勘查和评价过程。这一工作过程较为复杂，但同样可以用一个流程图（图 3-7）来表示。

三、油气资源勘查评价模型

油气资源勘查评价模型包含着许多子模型，例如油气资源勘探、含油气盆地分析、油气系统分析、圈闭评价、油气成藏动力学分析、盆地模拟、油气成藏模拟，等等。本书只简单介绍油气资源勘探、含油气盆地分析和油气系统分析等三个子模型。

1. 油气资源勘探子模型

油气勘探的基本任务是根据石油地质和油气勘探理论，综合运用各种勘探技术和方法，寻找和评价油气远景区带，发现具有工业价值的油气田，并探明其地质储量，取得开发评价所需的各项数据，为油气田的开发准备储量和资料。油气田勘查工作一般是分阶段进行的，主要分为盆地普查、区带详查、圈闭预探和评价勘探等四个阶段（胡朝元，1990）。

(1) 油气资源勘探子模型整体结构

油气资源勘探总实体集中也包含各子实体集（图 3-8）。其中勘查技术手段与盆地地质特征，盆地地质特征与盆地外部环境特征均为多对多关系。勘查技术手段与盆地地质特征间通过“原始资料采集”发生关系，盆地地质特征与外部环境特征间通过“地质技术经济条件综合评价”发生关系。这两种关系也都是描述对象集（子实体集）。盆地地质特征总对象中包含一系列下级个体对象，这些个体对象（子实体）间也是多对多的关系。

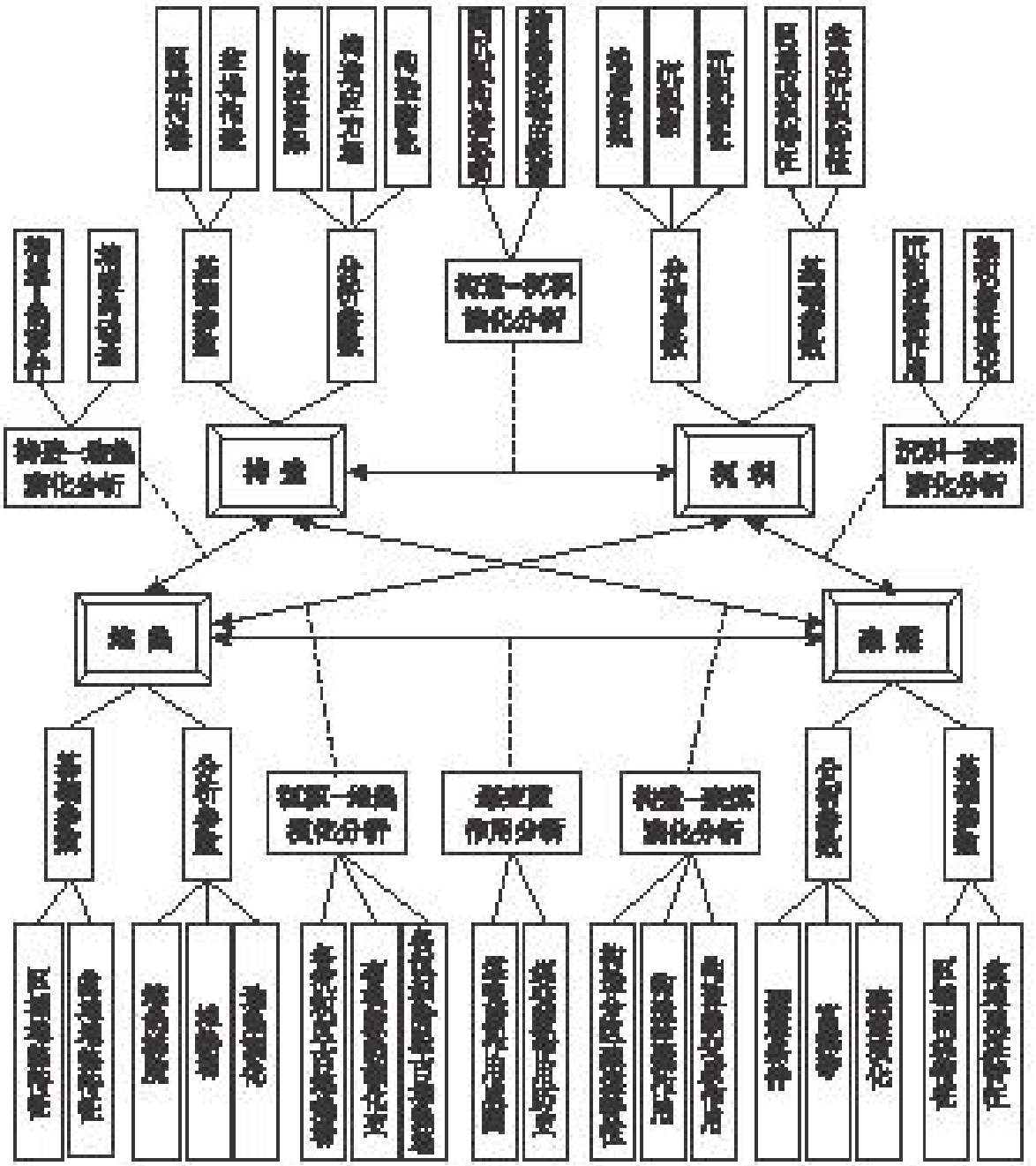


图 3-6　聚煤盆地分析实体模型

(据吴冲龙等，1996)

由于各个实体的复杂性，需要进一步分析其包含的子实体及其属性。

(2) 描述对象集

油气田勘查评价模型包含着如下 5 个描述对象集，即 5 个高级别的属性集:

盆地地理-环境特征属性集　主要包括: ①盆地地理位置; ②盆地自然地理特征; ③行政区划; ④盆地勘探史。

油气勘探技术手段属性集　主要包括: ①区域地质调查; ②地球物理勘探; ③地球化学勘探; ④油气钻探; ⑤分析化验与测试。

原始资料采集属性集　主要包括: ①原始资料的类型与信息源; ②各种原始资料采集的方式与方法; ③各种原始资料采集规范化与标准化; ④各种原始资料的预处理与定量化。

盆地地质特征属性集　主要包括: ①沉积与地层; ②盆地构造与构造应力场; ③盆地地热场; ④有机地球化学; ⑤盆地演化历史; ⑥油气系统与油气成藏动力学参数。

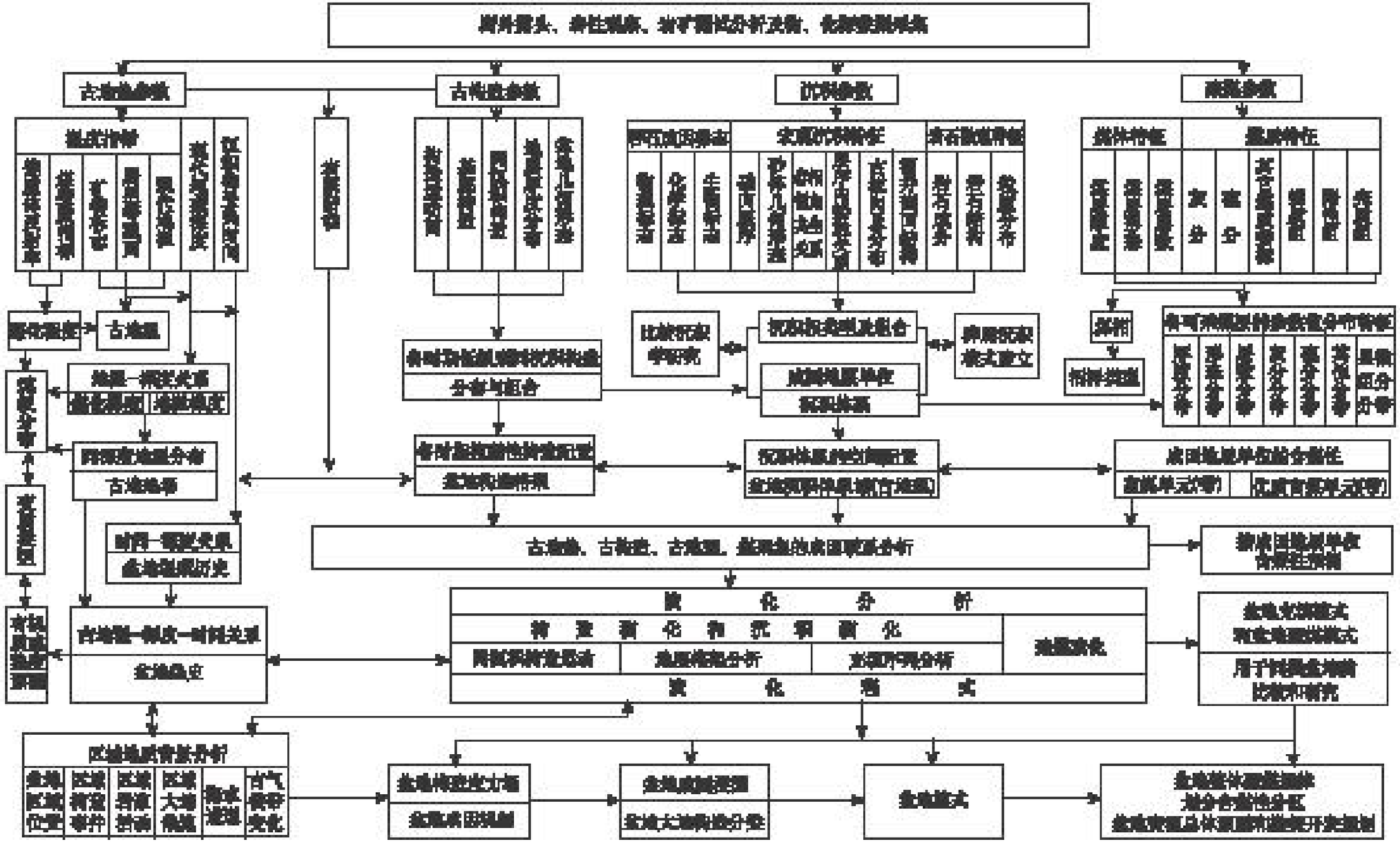


图 3-7　聚煤盆地分析工作的一般流程

(引自吴冲龙等, 1996)

地质技术经济条件综合评价属性集　主要包括: ①油气地质参数; ②政治经济 (国家需求状况、国际油价、能源政策、环境政策等等) ③自然地理参数 (地貌、气候、降水等等); ④开发技术参数 (油藏深度、井底压力、含水量、开采方法等等); ⑤技术经济条件综合评价方法 (主观概率法、模糊综合评价法、系统动力学评价法)。

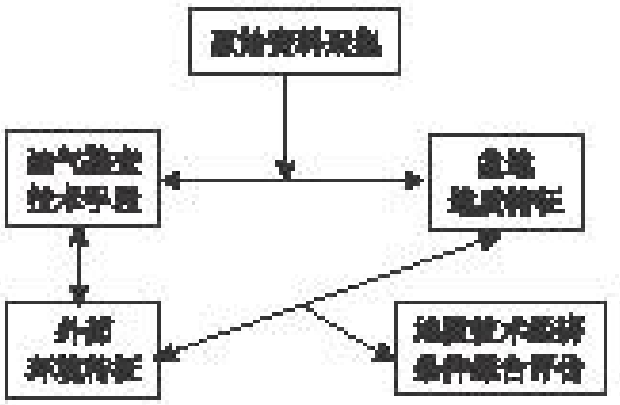


图 3-8　油气田勘查评价模型整体结构

(3) 总实体模型和工作流程

总实体模型　在上述概略模型的基础上, 对每个描述对象 (子实体集) 进行细分, 可得到油气勘探的总实体模型 (图 3-9)。其中一些子实体还可以进一步细分, 例如, 由于每一种勘探技术手段可以获取一种或数种勘探参数, 油气勘探技术手段子实体集中的子实体可进一步作如图 3-10 的细分。细分后的参数构成各实体的属性。

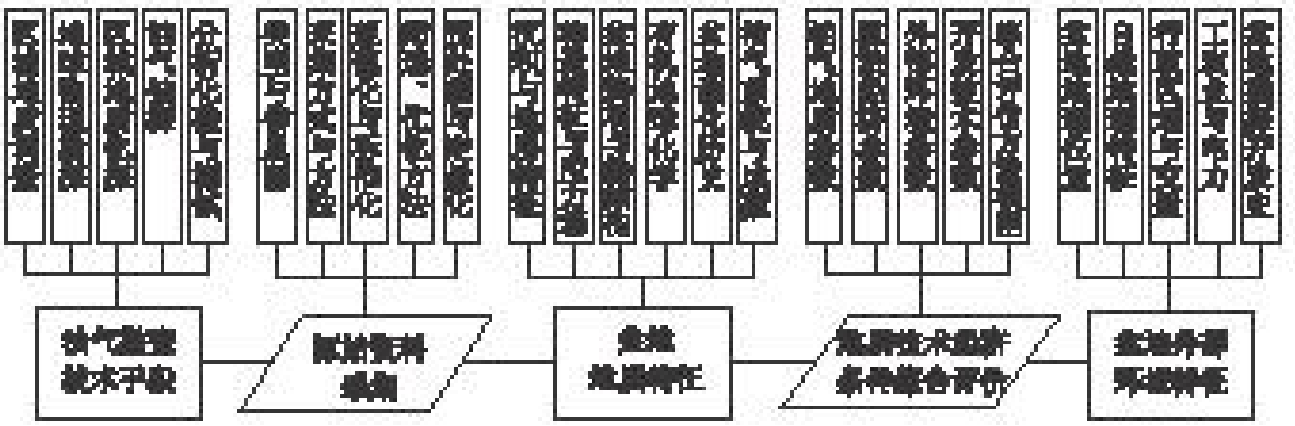


图 3-9　油气勘探总实体模型

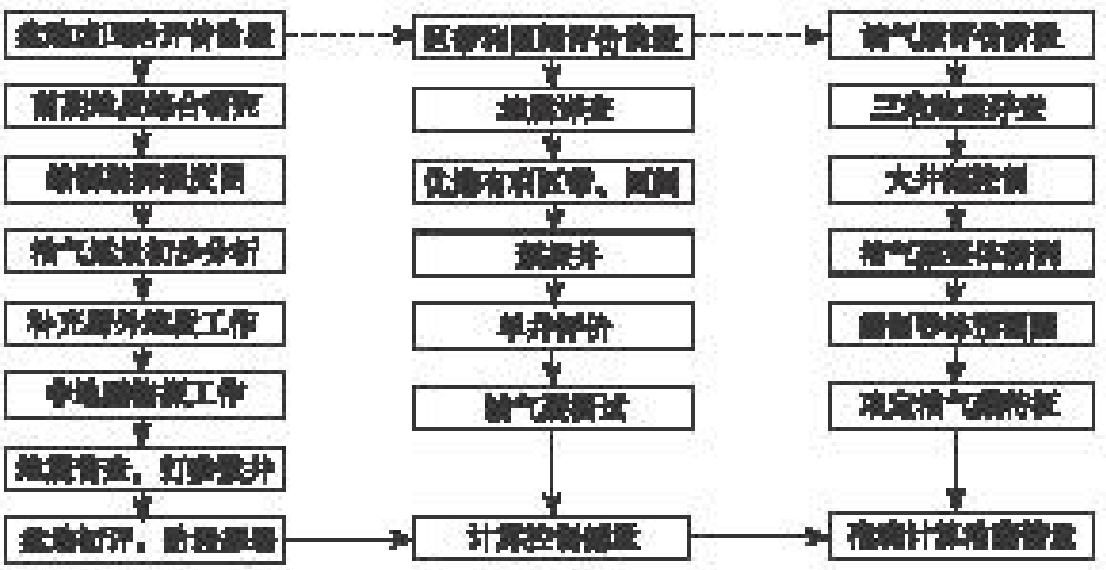


图 3-10　油气资源勘探技术手段子实体细分 (可以再分)

勘探工作流程　根据油气勘探的阶段划分, 可以用图 3-11 来简略描述各个阶段的工作流程。

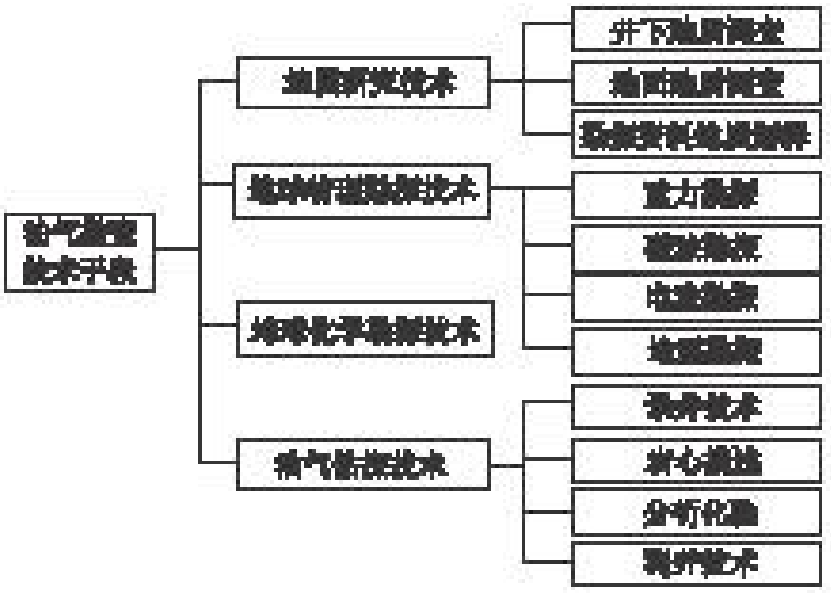


图 3-11　油气勘探工作流程略图

2．含油气盆地分析子模型

含油气盆地分析是以一个油气生成、聚集单元为对象，从盆地石油地质调查开始到优选出有利含油气区带的过程中使用的综合勘探方法。通过对盆地所处构造环境、盆地形成的力学机制、盆地充填的沉积学及沉积史、盆地地热演化史以及油气赋存、分布历史的研究，预测油气资源数量及其分布，优选出有利的含油气区带，为编制勘探计划提供科学依据。盆地分析主要应用于油气勘探的早中期（王益清等，1998），它要求综合各种资料，建立表征盆地特征的信息库，以及盆地演化和油气生排运聚的概念模型（丁贵明，1986）。

（1）含油气盆地分析子模型的整体结构

含油气盆地分析的基本参数包括 5 类：沉积参数、构造参数、地热参数、有机质参数和油气成藏参数，分别构成了 5 个子实体集。子实体集之间通过两两综合分析参数子集联系起来，其整体结构见图 3-12。

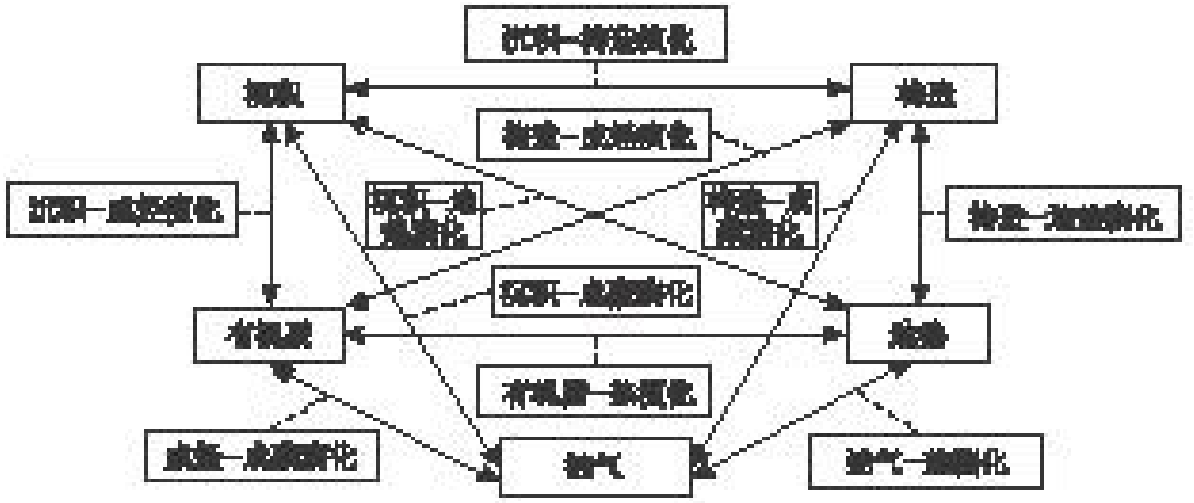


图 3-12　含油气盆地分析子模型的整体结构

（2）含油气盆地分析子模型的描述对象集

沉积参数属性集　主要包括：①基础参数：区域沉积特征；盆地沉积特征；②分析参数：地层格架；沉积相；沉积演化；体系域；层序地层；岩相古地理。

构造参数属性集　主要包括：①基础参数：区域构造特征；盆地构造特征；②分析参数：构造格架；构造应力场；构造演化。

有机质参数属性集　主要包括：①基础参数：有机质类型，有机质丰度；②分析参数：有机质分布。

油气成藏参数属性集　主要包括：①基础参数：油藏类型；油藏特征；②分析参数：

烃源岩生烃潜力；主生烃期与流体势；圈闭有效性与分布。

沉积-构造演化属性集　主要包括：①同沉积构造与同沉积构造运动；②沉积作用对同沉积构造运动的响应；③构造作用对沉积作用的控制。

沉积-地热演化属性集　主要包括：①不同沉积分区的热导率与地热场特征；②不同沉积演化阶段的古地热场特征；③压实-成岩过程中的热导率演化史。

沉积-成烃演化属性集　主要包括：①不同沉积分区的烃源岩特征；②不同烃源岩的有机质特征；③烃源岩的沉积史；④盆地充填史与有机质演化史、生烃史的相关关系。

沉积-成藏演化属性集　主要包括：①生储盖组合与空间分布；②盖层、储集层特征及其沉积史；③油气运移通道体系的形成、演化史；④岩性（地层）圈闭的类型及其时空分布特征；⑤岩性（地层）油气藏的形成、演化史。

有机质-热演化属性集　主要包括：①有机质特征与古地热场；②有机质成熟度；③有机质热演化史。

构造-地热演化属性集　主要包括：①不同构造分区的地热场特征；②不同构造演化阶段的古地热场特征；③构造-热事件演化史；④地幔热动态演化史。

构造-成烃演化属性集　主要包括：①不同构造分区的烃源岩特征；②不同沉降单元的有机质特征；③烃源岩的构造分割或破坏史；④各构造阶段有机质演化特征及相互关系。

构造-成藏演化属性集　主要包括：①不同构造分区的油气运移、聚集特征；②不同构造演化阶段的油气运移、聚集特征；③同沉积构造对油气成藏作用的控制或影响；④构造圈闭的形成、演化与油气运聚的时空匹配；⑤后期构造改造与油气再分配。

成烃-成藏演化属性集　主要包括：①油气的生成、排放、运移、聚集史；②有机质成熟度、油气相态与油气藏类型；③动力（幕式）排烃和运聚机制、效率及其相互关系；④二次生烃与油气成藏作用。

油气-热演化属性集　主要包括：①油气成藏期后的热演化史；②古油气藏的裂解、转化与破坏史。

(3) 含油气盆地分析子模型的总实体模型和工作流程

总实体模型　含油气盆地分析的总实体模型是含油气盆地分析子模型整体结构的细化。将含油气盆地分析的上述描述对象集与盆地分析整体结构组合起来，便可以得到充实了的含油气盆地分析总实体模型（图 3-13）。一般地说，应当根据具体盆地的实际情况制定实体模型，再从多个实体模型中归纳和抽提出概念模型和数据模型。由于图 3-13 所示的模型是从多个含油气盆地分析工作中总结出来的，也可以直接作为含油气盆地分析概念模型。

工作流程　含油气盆地分析涉及大量油气勘探资料的分析与综合。各种基础参数在分析过程中逐步引入，逐步形成各层次的综合参数和结论。含油气盆地分析的工作流程至今没有统一的模式，编者根据对多种模式的综合，大致可以用图 3-14 来表示。

3. 油气系统分析模型

油气系统是沉积盆地中与油气生、排、运、聚、散有关，且与周围环境密切联系着的若干相互作用、相互依赖的地质要素和过程（作用）及其描述参数和联系方式的集合体。油气系统分析是把油气从生成到成藏的过程中各种地质要素和条件，作为一个有机整体进

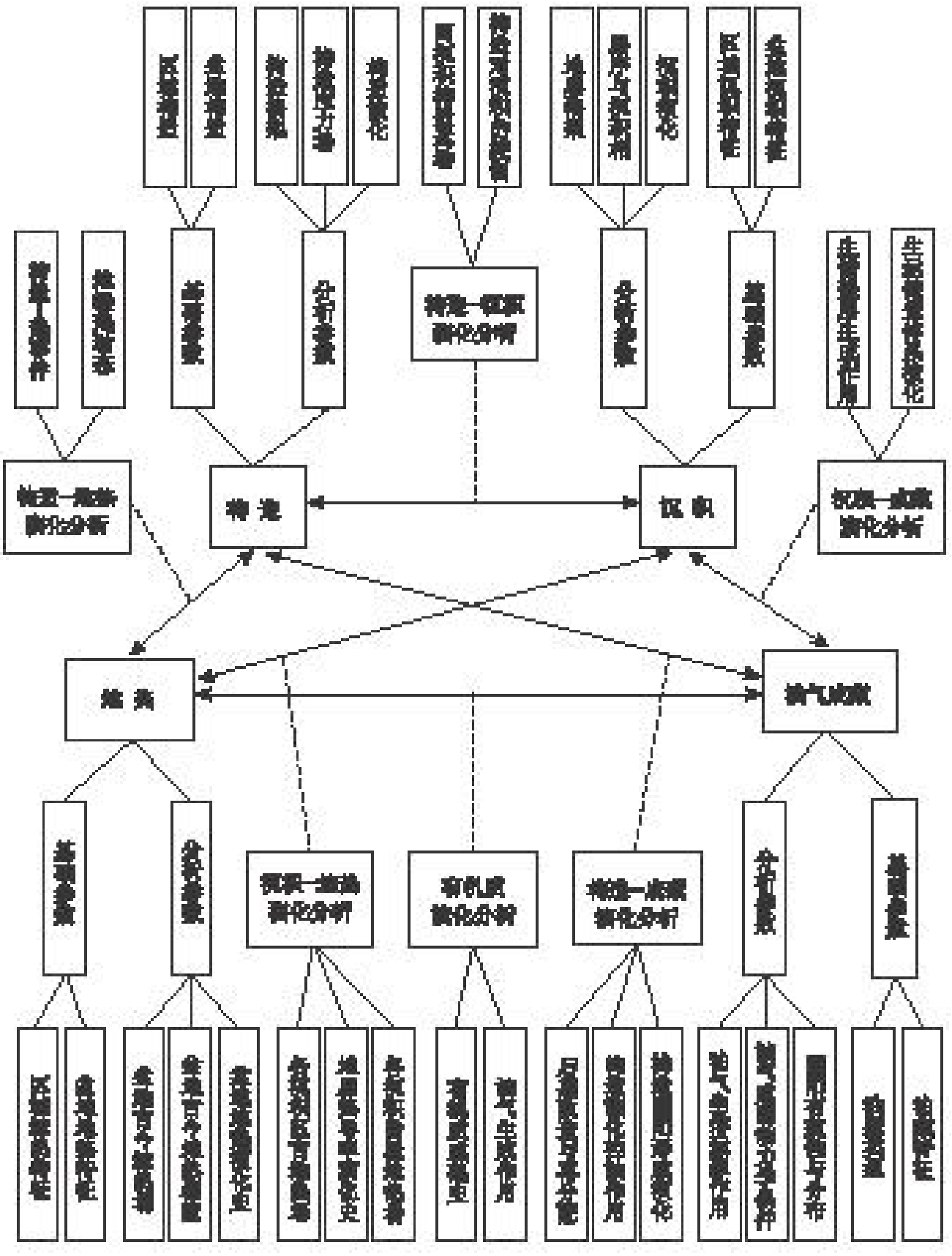


图 3-13　含油气盆地分析实体模型

行系统的综合研究，是含油气盆地评价的一种新的方法，研究内容包含以下几方面：①各种成因的天然气、凝析油、原油、重油及固态沥青；②油气源岩、储集层、盖层、上覆岩层；③圈闭的形成、烃类的形成、运移、聚集和保存；④上述诸因素在时间和空间上的成因关系。油气系统分析模型是对油气系统分析工作的概括与总结。

(1) 油气系统分析模型整体结构

油气系统分析的基本参数包括六大类，即有效烃源岩、储集层（体)、输导层（体)、盖层、驱动力和油气藏。这六类参数即为六大实体。油气成藏的基本过程也有六个，分别

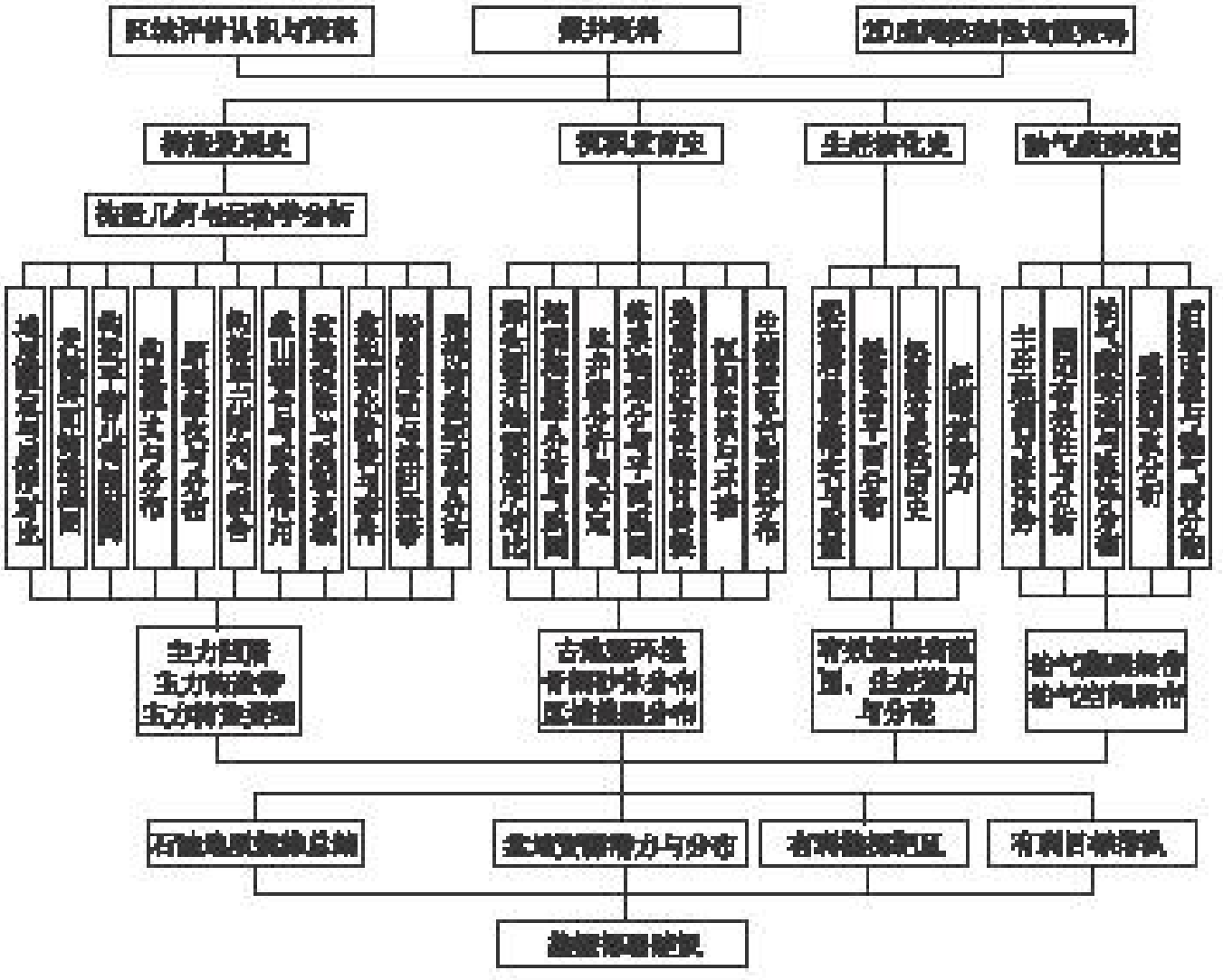


图 3-14　含油气盆地分析工作的一般流程

是油气生成过程、油气排放过程、油气运移过程、圈闭形成过程、油气聚集过程和油气藏破坏过程。这六大过程有先后也有交叉，将六大实体联系起来成为一体。这六大过程分别包含着一些子过程，因而也是实体集。油气系统分析模型的整体结构如图 3-15 所示。

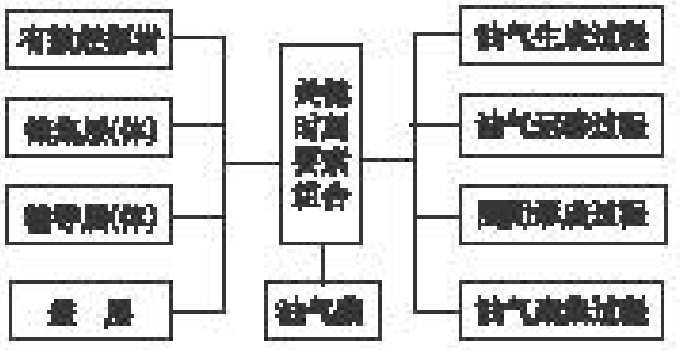


图 3-15　含油气系统分析整体结构

(2) 含油气系统分析描述对象集

油气藏属性集　主要包括：①油气藏类型；②原油和气的特征；③空间分布、规模及成藏规律。

有效烃源岩属性集　主要包括：①岩性特征；②有机地化特征；③生烃潜力；④空间分布及规模。

盖层属性集　主要包括：①盖层类型；②盖层性质特征；③几何参数和物性参数；④封堵性能评价。

输导层（体）属性集　主要包括：①输导层（体）类型；②输导层（体）性质特征；③裂隙和断层发育状况；④几何参数和物性参数；⑤输导性能评价。

储集层（体）属性集　主要包括：①储集层（体）类型；②储集层（体）性质特

征；③裂隙和断层发育状况；④几何参数和物性参数；⑤储集性能评价。

油气生成过程属性集　主要包括：①地热史；②关键时刻；③生烃范围与强度变化；④生烃史。

油气排放过程属性集　主要包括：①孔隙流体压力史；②构造应力场；③排烃方式与强度变化；④排烃史。

油气运移过程属性集　主要包括：①油气相态；②含烃流体动力学与运动学特征；③运移路径；④运移方式与效率；⑤运移史。

油气聚集过程属性集　主要包括：①圈闭类型；②圈闭规模与几何参数；③聚集期与圈闭形成期的匹配关系；④聚集方式与效率；⑤聚集史。

油气藏破坏过程属性集　主要包括：①构造破坏作用；②水动力破坏作用；③油气逸散作用；④油气散失史。

(3) 总实体模型和工作流程

总实体模型　将上述实体模型和描述对象集综合起来，就可以得到油气系统的总实体模型（图 3-16）。

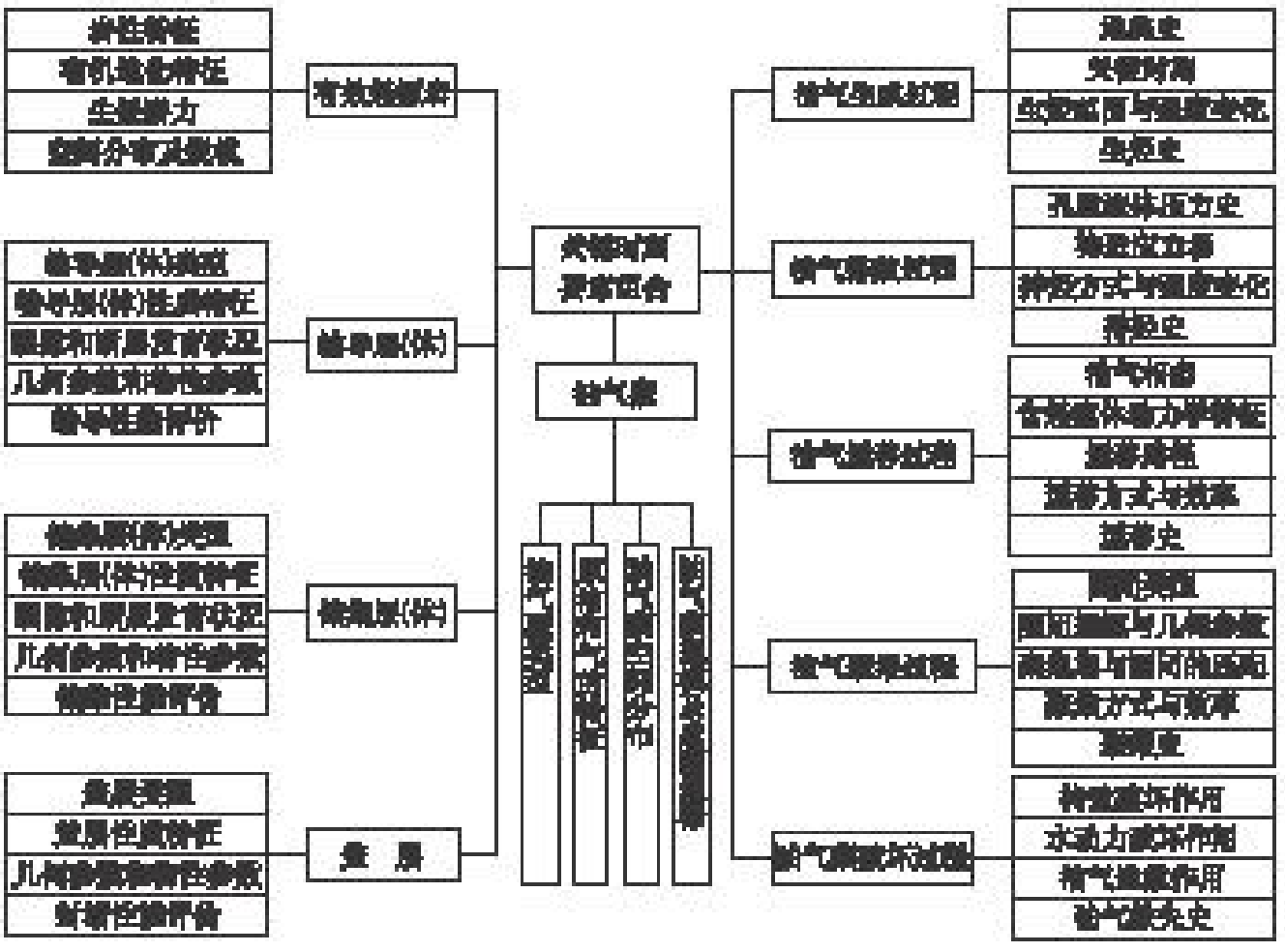


图 3-16　含油气分析总实体模型

工作流程　含油气系统分析过程一般在盆地分析的基础上进行，包括静态要素评价和动态过程评价，可以用流程图 3-17 来表示。

四、工程地质勘察模型

1. 工程地质勘察子模型

工程地质勘察的基本任务，是利用各种技术手段查清建筑地区的工程地质条件，指出

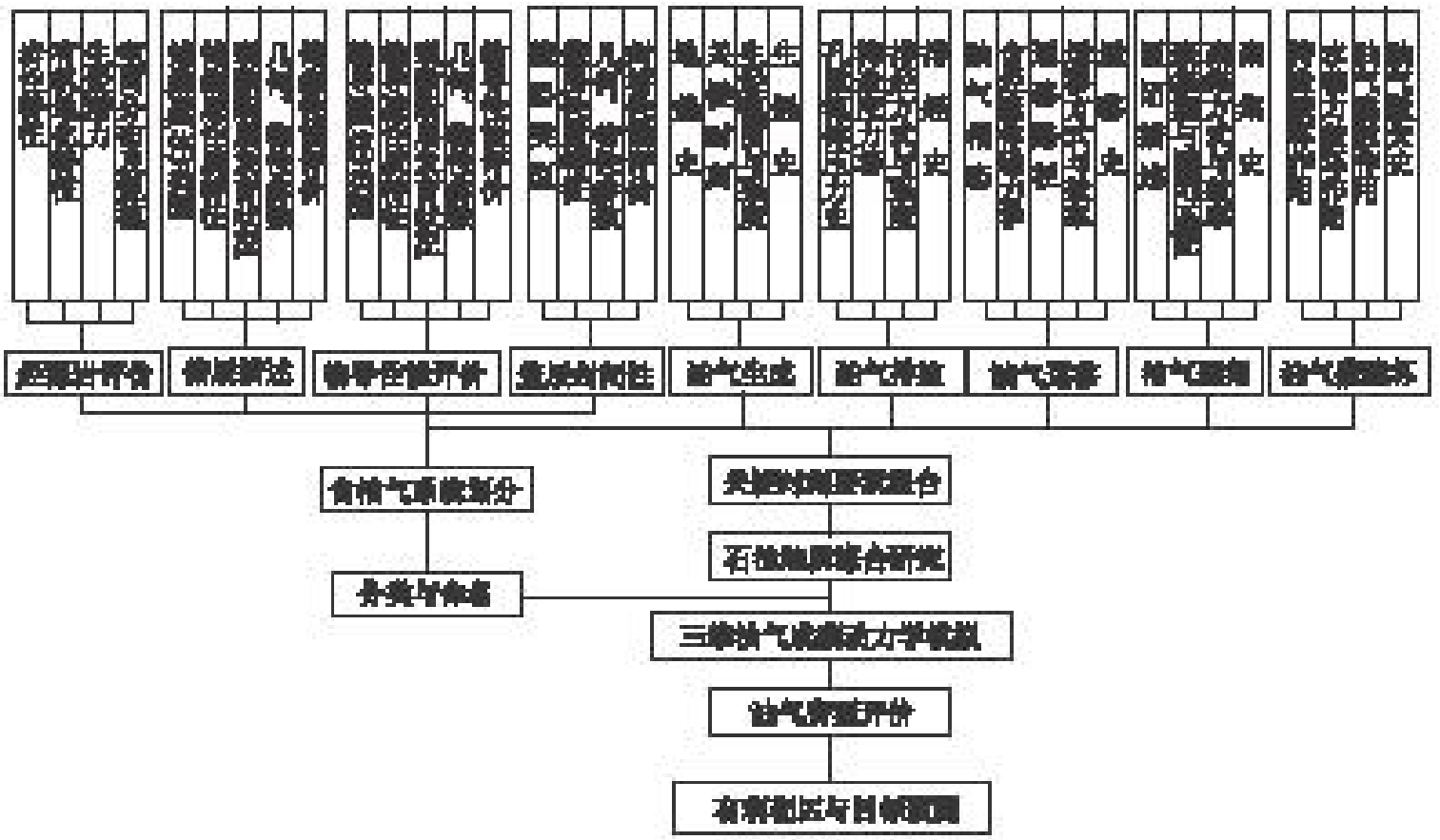


图 3-17　含油气系统分析工作流程

并解决所存在的工程地质问题，为建筑物的设计、施工、使用和维护提供所需的地质资料。其主要工作内容包括岩土工程性质研究、工程动力地质作用研究和区域工程地质研究，等等。工程地质勘察子模型应当力求全面而准确地概括这些基本工作任务和工作内容。

(1) 整体结构

工程地质勘察子模型包含原始资料采集、勘察区地理特征、勘察技术手段、勘察区地质特征、工程地质技术经济条件综合评价等 6 个描述对象子集，相互间通过两两综合分析参数子集相联系。其总体结构如图 3-18 所示。

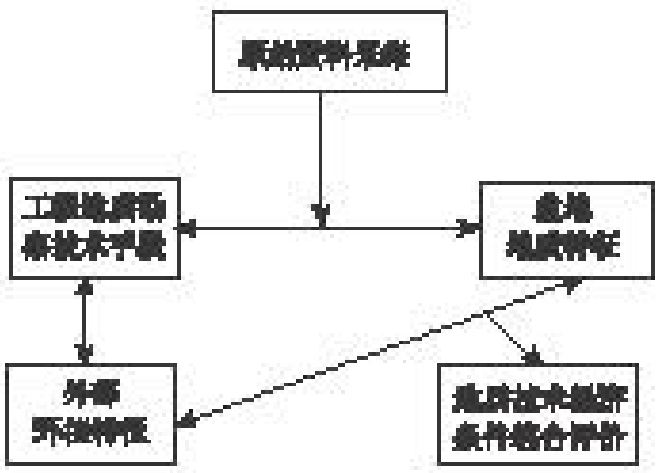


图 3-18　工程地质勘察子模型整体结构

(2) 描述对象集

原始资料采集属性集　主要包括：①前人成果资料；②野外地质资料；③野外物探资料；④分析测试成果。

勘察区地理特征属性集　主要包括：①地理位置及交通；②电力设施；③地形地貌条件；④气候特征。

勘察技术手段属性集　主要包括：①野外地质调查；②岩土力学测试；③工程地质测绘；④地面物探；⑤钻探、坑探、平硐等；⑥山地工程。

勘察区地质特征属性集　主要包括：①地形地貌条件；②物理地质现象；③地质结

构；④岩土类型及工程地质性质；⑤水文地质条件；⑥天然建筑材料。

工程地质技术经济条件综合评价属性集　主要包括：①工程地质参数（构造、基岩、盖层、岩溶、水文和地理参数，等等）；②政治经济和自然地理参数（国家需求状况、能源政策、环境政策，内外部建设条件，与经济区和中心城市的距离，等等）；③经济参数（蓄水量、引水量、发电量、水和电的价格、投资、成本以及利息率，服务年限，等等）；④施工技术参数（坝区和库区的构造复杂程度、断裂带性质与发育状况、溶洞的发育程度和充填程度、边坡与岩体的稳定性、软弱夹层的发育情况及其埋深和倾产状、河流的宽度和深度、河床坡角、水流速度和流量，等等）。

2．工程地质条件评价子模型

工程地质条件评价应当是一项对工作区的地质构造、基岩、覆盖层、岩溶、水文及地理条件进行全面分析与综合，从而对该项工程的地质条件做出合理评价的系统方法。工程地质条件评价子模型也是这项工作的归纳和概括。

（1）整体结构

以水利水电工程地质条件评价为例，工程地质条件评价子模型的基本结构参数有6大类，即构造参数、基岩参数、盖层参数、岩溶参数、水文参数和地理参数。对数据库管理系统而言，这6项基本参数同样是6个描述对象子集（实体子集）。每两个对象子集（实体子集）之间亦为多对多关系。参数集（实体集）之间分别通过综合分析参数子集来联系。由此得到21个描述对象集（图3-19）。

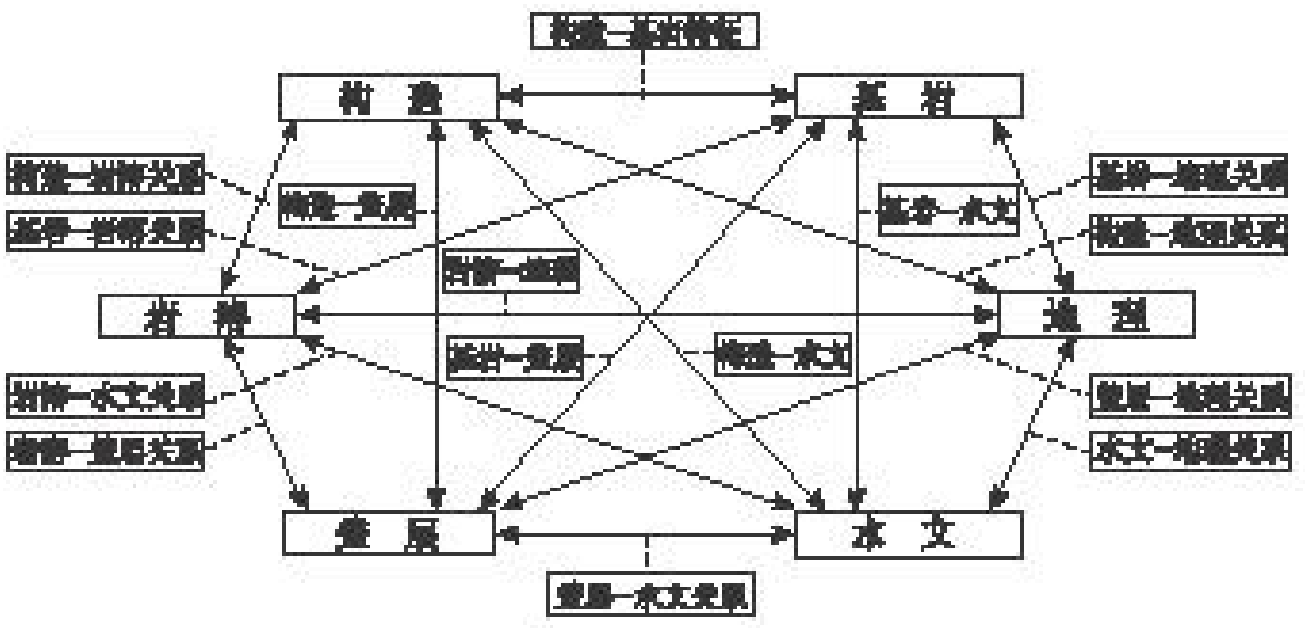


图3-19　工程地质条件评价子模型的整体结构

（据吴冲龙等，1996）

（2）描述对象集

构造参数属性集　主要包括：①基础参数：区域构造特征，地区构造特征；②分析参数：结构面组合特征，沉积结构面，构造结构面，岩体天然应力状态，构造演化。

基岩参数属性集　主要包括：①基础参数：区域基岩特征，地区基岩特征；②分析参数：基岩工程地质类型及特征，基岩结构构造及成分描述，基岩力学性质，基岩力学模型，构造应力场及构造应变场。

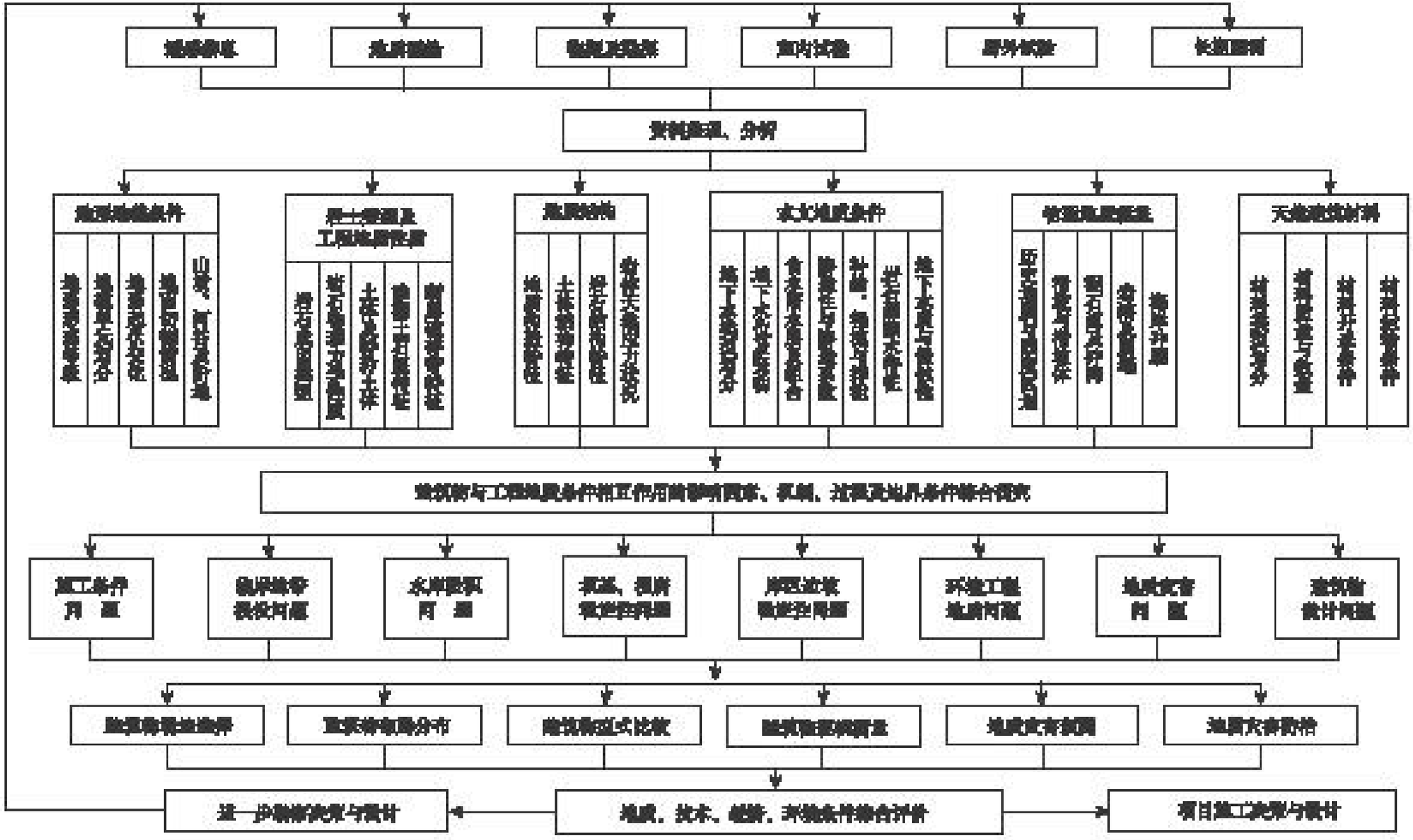


图 3-20　大型水利水电工程地质勘察评价工作的一般流程

（引自吴冲龙等, 1996）

盖层参数属性集　主要包括：①基础参数：区域盖层特征，地区盖层特征；②分析参数：盖层成因类型，盖层成分，结构构造，盖层厚度及空间分布，盖层物理力学性质。

岩溶参数属性集　主要包括：①基础参数：区域岩溶特征，地区岩溶特征；②分析参数：岩溶成因类型，岩溶地貌形态特征，岩溶空间分布及发育特征，岩溶水特征。

水文参数属性集　主要包括：①基础参数：区域水文特征，地区水文特征；②分析参数：含水层及隔水层特征，地表及地下水文地质特征，岩溶发育特征，水质分析及评价。

地理参数属性集　主要包括：①基础参数：区域地理特征，地区地理特征；②分析参数：地形地貌特征与成因类型，地貌单元空间分布及发育特征，地表水系特征，气候特征，交通，电力条件，工农业生产条件。

构造-基岩关系属性集　主要包括：①不同构造带的基岩特征；②不同基岩分布区的构造特征；③构造发育与基岩的关系；④构造对基岩的控制作用或影响。

构造-盖层关系属性集　主要包括：①不同构造带的盖层特征；②不同盖层分布区的构造特征；③构造发育与盖层的关系；④构造对盖层特征及盖层分布的影响。

构造-岩溶关系属性集　主要包括：①不同构造发育区的岩溶特征；②不同岩溶发育区的构造特征；③构造发育与岩溶的关系；④构造对岩溶的控制作用或影响。

构造-水文关系属性集　主要包括：①构造-水动力条件分析；②构造对水文地质条件的影响；③水文地质条件对新构造运动的影响。

构造-地理关系属性集　主要包括：①不同构造发育区的地形地貌特征；②不同地貌单元的构造特征；③构造对地形地貌的控制作用或影响；④地表水系与构造的联系。

基岩-岩溶关系属性集　主要包括：①不同类型基岩分布区的岩溶特征；②不同岩溶发育区的基岩特征；③基岩对岩溶的控制作用或影响。

基岩-水文关系属性集　主要包括：①基岩分布区的水文地质特征；②基岩水文地质条件评价；③基岩水质类型及评价；④水文地质单元的基岩特征；⑤基岩对水文地质条件的控制作用或影响。

基岩-盖层关系属性集　主要包括：①不同基岩发育区的盖层特征；②不同盖层分布区的基岩特征；③基岩对盖层的控制作用或影响。

基岩-地理关系属性集　主要包括：①不同基岩发育区的地形地貌特征；②不同地貌单元的基岩特征；③基岩对地形地貌的控制作用或影响；④地表水系与基岩分布的联系。

盖层-水文关系属性集　主要包括：①不同盖层分布区的水文地质特征；②盖层水文地质条件评价；③盖层水质类型及评价；④不同水文地质单元的盖层特征；⑤盖层对水文地质条件的控制作用。

盖层-岩溶关系属性集　主要包括：①不同盖层发育区的岩溶特征；②不同岩溶发育区的盖层特征；③盖层与岩溶的可能联系。

盖层-地理关系属性集　主要包括：①不同盖层发育区的地形地貌特征；②不同地貌单元的盖层特征；③盖层对地形地貌的控制作用；④地表水系与盖层分布的联系。

岩溶-水文关系属性集　主要包括：①不同类型岩溶分布区的水文地质特征；②岩溶分布区水文地质条件评价；③岩溶水质类型及评价；④不同水文地质单元的岩溶特征；⑤岩溶对水文地质条件的控制作用或影响；⑥水文地质条件对岩溶的控制作用或影响。

岩溶-地理关系属性集　主要包括：①不同岩溶发育区的地形地貌特征；②不同地貌

单元的岩溶特征；③岩溶对地形地貌的控制作用；④地表水系与岩溶分布的联系；⑤地形地貌对岩溶的影响。

水文-地理关系属性集　主要包括：①不同地貌单元的水文地质特征；②不同水文地质单元的地形地貌特征；③地形地貌对水文地质条件的控制作用或影响；④地表水系对地形地貌的控制作用或影响。

3. 工程地质勘察评价总实体模型及工作流程

综合上述两个实体模型，可以得到工程地质勘察评价的总实体模型。

在实际工作中，水文地质与工程地质勘察资料的分析与综合，通常也分为若干个层次。各种基础参数是在分析过程中按层次逐步引入的，而各层次的综合参数和结论在分析与综合过程中逐步形成。下一层次的综合参数和结论是上一层次分析与综合的基础，上一层次的分析与综合又可反过来指导下一层次的分析与综合。

从某种意义上说，水文地质与工程地质勘察、预测和评价过程，也是基础参数的数据采集、调用、分析、综合和反馈的反复过程。这一过程可以用一个综合性的流程图来表示。以大型水利水电工程地质条件勘察评价为例，其流程图表示如 3-20 所示。

思　考　题

1. 简述开展地矿资源勘查系统分析的目的和意义。地矿资源勘查单位的应用软件开发者是否可以不做系统分析而直接进行地矿资源勘查信息系统设计？为什么？

2. 为什么说系统需求分析与工作环境分析是系统调查阶段和系统分析的重要内容？

3. 业务现状与数据现状分析的基本内容是什么？就你所知的某一项地矿资源勘查（调查）工作，谈谈其业务现状与数据现状，并分析其数据流向。

4. 在地矿资源勘查信息系统中常用的术语代码体系有哪些种类？说说它们的特点和命名规则，并指出其适用范围和应用条件。

5. 术语代码的用途是什么？为什么在地矿资源勘查区点源信息系统和国家信息网络中，要求使用国家标准术语代码？

6. 谈谈实体模型研究在地矿资源勘查信息系统设计和开发过程中的意义。

7. 具体而简要地编制出你所接触过的某一项地矿或地质勘查（调查）工作的实体模型。

8. 试比较地矿与地质勘查实体模型与地矿与地质综合评价实体模型的异同点。

第四章　地矿资源属性数据库子系统的设计与应用

地矿资源属性数据库子系统是地矿资源勘查信息系统的核心组成部分，其总体设计工作是在系统分析的基础上进行的。该子系统总体设计的主要内容是根据系统分析所建立的系统模型来选择硬、软件配置，进行数据库的逻辑结构和数据字典设计。

第一节　数据库系统原理与方法概述

一、数据库系统的基本概念

在系统地了解数据库知识之前，应先了解和熟悉数据库的一些最基本的术语和概念，它们包括：数据、数据库、数据库管理系统和数据库系统。

1．数据库与数据库管理系统

数据库技术是一种计算机软件技术，发明于20世纪60年代初。该项技术根据一次输入多次使用的原理，使庞杂的数据能按一定格式组织、描述和存储，具有较小的冗余度、较高的数据独立性和易扩展性，并能为多个用户所方便地使用和共享，因而可以充分地发挥这些数据资源的价值（James，1977；Pratt & Adamski，1997）。

（1）数据库

数据库（DB：data base）是被长期存放在计算机内、有组织的、可以表现为多种形式的可共享的数据集合。这里“共享”是指数据库中的数据可为多个使用多种不同的语言、为了不同目的用户同时存取；“集合”是指某特定应用环境中的各种应用的数据及其数据之间的联系（联系也是一种数据）全部集中地按照一定的结构形式进行存储。

1）数据库的组成。一个数据库有4个主要成分：数据、关系（Relationship）、约束（Constraint）和模式（Schema）（图4-1）。数据是所存储的逻辑实体在计算机中的二进制表示；联系表示数据项之间的某种对应；约束是定义正确数据状态的断言；模式描述数据库中数据的组织和联系。

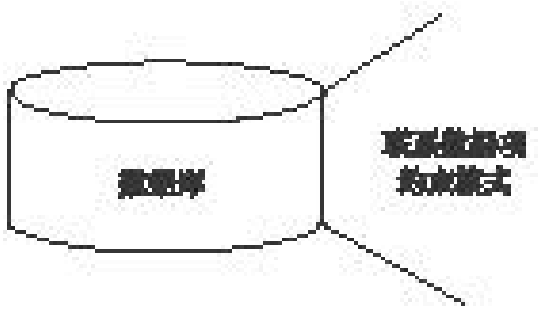


图4-1　数据库组成

2）数据库的优点。与数据文件管理方式相比较，数据库管理有许多优越性。这些优越性来自于数据库特有的数据结构和数据存贮方式。归纳起来，数据库有如下优点：

查询迅速、准确，且有多种表达与传输方式

以一个大型仓库管理为例：用手工操作，如要查找“某地区生产的商品的名称、规格、数量”，就可能要先翻阅账本，弄清分类目录和分类形式，然后按地区或商品分类、商品名称找到其他附带的属性，再抄写下来。如果要查找的内容较多，则查找与抄写既费时又

费力。数据库系统能根据给定的条件自动地按一定途径以毫秒级速度进行扫描查找，可以在瞬间将符合要求的数据一一用表格或其他方式显示出来，还可以自动地打印出来或通过网络传输到指定地址，而且不会出现错误。

数据结构化且统一管理　在数据库中，数据按逻辑结构组织起来，而按物理结构存放在磁介质中，并且由数据库管理系统（Database Management System）统一管理，既考虑了数据本身的特点，也考虑了数据之间以及文件之间的联系，数据的查询、检索和处理很方便。在传统的文件系统中，尽管记录内部存在某种结构，但记录之间没有联系，数据的查询、检索和处理十分繁琐、困难。实现数据的整体结构化管理，是数据库的主要特征之一，也是数据库系统与文件系统的本质区别。

数据冗余度小　在文件系统中，为了满足一个应用程序对数据的需要，常常在不同地方重复存放同一个或同一组数据。这样一来，如果一个多处存放的数据出现错误，就必须同时修改几个地方，否则将造成数据之间的不一致性。在数据库系统中，数据不仅可以面向某个局部应用而且可以面向整体应用，从而大大减少数据冗余，节约了存储空间，有效地避免了数据之间的不一致性。

具有较高的数据独立性　数据独立性是指用户应用程序与存储在数据库中数据的相互独立性。当人们利用应用程序调用数据库进行数据处理时，只涉及数据的逻辑结构，而不涉及其存储方式和物理结构。而当数据的物理存储方式和结构改变时，数据库管理系统将自动处理这种改变，而应用程序不必改变。近期甚至发展到数据库的逻辑结构改变了，用户程序也可以不变。用户程序不随数据逻辑结构改变而改变的特性，可称为数据的“逻辑独立性”。数据独立性（物理的和逻辑的）是数据库的重要特征和优点，它有利于在数据库结构修改时保持应用程序的稳定性，可以大大的减少应用程序员的软件开发工作量。

数据的共享性好　存放于数据库中的数据的共享性包括系统内部共享性和外部共享性两种，这是数据库管理方式区别于手工管理和文件管理方式的最本质特征和优点。

系统内部的共享性是指同一个（组）数据在一次处理中可以多次被调用的性能，而系统外部的共享性是指同一个（组）数据可以同时供多个用户调用。这两种共享性的原理是一致的，它使得多种作业、多种语言、多种用户可以相互覆盖地使用数据集合。

内部共享性有效地降低了数据的冗余度，系统很容易进行维护和扩充，而且能够使应用程序的编写更加方便。系统外部共享性能够促进并实现信息社会化服务，可以充分发挥信息的价值。举一个简单的事例，在没有使用数据库技术时，火车票和飞机票发售的数据不能共享，常常会出现有的售票处某线路车票或机票已售完，而另一售票处还有很多同一线路的车票或机票无人购买的情况。为了避免上述现象的出现，火车客运站和航空公司往往采取让不同售票处出售不同路线车票和机票的做法，结果给乘客带来许多不便。如今采用数据库来管理车票和机票发售工作，各个售票处可以互通有无，乘客不但可以在一个城市的任何售票处方便地查询并购买当地出发的某路线车票和机票，而且随着网络技术的发展，人们甚至可以在全国任一城市查询并购买异地出发的任一路线车票和机票。

(2) 数据库管理系统

数据库管理系统（DBMS: Data base management system），是数据库系统的核心软件，用于指挥计算机进行有效的组织和存储数据、管理和处理数据。它位于用户与计算机操作系统之间，主要功能包括①数据定义功能；②数据操纵功能；③数据库的运行管理功能；

④数据库的建立和维护功能；⑤数据字典功能；⑥数据通信功能等几个方面。

2．数据库语言

数据库语言是用于对数据库进行应用开发的基本编程工具。一般来讲，数据库管理系统的数据处理能力较弱，需要提供与数据库接口的高级语言及其编译系统，以便于开发应用程序。常用的数据库语言有两种类型：一种是交互式命令语言，它语法简明，可以独立使用，并且通常是专用的，称为自主型语言；另一种是嵌入式语言，如 C、FORTRAN、COBOL、Pascal 等，属于通用性语言，可以嵌入到某种程序设计语言中，称为宿主型语言。DBMS 对于自主型语言通常采用解释执行方式；对于宿主型语言则提供两种编译方法：

预编译方法　DBMS 提供预处理程序，对源程序进行扫描，识别数据库语言的语句并把它们转换成主语言调用语句，以便原来的编译程序能接受和执行它们。

修改并扩充主语言编译程序方法（也称增强编译方法）　未来的数据库语言应该是将数据库语言与通用的程序设计语言相结合，除能对传统的数据库功能提供透明访问外，也应能支持面向对象的程序设计方法。

3．数据库系统

数据库系统（DBS: database system）是一个由硬件、软件（操作系统、数据库管理系统和编译系统等）、数据库和用户构成的完整计算机应用系统。数据库是数据库系统的核心和管理对象。因此数据库系统的含义已经不仅仅是一个对数据进行管理的软件，也不仅仅是一个数据库，数据库系统是一个实际运行的，按照数据库方式存储、维护和向应用系统提供数据支持的系统。数据库系统的组成如图 4-2 所示。

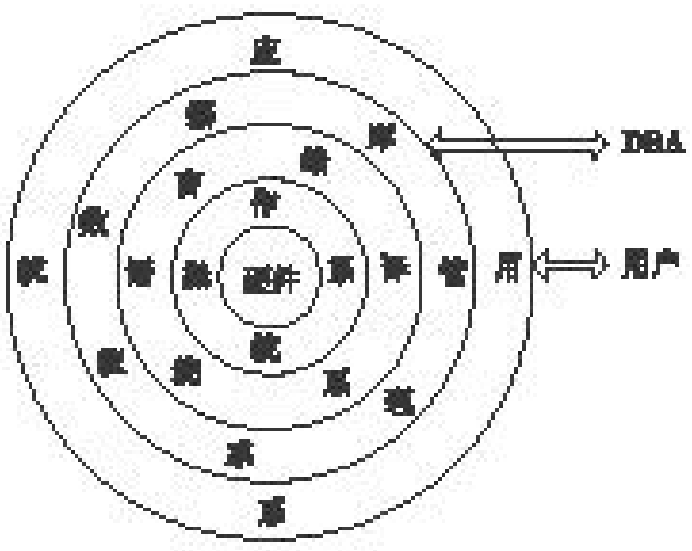


图 4-2　数据库系统的组成

（1）硬件

数据库系统的硬件包括计算机的主机、键盘、显示器和外围设备（例如打印机、光盘机、磁带机等）。由于一般数据库系统所存放和处理的数据量很大，加之 DBMS 丰富的功能软件，使得自身所占用的存储空间很大，因此整个数据库系统对硬件资源提出了较高的要求，这些要求是：①有足够大的内存以存放操作系统、DBMS 的核心模块、数据缓冲区和应用程序；②有足够大的直接存取设备存放数据（如磁盘），有足够的存储设备来进行数据备份；③要求计算机有较高的数据传输能力，以提高数据传送率。

（2）软件

数据库系统的软件除了数据库管理系统（DBMS）之外，还包括操作系统各种高级语言处理程序（编译或解释程序）、应用开发工具软件和特定应用软件等。应用开发工具包括应用程序生成器和第四代语言等高效率、多功能的软件工具，如报表生成系统、表格软件、图形编辑系统等。它们为数据库系统的应用开发人员和最终用户提供了有力的支持。特定应用软件是指为特定用户开发的数据库应用软件，如基于数据库的各种管理软件、管

理信息系统（MIS)、决策支持系统（DSS）和办公自动化（OA）等。

(3) 数据

数据是数据库的基本组成，是对客观世界所存在事物的一种表征，也是数据库用户的操作对象。数据是数据库系统，也是企业或组织的真正的财富。数据应按照需求进行采集并有结构地存入数据库。由于数据的类型多样性，数据的采集方式和存储方式也会不同。数据作为一种资源是数据库系统中最稳定的成分，即硬件可能更新，甚至软件也可以更新，但只要企业或组织的性质不改变，数据将是可以长期使用的财富。数据库中的数据具有如前所述的集合、共享、最少冗余和能为多种应用服务的特征。

(4) 用户

数据库用户是管理、开发、使用数据库的主体。根据工作任务的差异，数据库用户通常可以分成终端用户、应用程序员和数据库管理员等 3 种不同类型。

终端用户　这类用户使用数据库系统提供的终端命令语言，或者菜单驱动、表格驱动、图形显示和报表生成等对话方式，来存取和应用数据库中的数据。这类人员是一些并不精通计算机和程序设计的各级管理人员，但必须接受必要的数据库应用培训。

应用程序员　这类人员负责设计和编制应用程序。他们通常使用 C 语言、数据库语言或 4GL（第四代语言）等来设计和编写应用程序，供终端用户使用。应用程序员不仅要求具有较高的技术专长，而且还要具备较深的资历，熟悉部门全部数据的性质和用途，兼有系统程序员、系统分析员的能力。其具体职责是：①决定数据库的内容和结构；②决定数据库的存储结构和存取策略，使数据的存储空间利用率和存取效率均较优；③定义数据的安全性要求和完整性约束条件；④根据终端用户的需要，设计和编制各种功能强劲的应用程序。

对于应用程序员而言，有两个必备的重要工具，一个是系列的实用程序（DBMS 的装配、重组日志、恢复、统计等程序)；另一个是数据字典（存储数据库结构的定义、记录类型和字段定义等信息)。数据库系统建立时，DBA 还必须和系统分析员一起负责应用系统的需求分析和规范说明，确定系统的软硬件配置并参与数据库的设计。

数据库管理员　数据库管理员（DBA：Data Base Administrator）是指全面负责数据库系统的日常管理、维护和运行的人员。DBA 处于终端用户与应用程序员之间，是数据库系统能否正常运转的关键，大型数据库系统需要设置专门的管理办公室。其职责是监督控制数据库的使用和运行，实施数据库系统的维护、改进和重组，开展信息社会化服务。

对于不同规模的数据库系统，用户的人员配置是不相同的。只有大型数据库系统才配备有应用程序员和数据库管理员。应用型微机数据库系统比较简单，其用户通常兼有终端用户和数据库管理员的职能，但必要时也应当兼有应用程序员的能力。

二、数据结构的关系模型与关系数据库

数据模型是数据库中的数据类型、关系、集合、结构和表达形式的综合。数据库系统中所支持的主要数据模型是：层次模型（Herarchical model)、网状模型（Network model)、关系模型（Relational model）以及最近提出的面向对象模型（Object-oriented model)。与之对应的数据库系统为关系数据库系统、层次数据库系统、网状数据库系统和面向对象的数据库系统；而对应的数据库管理系统为关系数据库管理系统、层次数据库管理系统、网状数据库管理系统和面向对象的数据库管理系统。

1. 数据模型与数据结构

层次模型、网状模型是早期的数据模型，通常称为格式化数据模型。由于关系模型比网状模型、层次模型更为简单灵活，所以基于关系模型的关系数据库在市场上逐渐占据统治地位。致使大多数数据库系统，特别是基于微型机的数据库都是关系模型支持的。然而，现实世界中许多领域的数据，如图像、图形、声音等数据，用传统的关系模型表示较为困难，而面向对象模型却甚为得心应手。这导致近期面向对象模型的应用日渐增多。

(1) 网状模型 (network model)

用从结构 (网络结构) 表示实体类型及其实体间联系的数据模型称为网状模型。其数据结构是记录类型和系类型 (set type) 的集合。一个系类型是一棵二级树，而记录类型是结点。根结点是主记录类型，叶结点是成员记录类型。系类型 A - B 表示主记录类型 A 和成员记录类型 B 之间是 1∶ M 联系。即一个 A 记录可以与零个或多个 B 记录有联系。一个主记录和与之联系的所有成员记录构成了系类型的一个系值。系值表示记录之间的链接联系。

在网状模型的网络结构中，一个子结点可以有多个父结点，任意两个结点之间可以有一种或多种联系。网状模型的记录之间的 M∶ N 联系通过指针索引，因此其数据结构在物理上易于实现，效率也较高，但是编写应用程序较复杂，程序员必须熟练。网状模型有许多成功的 DBMS，20 世纪 70 年代的 DBMS 产品大部分是网状系统。CODASYL 组织提出的 DBTG 报告中的数据模型是网状模型的主要代表。

(2) 层次模型 (hierarchical model)

用树型 (层次) 结构表示实体类型及其实体间联系的模型叫层次模型。层次结构是一棵有向树，树的结点是记录类型。根结点以外的结点有且只有一个父结点。上一层记录类型和下一层记录类型的联系是 1∶ M 联系 (包括 1∶ 1) 联系。

在现实世界中，许多实体的联系本身就是一个自然的层次关系。例如家庭关系、人事组织关系和行政机构隶属关系，等等。在层次系统中可以用许多辅助手段来实现记录类型间的 M∶ N 联系，但由于层次顺序非常严格，引起数据的插入、删除较为复杂，应用程序的编写也较为麻烦。1969 年美国 IBM 公司的 IMS 系统产品便是最典型的层次模型系统，在 20 世纪 70 年代商业上曾得到广泛的应用。

(3) 关系模型 (relation model)

用二维表结构来表示实体及实体间联系的模型称为关系模型 (图 4-3)。关系模型是用二维表格结构来表示实体以及实体间联系的模型。它的数据结构是一个“二维表框架”组成的集合。由于关系模型采用了二维表格的数据结构来表述，每个关系在用户的眼中被看成是一个二维表格，因此每个二维表又可反过来称为关系。关系模型的实例由若干关系组成，或者说，由若干表格组成。

关系模型很容易被人们理解和应用。自 20 世纪 80 年代以来，计算机厂商推出的数据库管理系统的产品几乎都是基于关系模型的。

2. 关系模型及其相关概念

(1) 关系模型的表达

关系模型的数据结构较为明晰、单一，对它的表达涉及如下几个基本概念:

关系 (Relation)　关系是笛卡儿积的一个有意义的子集，即满足一定规范要求的二



图 4-3　关系模型及术语

维表。每个关系都有一个关系名。所有关系可理解为由关系的框架和若干元组构成。

元组（Tuple)　二维表格中的一行，相当于通常所说的记录值，简称为记录。

记录型　指表格的框架，一个表格数据相当于一个同质文件。

关系模式　行的定义（记录的型）称为关系模式。关系模式和关系也常常通称为关系。

属性（Attribute)　二维表格中的一列，相当于记录中的一个数据项，也称为字段。属性包括属性名、属性值和属性值类型三部分内容。

属性名　属性的名称（给属性起的名)。

分量　元组中的一个属性值。

属性值的类型　属性值分为数值型（N)、文字型（C)、逻辑型（L）和日期型（D)等 4 种。

属性的值域　属性的取值范围称为值域（Domain)。

(2）关键字及其类型

关键字（Key Word)　关键字是一个语义概念。能够用来唯一标识元组的属性（或属性组合）称为关键字。关系中的元组由关键字的值来唯一确定。在图 4-3 中，钻遇地层关系的关键字是“钻孔编号 + 层号”；在表 4-1 中，钻孔关系的关键字是钻孔编号。

在一个关系中，关键字的值不能为空，即关键字的值为空的元组在关系中是不允许存在的。但在一些桌面型的系统中，如 xBase 系列却没有这种严格的限制。有些关系的关键字是由单个属性组成的，还有一些关系的关键字由若干个属性组合而成。后者的元组不能由任何一个属性唯一标识，必须由多个属性的组合才能唯一标识。例如，钻遇地层关系：工程勘察中的钻遇地层包含钻孔编号、层号、底面深度、风化带类型、岩石地层单位、岩石基本名称、样品编号等 7 个属性。它的检索关键字应当由属性组合“钻孔编号 + 层号”构成。

表 4-1　钻孔基本信息

钻孔编号	钻孔位置	坐标系	纵坐标 (°)	横坐标 (°)	孔口标高 m	钻孔深度 m	覆盖层厚度 m	地面高程 m	风化带厚度 m	钻机编号
1348	长江右岸白岩尖山坡下	1	3410984.400	499541.940	107.570	74.77	0.00	107.570	0.00	
1642	白岩尖脚下	1	3411050.710	499607.690	83.610	49.74	0.00	83.610	1.95	
1643	白岩尖山顶	1	3410706.450	499286.840	215.230	33.79	0.00	215.230	33.70	
…	…	…	…	…	…	…	…	…	…	

候选关键字（Candidate Key）　如果在一个关系中，存在多个属性或属性组合都能用来唯一标识该关系的元组，这些属性或属性组合都称为该关系的候选关键字。

主关键字（Primary Key）　在一个关系的若干个候选关键字中，经指定具有唯一标识作用的属性或属性组合，称为该关系的主关键字。例如，在钻遇地层关系中，当风化带类型字段的值在每层中都是唯一的话，则“（钻孔编号，层号）”和“（钻孔编号，风化带类型）”是该关系的两个候选关键字，根据唯一标识作用的条件，可选择“（钻孔编号，层号）”作为该关系的主关键字。

非主属性（Non-Primary Attribute）　关系中不组成关键字的属性均为非主属性。

外部关键字（Foreign Key）　当关系中的某个属性或属性组合虽不是该关系的关键字或只是关键字的一部分，但却是另一个关系的关键字时，称该属性或属性组合为这个关系的外部关键字。例如，对表 4-1 中的钻孔关系，钻机编号不是关键字，但钻机编号是钻机关系的关键字，所以钻机编号是钻孔关系的外部关键字。

主表与从表　主表和从表是指以外部关键字相关联的两个表。以外部关键字作为主关键字的表称为主表；外部关键字所在的表称为从表。在表 4-1 中，以“钻孔编号”作为主关键字的“钻孔关系表”是主表，而以“钻机编号”作为外部关键字的“钻机关系表”是从表。关系模式是稳定的，而关系可随时间和数据库更新而不断变化。

3. 关系数据库

采用关系型数据结构模型来组织数据的数据库称为关系数据库。数据结构的关系模型是建立在严格的关系代数基础之上的，其组织数据采用数学方法来实现。

关系数据库的概念简单清晰，语言高度非过程化，用户不需了解复杂的存取路径，不需说明“怎么干”，而只需指出“干什么”就能方便地操作数据库。经过近 30 年的发展，目前已经涌现出许多性能优良的商品化管理系统（RDBMS），例如著名 DBZ、ORACLE、INGRES、SYBASE、INFORMIX 等。关系数据库产品也已经从单一的集中式系统，发展成为可在网络环境中运行的分布式系统；从封闭式系统逐步发展成为开放式系统，从联机事务处理发展到能够支持信息管理和辅助决策。系统的功能不断完善，应用领域也迅速扩大。

(1) 关系操作

关系操作包括传统的集合操作和专门的关系操作两类，分别通过关系集合运算和关系代数运算来实现。下面将分别对这两类关系操作做一些说明。

1）传统的集合操作。传统的集合操作包括并、交、差、广义笛卡儿积等。这类操作将关系看成元组的集合。其操作是从关系的水平方向，即是对关系的行来进行的。

设关系 R 和关系 S 具有相同数目的属性列（n 列属性），并且相应的属性取自同一个域，则可定义以下四种集合运算:

并（Union）　关系 R 与关系 S 的并，它是属于 R 或属于 S 的元组组成的集合，结果为 n 列属性的关系。

交（Intersection）　关系 R 与关系 S 的交，它是既属于 R 又属于 S 的元组组成的集合，结果为 n 列属性的关系。

差（Difference）　关系 R 与关系 S 的差，它是属于 R 而不属于 S 的元组组成的集合，结果为 n 列属性的关系。

广义笛卡儿积（Extended Cartesian product）　关系 R（假设为 n 列）和关系 S（假设为 m 列）的广义笛卡儿积是一个（n + m）列元组的集合，每一个元组的前 n 列是来自关系 R 的一个元组，后 m 列是来自关系 S 的一个元组。若 R 有 k1 个元组，S 有 k2 个元组，则关系 R 和关系 S 的广义笛卡儿积有 k1× k2 个元组。

图 4-4 中（a）、（b）分别为具有三个属性列的关系 R 和 S，（C）为关系 R 与 S 的并，（d）为关系 R 与 S 的交，（e）为关系 R 和 S 的差，（f）为关系 R 和 S 的广义笛卡儿积。

R

A	B	C
a1	b1	c1
a1	b2	c2
a2	b2	c1

（a）

S

A	B	C
a1	b2	c2
a1	b3	c2
a2	b2	c1

（b）

R 与 S 的并

A	B	C
a1	b1	c1
a1	b2	c2
a2	b2	c1
a1	b3	c2

（c）

R 与 S 的交

A	B	C
a1	b2	c2
a2	b2	c1

（d）

R 与 S 的差

A	B	C
a1	b3	c2

（e）

R 和 S 的广义笛卡儿积

A	B	C	A	B	C
a1	b1	c1	a1	b2	c2
a1	b1	c1	a1	b3	c2
a1	b1	c1	a2	b2	c1
a1	b2	c2	a1	b2	c2
a1	b2	c2	a1	b3	c2
a1	b2	c2	a2	b2	c1
a2	b2	c1	a1	b2	c2
a2	b2	c1	a1	b3	c2
a2	b2	c1	a2	b2	c1

（f）

R 和 S 的等值连接

A	B	C	A	B	C
a1	b2	c2	a1	b2	c2
a1	b2	c2	a2	b2	c1
a2	b2	c1	a1	b2	c2
a2	b2	c1	a2	b2	c1

（g）

R 和 S 的自然连接

A	B	C	A	C
a1	b2	c2	a1	c2
a1	b2	c2	a2	c1
a2	b2	c1	a1	c2
a2	b2	c1	a2	c1

（h）

图 4-4　集合运算示例

2）专门的关系操作。包括选择、投影、连接等。这类操作不仅涉及行，而且涉及列。

选择（Selection）　选择操作是指在关系中选择满足某些条件的元组。例如，要找出年龄为 21 岁的所有学生的数据，可以用 21 岁为条件，对学生基本信息表做选择操作。

投影（Projection）　投影操作是在关系中选择某些属性列。例如，要找出所有教师

的姓名、电话，则可以对教师基本信息表做投影操作，将表数据投影到教师名和电话列。

连接（Join） 连接操作是从两个关系的笛卡儿积中，选取属性间满足一定条件的元组。例如，对图4-4中关系R和S做连接操作，连接条件是S·B-R·B，即在图4-4（f）所示的两个关系的笛卡儿积中选取S·B-R·B的元组，得到的结果关系如图4-4（g）。连接条件中的属性称为连接属性，两个关系中的连接属性应属于同一数据类型，即都是数字型或都是字符型。连接条件中的算符为算术比较运算符，当此算符取“-”时，为等值连接。图4-4（g）便是关系R和S做等值连接后得到的结果。若等值连接属性为相同属性（或属性组），且在结果关系中去掉重复组，则此等值连接为自然连接。图4-4（h）即是关系R和S做自然连接得到的结果关系，其连接条件为S·B-R·B。自然连接是最常用的连接操作。

由此可见，关系操作的特点是集合操作，即操作的对象是元组的集合——关系，而操作的结果也是元组的集合——关系。

（2）关系模型的完整性

关系模型中的完整性是指数据库数据的正确性和一致性。数据完整性由数据完整性规则来维护。完整性规则包括系统定义的结构完整性（实体完整性、参照完整性）规则和用户定义的一般完整性规则。完整性规则的集合构成完整性约束条件，用来限定数据库状态及其变化，是在给定的数据结构模型中，数据及其联系的制约和依存基础。

1）实体完整性规则。实体完整性（Entity Integrity Rule）是指关系的主属性，即主码的组成不能为空，也就是关系的主属性不能是空值（null）。在关系模型中，一个关系通常对应一个表。在机器上实际存储数据的表称为基本表，查询结果表是临时表，而视图表是一种虚表，并不实际存储数据。实体完整性是针对基本表的。因此，换句话说，实体完整性是指在实际存储数据的基本表中，主属性不能取空值。

2）参照完整性规则。如果在基本关系R1的属性中，包含与另一基本关系R2的主属性K相对应的属性组F，则称F为基本关系R1的外主属性。参照完整性（Reference Integrity Rule）就是指在含有外主属性F的关系R1中，外主属性值（F值）只能在空值和另一个关系R2的对应主属性值（K值）之间进行选择的特性。

实体完整性与参照完整性是由系统自动支持的，即在建立关系（表）时只要说明了“谁是主属性”、“谁参照于谁”，系统将自动进行此类完整性的检查。

3）用户定义的完整性规则。用户定义的一般完整性规则，是针对某一应用环境的完整性约束条件，它反映了某一具体应用所涉及的数据应满足的要求。在实际系统中，这类完整性规则一般在建立库表的同时进行定义，应用编程人员不需再作考虑。如果某些约束条件没有建立在库表一级，则应用编程人员应在各模块的具体编程中通过程序进行检查和控制。

（3）关系数据库模式

1）模式。模式为数据库管理系统各部件所定义的各种数据库视图。一种模式将数据存储的物理外表与逻辑表示分开（图4-5）。内部模式定义数据在物理数据存储区中如何组织以及放在何处；概念模式按照适当的数据库数据模型（如关系模型或对象模型）定义所存储数据的结构；外部模式为特定用户（们）所定义的数据库的一个或多个视图。一个数据模型是数据库的一种抽象定义，它提供一种描述数据的表示法。

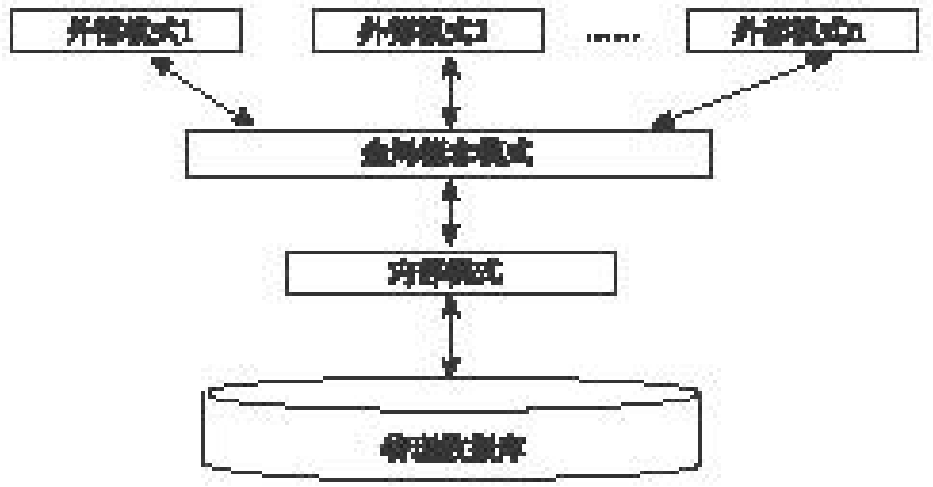


图 4-5　数据库模式的概念

2）关系模式。关系模式就是关系的框架，也称表框架，相当于记录格式。它是对关系结构的描述。一个关系模式实际上确定了这个关系的二维表的形式（二维表的框架）。

当一个关系模型被确定之后，就可以用 DBMS 提供的关系数据定义语言来描述和定义关系模型和组成模型的每一个关系，其中每一个被定义的关系模式包括关系名、属性名、属性的数据类型、属性宽度、属性值的约束等，最后指出每个关系的主键。用关系 DBMS 提供的关系数据定义语言来描述关系二维表的框架（相当于记录格式），这有点像我们做表填写数据时，首先要画出表框（每栏的栏名对应属性名，宽度对应属性的长度等）的情形。为了便于在书面上讨论和描述关系模式，将关系模式用下面最简单的式子表示出来:

R（A1，A2，…，An）

其中，R 为关系名，A1，A2，…，An 为属性名。

关系数据库模式是关系数据库结构的语言描述（即程序化），是一组关系模式的集合。

3）视图。视图（view）是用户所用到的那部分数据的描述，也称为关系子模式。视图是一个虚表，所对应的数据存储在基本表中。一旦需要，可以从一个或几个基本关系（表）中导出来。视图一经定义就可以和基本表一样进行查询操作，但更新（增、删、改）操作将有一定限制。

三、数据库的结构与工作原理

数据库系统对数据（也包括传统意义上的资料和信息）的强大管理与处理功能，源自于数据结构模型的合理选择、数据库内部层次结构的妥善配置、数据模式之间映像关系的巧妙构思和数据库管理系统的精彩设计。有关数据结构模型已经在前面作了较多介绍，下面将围绕另外 3 个方面来简单地介绍数据库的工作原理。

1. 数据库结构与映像关系

（1）数据模型的层次划分

从结构层次上看，数据模型通常有全局模型、外模型和内模型之分。全局数据模型是数据库的全局逻辑数据集合，通常也叫做概念模型。它是数据库管理员所理解的实体、实体属性和实体间的关系。外模型是一种局部的逻辑数据集合，也是用户所理解的实体、实体属性和实体间关系。而内模型是物理数据的集合，也是数据库设计者所理解的实体、实

体属性和实体间的关系。采用 3 个层次的数据模型，是为了适应计算机管理的需要。

数据库管理系统在将实际数据存入计算机的过程中，要对数据进行两次变换。其一，为了减少冗余、实现共享，把所有用户的数据综合成一个统一的全局数据模型；其二，为了提高存取效率，改善性能，把全局模型的数据按照最优的物理组织形式来存放，即转化为内模型。当数据库管理系统向用户提供数据时，则作相反的变换，成为外模型。外模型可以有多个，而概念模型和内模型均只有一个。在进行数据库应用设计时，用户只需全面而准确地提供自己所使用的外模型，而不必干预数据物理组织的形式；反之，数据库设计者应当关注概念模型和数据物理组织的最优形式，而不必干预用户外模型的构筑。

(2) 数据模式的层次划分

用数据描述语言精确定义的数据模型称为数据模式。与 3 个级别的数据模型相对应，数据库设计中采用了 3 个级别的数据模式，即外模式、内模式和概念模式 (图 4-6)。外模式又称为子模式或用户模式，是数据的局部逻辑结构视图，也是各种数据库用户从不同角度所看到的各种数据视图。内模式又称存储模式，是数据在数据库中的内部表达，即数据的物理结构和存储方式的描述。概念模式又称全局模式或逻辑数据模式，是数据库中全体数据的全局逻辑结构和特征的综合描述，也是所有用户的全局公共数据视图。

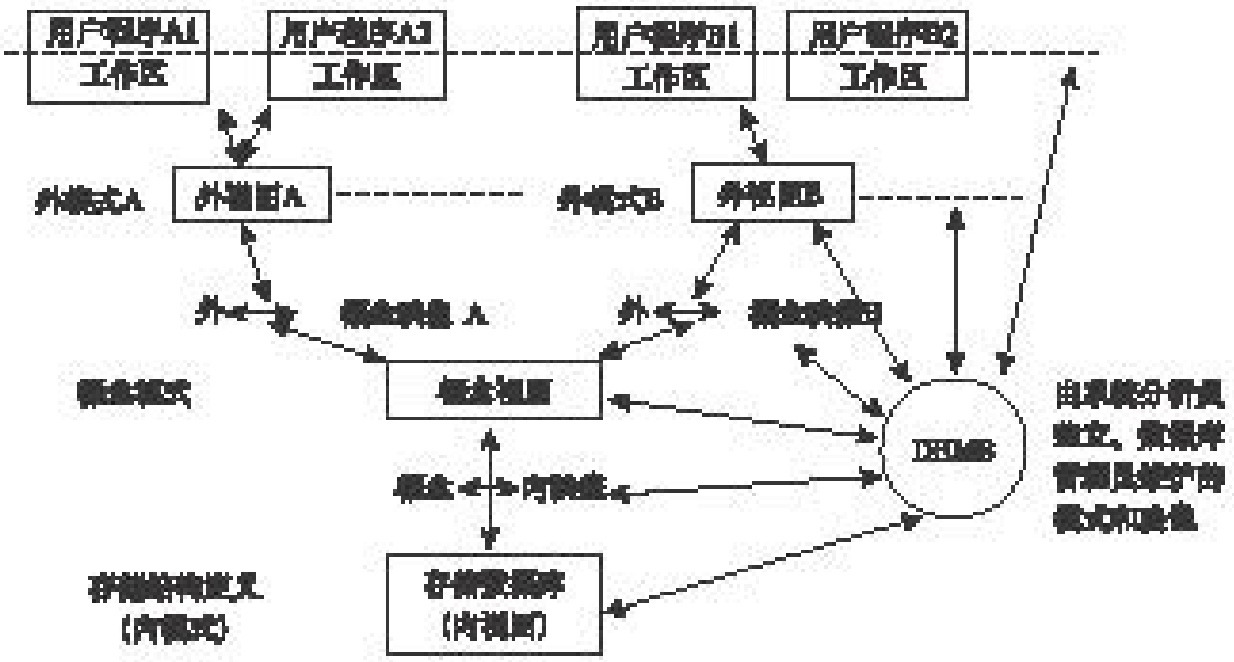


图 4-6　数据库系统结构示意

无论那一级模式都只是数据库的一种框架。在数据库系统内，框架和数据是放在不同的地方的。以外模式为框架的数据库是用户数据库，以内模式为框架的数据库是物理数据库，而以全局模式为框架的数据库是概念数据库。用户数据库是概念数据库的部分抽取，物理数据库是概念数据库的具体实现，而概念数据库是物理数据库的抽象表示。其中，只有物理数据库才是真实存在的——存放在外存储设备上的实际数据。其他两个数据库在外存储设备上是不存在的，但数据库管理系统 (DBMS) 能够根据需要，从相应的物理数据库中快速而准确地构造出概念数据库和用户数据库。因此，用户可以放心地使用它的数据库。

在设计数据库时，软件人员主要是考虑整个数据库的框架；而在使用数据库时，用户主要关心数据库的内容。在数据库生命周期中，数据库的框架是相对稳定的，而数据库的

内容是经常变化的。概念模型或概念数据库是按一定的规则建立起来的，不能等同于用户模型或用户数据库。在目前的数据库技术条件下，当出现强行使概念模型或概念数据库等同于用户模型时或用户数据库时，数据库的设计便难以顺利进行，造成适得其反的结果。

(3) 数据模式之间的映像关系

映像关系即对应关系。在数据库内部的数据子模式与数据全局模式之间、数据全局模式与数据物理模式之间，存在着两两对应的映像关系。这种映像关系既保证了用户数据与物理数据之间的一致性，又保证了数据库数据与应用程序之间的独立性。

子模式←→模式映像　它定义了各子模式（外模式）与全局模式之间的映射关系。这种映像关系使得用户数据能够独立于全局逻辑数据，保证了系统的逻辑数据独立性。当整个系统数据管理的范围扩大或某些管理的要求发生改变，而要求改变全局逻辑结构及其概念模式时，可以只改变映射关系而保持原有的子模式不变，同时扩充新的子模式。

模式←→物理模式映像　它定义了全局模式与物理模式（内模式）的映射关系。这种映像关系使得全局的逻辑数据能够独立于物理数据，保证了系统的物理数据独立性。于是，当数据库的硬件设备或存储方法改变，或者为了提高文件的存取效率，可以只改变物理模式（内模式）和模式←→物理模式映像，而保持模式和子模式（外模式）不变。

上述两种映像关系把用户数据和物理模式完全分开，使用户得以逻辑地处理数据，而不必考虑数据在计算机中的表示和存储方法，摆脱了繁琐的物理存贮细节。同时，由于用户程序不依赖于物理数据，可以大大减地少应用程序的维护工作量和经费开支。

2. 数据库管理系统的功能及操作过程

在数据库中，数据操作主要有检索和更新（包括插入、删除、修改）。数据结构模型要定义这些操作的确切含义、操作语言、操作符号和操作规则。

(1) 数据库管理系统的功能

数据库管理系统（DBMS）执行对数据的操作，它通常具备以下几种功能。

数据库定义功能　DBMS 提供数据定义语言（DDL: data definition language），用户通过它可以方便地对数据库中的相关内容（如数据库结构、基本表、视图和索引）进行定义。用 DDL 编写的概念模式、外模式和内模式分别称为源概念模式、源外模式和源内模式。各种源模式通过相应的模式编译程序转换为机器内部代码表示形式，分别称为目标概念模式、目标外模式和目标内模式。这些目标模式是对数据库结构的描述，并被保存在数据字典中。

数据操纵功能　DBMS 向用户提供数据操纵语言（DML: data manipulation language），利用 DML 来书写命令，可以实现对数据库的各种基本操作，例如，对数据库中数据的查询、插入、删除和修改；也可以补充实现对数据库的各种特殊操作，例如，对数据进行各种统计、分析、处理和可视化输入输出的操作。

在现行的数据库产品中，DBMS 一般同时拥有自主型（如 XBASE）与宿主型（如 COBOL，FORTRAN，PASCAL，VB 或 C 等高级语言）两种类型。

数据库运行和控制功能　这是 DBMS 功能的核心部分，包括通过数据安全性、完整性、并发性和数据恢复等的控制，来对数据库进行有效的管理，以便确保数据库数据正确有效和数据库系统的正常运行。

数据的安全性保护　是指为防止不合法的使用造成数据泄密和破坏，使每个用户只能按

规定对某些数据以某些方式进行使用和处理。例如，划分不同的用户权限，有的用户只能查询而不能修改数据，而有的用户却有修改权等。数据的完整性约束是对数据的正确性、有效性和相容性要求，即将数据有效性控制在一定范围内，或要求数据间满足一定的关系。

并发控制 是对多个用户同时存取、修改数据库的协调和控制。当多个用户同时存取、修改数据库时，可能造成相互干扰而得到错误结果，甚至使数据库的完整性遭到破坏。DBMS 提供了对多用户的并发操作加以控制和协调的功能。

数据库恢复 计算机的硬件故障、操作员的失误以及人为的破坏，都会影响数据库中数据的正确性，甚至造成数据库部分或全部数据的丢失。DBMS 具有将数据库从错误状态恢复到某一已知的正确状态（亦称为完整状态或一致状态）的功能。

数据库的建立和维护功能 包括数据库初始数据的安装，数据库的转储、恢复、重组织，数据库的性能监视、分析功能，等等。这些功能是数据安全性、完整性的重要保证，通常由一些应用程序来完成的，但仍属于数据库管理系统的重要组成部分。

数据字典功能 数据字典（DD: data dictionary）也称为数据目录或系统目录，它以表格的形式存放着数据库的各种有关信息（数据库三级模式、数据类型、用户名表、用户权限、程序与其用户联系等有关数据库系统的信息），能帮助用户、数据库管理员和数据库管理系统本身使用和管理数据库。对其作适当的开发，可用于提高自学水平并解决数据库软件复用与自动编程方面的难题。在大型系统中，DD 也可单独成为一个系统。

数据通信功能 用于实现用户程序与 DBMS 之间的通信的功能，包括与 OS 的联机处理、分时处理和远程作业传输的相应接口等。这些功能对分布式数据库系统尤为重要，通常与操作系统协调完成。在当前的各种 DBMS 中，数据通信功能都被列为重点开发内容。

(2) 数据操作过程

数据库管理系统回应用户访问的操作过程，可以用应用程序从数据库中读取一个数据记录（图 4-7）为例来加以说明。该事例还具体地反映了数据库各部分所起的作用以及它们之间的相互关系。归纳起来数据库的操作过程，大致有如下 10 个步骤:

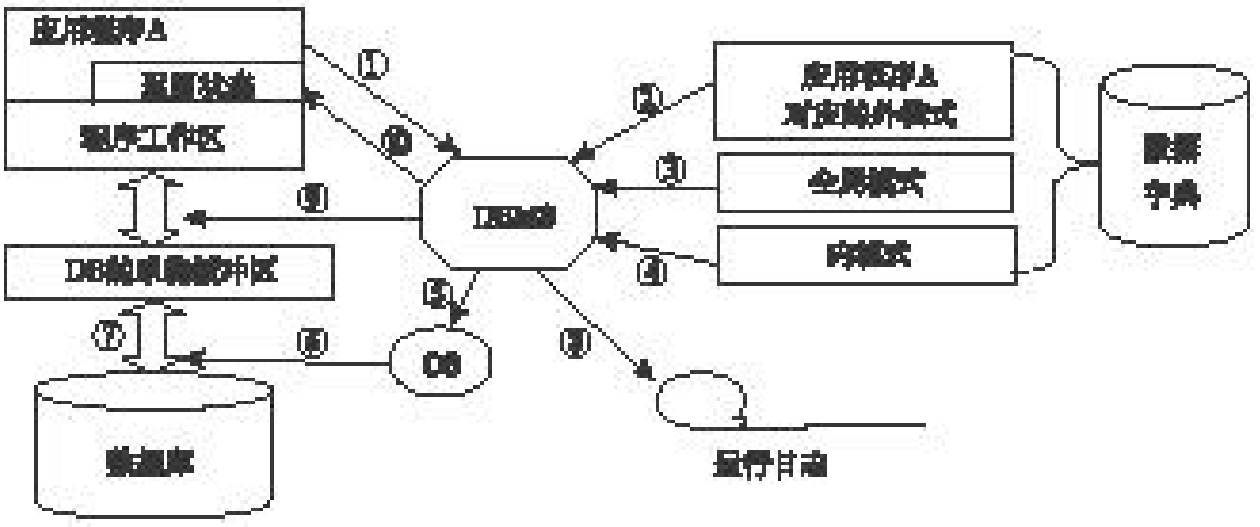


图 4-7 用户访问数据库的主要步骤

①应用程序 A 向 DBMS 发出读取数据请求，并给出记录名及其关键字;

②DBMS 接到请求后，利用应用程序 A 所使用的外模式来分析这一请求;

③DBMS 调用全局模式，进一步分析请求，根据外模式与模式之间变换的定义，决定应读入哪些模式记录；

④DBMS 通过内模式，将数据的逻辑记录转换为实际的物理记录；

⑤DBMS 向操作系统发出读所需物理记录的请求；

⑥操作系统对实际的物理存储设备启动读操作；

⑦读出的记录从保存数据的物理设备送到系统缓冲区；

⑧DBMS 根据模式和外模式的规定，将记录转换为应用程序所需的形式；

⑨DBMS 把数据从系统缓冲区传送到应用程序 A 的程序工作区；

⑩DBMS 向应用程序 A 发出请求执行情况的信息。

以上给出应用程序 A 读取数据库中数据的一般步骤和过程，并体现了三级模式的作用。不同数据库管理系统的操作细节可能存在差异，但其基本过程大体一致。其他数据操作的步骤，如写入数据、修改数据、删除数据等，在总体上与此相近。

第二节　地矿资源属性数据库结构设计

数据库结构的设计是数据库系统设计的核心。其主要任务是确定整个数据库的数据逻辑结构。地矿资源勘查信息系统子系统的设计过程，实质上就是将反映勘查区的有关地矿资源属性的数据，按照系统分析的结果组织成符合于某个特定的数据库管理系统（例如，FoxBASE +、FoxPRO、Oracle 或 SyBASE 等）的数据模式所要求的数据结构。这一设计也称为关系式数据库的逻辑结构设计，下面以煤炭资源勘查区数据库为例加以说明。

一、地矿资源勘查区数据模式的建立

为了使设计出来的数据库能很好地满足勘查工作的需要，在进行具体的数据库逻辑结构设计之前，必须把系统分析中所建立的各个地质矿产模型都转化为相应的一系列数据模式。然后，再将这些反映客观情况的勘查区数据模式转换成 FoxBASE +、FoxPRO、Oracle 或 SyBASE 等要求的数据模式，即逻辑数据模式。

勘查区地矿资源信息本身的复杂性和多样性，决定了其数据模式不可能是单一的。由勘查区地矿模型向勘查区数据模式的转化过程是一个分析、综合的过程。它的大致步骤是：先将地矿模型中各种实体集的全部属性，按其来源和取值方式分成许多类别，再在各个类别中进行比较，合并其中同种或同名的属性，然后再与有关的“标准”、“规范”及现行的生产记录表格作比较，归纳为几个大类，得到了一系列数据模式。每个数据模式便是一个新的实体集，而新实体集内的每个数据项对应着一个属性。

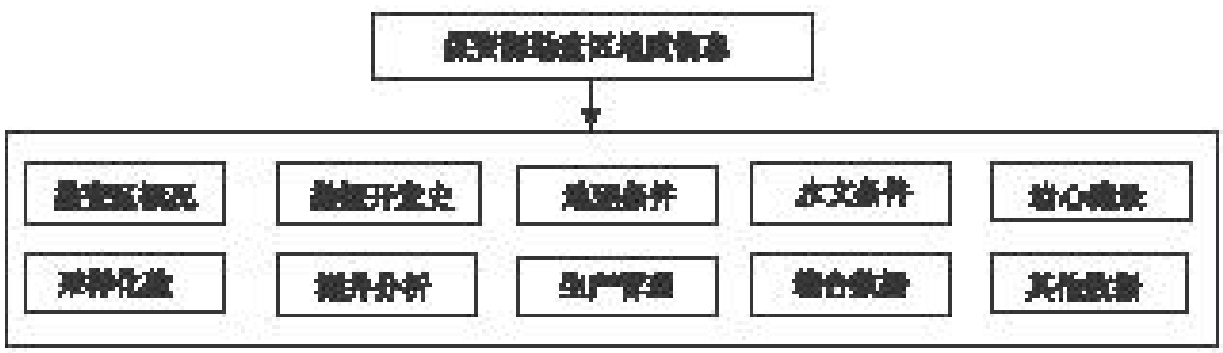


图 4-8　煤炭资源勘查区全局数据模式

这些数据模式的集合，可作为勘查区全局数据模式来看待（图 4-8）。它们一方面成为本数据库采集数据的标准格式，另一方面又成为本数据库系统文件设计的基础。

二、数据文件结构的设计

数据文件结构即通常所说的记录型式，其设计过程就是将勘查区数据模式转换为 Fox-BASE + 、FoxPRO、Oracle 或 SyBASE 等所要求的数据模式的过程。设计中既要顾及具体数据库管理系统的数据组织方式，又要考虑数据的存贮效率和访问效率。其设计方法与步骤，可以通过岩心编录数据模式的转换来说明。

1. 划分实体集，给定初选文件名称

岩心编录资料是勘查区最重要的原始资料之一。煤炭资源勘查区的岩心编录数据模式可以根据工作需要分为沉积岩岩心编录、煤心编录、岩浆岩岩心编录和构造岩岩心编录四个大的实体集。每个实体集还可以再划分为一些小实体集，即实体子集。

实体子集划分的根据主要是实体的类型、数据出现的频度和规范化要求。例如，沉积岩编录这一实体集，可依岩性类别划分为“细碎屑岩描述”、“粗碎屑岩描述”和“碳酸盐岩描述”等 3 个实体子集。考虑到其中的某些属性在岩心中并非经常和同时有值可取，为了节约存贮空间，在细碎屑描述、粗碎屑描述和碳酸盐岩描述等 3 个实体中，仅保留经常且同时出现的基本属性（数据项），而将那些非经常、并非同时出现的属性单独提出来，再加以归纳，合并成 6 个实体子集即沉积构造描述、岩层结核描述、同生角砾描述、异化粒描述、实体化石描述和遗迹化石描述。每个实体子集对应着一个数据文件，因而实体子集的名称便可按规定格式转化为数据文件的名称。

2. 选择描述实体集的全部属性，确定索引关键字

这一步骤实际上是与上一步骤交叉进行的。当前 3 个实体子集按照岩性类别划定后，其相应的属性就应该初步选择出来；而当后 6 个实体子集被分离出来，重新划定之后，描述各个实体子集的全部属性也应该随之确定。

实体子集的属性选择与确定仍然以勘查区数据模型为基础。由于实体子集是同一关系下属性的集合，因而符合某一关系的属性便被选入某一实体子集中去。这里需要着重解决的是一个消除“导出依赖”（James，1982）的问题，即凡是可以由子集内的其他属性计算出来的属性都不应该保留在实体子集内。前 3 个实体子集的属性选择便是这样进行的，当后 6 个实体子集被强制分割出来后，属性的选择也随之变化。这时应先满足后 6 个子集的需要，余下的仍归原来的 3 个实体子集。当然，作为各个实体子集不可缺少的、用于联结文件的共同关键字，则各自保留。例如，当沉积构造描述被从前 3 个实体子集分离出来之后，满足“沉积构造特征”这一关系的属性，即层理级别、层理类型、层系形态、纹层形态、纹层倾角、其他沉积构造、其他沉积构造的发育程度和层面构造等 8 个属性，便随之被划入“同沉积构造描述”子集中，尽管它们同时也都满足“细碎屑岩特征”和“粗碎屑岩特征”这两个关系。作为各个实体子集不可缺少的、用于联结文件的关键字，如钻孔编号、岩层序号等属性，则应当各自保留。

地矿资源变量多属于空间变量，空间定位参数往往是我们查找和检索数据的主要线索。因此，岩心编录数据的主关键字应当是钻孔编号、岩层序号和样品编号以及它们的组合。索引关键字的确定，应在数据模式规范化完成之后。

3. 分析实体集内各属性的依赖关系，进行数据规范化处理

开展地矿资源数据库系统的数据模式规范化工作，一方面是为了减少数据冗余；另一方面是为了提高数据的灵活性。这项工作也是与前两个步骤交叉进行的。在实体集初步划分出来并且初步选择了相应的描述属性之后，数据规范化工作就应着手进行。前一步骤中的属性选择过程，实际上就是数据模式规范化的最初过程。

采用数据规范化技术可以消除各属性之间的不必要牵扯，为插入、修改、删除等数据操作带来极大的方便。而且，被规范化了的数据模式还具有较大的灵活性和扩展能力(李锦锋，1984)。在一个实体集内，各属性之间的关系有6种范式，即第1范式（1NF)、第2范式（2NF)、第3范式（3NF)、BOYCE/CODD范式（BCNF)、第4范式（4NF）和第5范式（5NF)。由第1范式向第5范式规范化过程，实际上是一个数据间关系由函数依赖→完全函数依赖→非传递依赖→非多值依赖→隐含联结依赖的逐步简化过程。

一般地说，范式越高，数据的冗余度越小，独立性越大，可操作性与可修改性也越好。但是，在实际应用中，并非规范化程度越高越好。因为随着范式增高，数据间关系趋向简单，必然引起数据文件急剧增多，加长了查找途径，提高了联结代价，反而降低了访问效率。为此，地矿数据库系统的数据模式规范化的最高目标宜定在第3范式，多数情况下达到第2范式即可。在多数情况下，只要使实体集内部的各属性之间满足完全函数依赖关系，便可满足本数据库系统的要求。在具体的规范化工作中，究竟选择第2范式还是第3范式，可视各个地质模型中各个子型所涉及的属性多寡及属性间传递依赖关系的多寡而定。模型所涉及的属性多且属性间的传递依赖关系也存在多个，则取第2范式，否则取第3范式。岩心编录数据模式正是遵循这一原则，选择了第2范式。下面以此为例，简单说明这种规范化方式。

从地质模型转化而来的岩心鉴定表是一种用户的视图模式。在这种关系模式中，不仅数据相互不独立，冗余度也很高，不具备最低级规范化的第1范式要求。对于一个具体钻孔的岩心描述来说，有描述数据的项通常只占岩心鉴定表中所列的项的10%。而且，不同的岩层和煤层所描述的内容不一样，甚至相同的岩层和煤层所描述的内容也不一样。这些情况，将造成存贮、检索和预处理的困难，严重地影响了数据库的运行，给数据库的管理和维护带来无穷无尽的麻烦。为此，根据实现属性间完全函数依赖的要求，把岩心鉴定表的关系模式作了两次分解，得到如下10个新的关系模式:

岩层一般参数——MDBN502
岩石结构描述——MDBN506
沉积构造描述——MDBN507
碎屑成分描述——MDBN509
同生砾描述——MDBN510
异化粒描述——MDBN511
结核描述——MDBN512
实体化石描述—MDBN513
遗迹化石描述—MDBN514
节理描述——MDBN520

经过这样的分解，就使关系模式达到了第2范式（2NF)，较好地解决了上述几方面的问题。对矿心鉴定表、火成岩鉴定表和构造岩鉴定表，也可进行类似的规范化处理。

为了输入方便，可以设计出一个与原始记录表格一致的输入模块（吴冲龙等，1996），让用户按照已经习惯了的数据表格式，方便地输入数据，而不必按照一般数据库的格式要求重新进行数据整理。这种输入模块能够自动地把所输入的数据分解为规范化的数据文件，然后删除原始数据文件。这一技术措施，实现了数据关系模式的计算机辅助规范化。

经过规范化处理后，全部勘查区数据模式转换成了 8 个部分，共 144 个基础数据记录型式。此外，为了满足煤炭资源勘查工作和盆地分析的当前需要，在不给数据库造成太大负担的前提下，依照“控制冗余”的规则，逆规范化方向，专门设置了 34 个综合数据记录型式，以供快速实现专题检索之用。应该指出的是，只要给定一个合理的计算规则，通过一系列联结运算，这些数据是能从其它 144 个数据文件中得到的。这 178 个数据记录型式的全集合，构成了煤炭资源勘查区点源数据库的全局数据逻辑结构：

（1）勘查区概况，GMBN101。

（2）勘查区开发史，见表 4-2。

表 4-2　勘查区开发史

201	钻探史	207	煤矿调查（一）
202	物探史	208	煤矿调查（二）
203	勘探工程布置	209	矿井调查
204	钻孔设计	210	采空区调查
205	水文孔设计	211	勘查线与钻孔
206	地震工程量设计		

（3）地理条件，见表 4-3。

表 4-3　勘查区地理条件

301	地貌条件	307	工业生产
302	交通运输	308	农业生产
303	降水量统计	309	乡镇人口
304	气象资料	310	测量控制点成果表
305	气象条件	311	工程测量图根锁网精度统计表
306	地震灾害		

（4）水文条件，见表 4-4。

表 4-4　勘查区水文条件

401	区域水文地质	411	河流调查	421	钻孔抽（放）水试验止水情况
402	含水层描述	412	水井调查	422	钻孔注水实验基本参数
403	隔水层描述	413	水井简易抽水试验	423	钻孔注水实验观测记录
404	水文地质观测点	414	地下水长期观测点	424	水位观测记录
405	矿井水文调查	415	地下水长期观测记录	425	钻孔压水试验基本参数
406	矿井涌水量	416	钻孔简易提筒抽水基本参数	426	钻孔压水试验观测记录
407	开采过程的水文地质情况	417	钻孔简易提筒抽水观测记录	427	钻遇含水层情况
408	泉水调查	418	钻孔抽水实验观测参数	428	钻孔简易水文地质观测记录
409	湖泊、水池、水库调查	419	钻孔抽水实验观测记录	429	矿井各开采水平涌水量调查
410	河流	420	钻孔抽水试验综合资料		

(5) 岩心编录，见表 4-5。

表 4-5　勘查区岩心编录

501	钻孔概况	515	火成岩综合描述
502	岩层一般参数	516	火成岩矿物成分描述
503	矿心与矿层（体）描述	517	火成岩包裹体描述
504	煤心与煤层描述	518	火成岩的孔洞构造描述
505	岩性一般参数及其他关联关系	519	接触变质岩描述
506	岩石结构描述	520	节理描述
507	沉积构造描述	521	钻孔控制的断层描述
509	碎屑成分描述	531	岩心综合鉴定表
510	同生砾描述	532	煤心鉴定表
511	异化粒描述	533	沉积岩岩心综合鉴定表
512	结核描述	534	火成岩岩心综合鉴定表
513	实体化石描述	535	变质岩岩心综合鉴定表
514	遗迹化石描述		

(6) 采样分析，见表 4-6。

表 4-6　勘查区采样分析

601	观测点	621	燃点试验成果
602	岩矿样品登记	622	煤的硬度测试成果
603	土样登记	623	煤层瓦斯测试成果
604	水样登记	624	岩石物理力学测试基本参数
605	送样记录	625	岩石物理力学测试成果
606	煤样分析要求（设计）	626	土层物理力学试验结果
607	煤层顶底板力学性质试验要求	627	水质分析基本成果
608	矿井煤质调查成果	628	常见离子分析成果
609	煤岩鉴定成果	629	泉的气体成分分析成果
611	原煤基本分析成果	630	细菌分析成果
612	原煤抽测成果	631	伴生矿产化学成分分析成果
613	风化、氧化煤样特殊测试成果	632	岩石薄片结构鉴定成果
614	精煤基本测试成果	633	岩石薄片石英组分鉴定成果
615	精煤抽测成果	634	岩石薄片长石组分鉴定成果
616	筛分总样化验结果	635	岩石薄片岩屑及重矿物组分鉴定成果
617	筛分试验结果	636	岩石薄片胶结物鉴定成果
618	浮沉试验参数	637	岩石薄片交代矿物鉴定成果
619	浮沉试验结果	638	岩石名称薄片鉴定成果
620	煤尘爆炸性试验成果		

(7) 测井分析，见表 4-7。

表 4-7 勘查区测井分析

701	测井概况	711	测井质量评价表（二）
702	测井数据登录	712	钻孔测斜技术数据表
703	测井综合解释成果（一）	713	钻孔测斜记录表
704	测井综合解释成果（二）	714	数字测斜野外记录表（二）
705	煤层分项解释成果	715	数字测斜野外记录表（二）
706	钻孔地温测量概况	716	数字测斜野外记录表（三）
707	地温测量记录	717	数字测斜野外记录表（四）
708	普通测井技术数据记录表（一）	723	煤层综合解释成果
709	普通测井技术数据记录表（二）	724	含水层综合解释成果
710	测井质量评价表（一）		

(8) 生产管理，见表 4-8。

表 4-8 勘查区生产管理

801	人员编制	806	泵设备情况
802	钻孔事故记录	807	台班记录情况
803	仪器设备情况	808	钻孔结构
804	钻机设备情况	809	套管结构
805	管状水文实验设备情况	810	钻孔取心情况

(9) 综合数据，见表 4-9。

表 4-9 勘查区综合数据

901	钻遇煤层综合数据	918	褶皱构造
902	钻遇煤层一览表	919	同沉积褶皱构造
903	煤层综合数据表	920	断裂构造
904	钻孔地质点投影计算结果	921	同沉积断裂
905	煤层储量计算采用数据	922	伴生矿产
906	储量计算基本参数	923	煤田概况
907	空间三角形网络法参数	924	聚煤盆地
908	储量计算成果表	995	盆地热特征
909	煤田勘查类型定量评价参数	926	聚煤作用
910	含煤段（煤组）综合数据	927	图件
911	富煤带	928	煤炭资源的地质、技术、经济综合评价表
912	钻孔地层统计表	929	断层点测井解释成果表
913	钻遇标志层情况	930	钻孔地质点综合数据
914	标志层综合数据	931	区域地层
915	钻遇断层岩描述	932	区域构造特征（总体特征）
916	含煤岩系	933	岩浆活动特征
917	勘查区构造	934	变质作用特征

4. 定义数据完整性约束

在实体集及其全部属性确定之后，数据库的各数据记录型式也就确定了。随后，便可以对数据的完整性约束进行定义了。

如前所述，数据完整性约束包括系统定义的结构完整性约束和用户定义的一般完整性约束。结构完整性约束主要指维持用户指定的函数依赖关系，特别是包含关键字的函数依赖关系。在设计关系模式时，应当对所处理的属性间关系和背景非常清楚，以便在定义数据库模式时，指明属性间的依赖关系。确定了一个关系模式的关键字，就意味着找到了该关系的决定属性。数据的结构完整性约束，由 DBMS 执行。当每一个元组输入时，都由 DBMS 对其进行检查，若符合所定义的关键字函数依赖关系，就允许装入数据库，否则拒绝装入。这样，就可以确保数据库中的所有关系，都满足业已定义的函数依赖关系。例如，岩层一般参数的数据结构完整性约束，包括建立以钻孔编号和岩层序号为关键字的函数依赖关系和建立以关键字为索引的函数依赖关系。显然，为了保证钻孔数据文件的所有一般参数与关键字具有指定的完全函数依赖关系，每一个钻孔的每一个岩层和岩层序号都必须具有唯一性。

数据的一般完整性约束，主要指每个实体集（子集）内部各属性的字段名、单位、字段类型及其取值范围。这一工作的基础是域分析。属性的字段名即数据项的标准代码。通常为便于查询、联结运算和批处理，在同一数据库内，同名字段的单位及字段类型、取值范围都必须保持一致。但有些时候，为了不同的用途，也允许出现不一致。例如，全硫

表 4-10　岩层一般参数的数据文件结构

数据文件名　MDBN502. DBF 索引文件名: MDBN502. CDX

索引关键字　钻孔编号 + 岩层序号

序号	字段说明	字段名	约束	单位	类型	长度	小数
1	钻孔编号	MDBTAG			C	5	0
2	岩层序号	MDBWAA			C	5	0
3	岩石地层单位	DSB	专项		C	6	0
4	底面深度	MDBWAC		m	N	7	2
5	岩心长度	TKAJAI		m	N	5	2
6	岩心倾角	MDBWAD		度（°）	N	4	1
7	岩石基本名称	MDAEI	专项		C	8	0
8	原生色	YSHBA	专项		C	7	0
9	次生色	YSHBB	专项		C	7	0
10	层厚型式	MDBWAK	专项		C	7	0
11	底部接触关系	MDBWAQ	专项		C	8	0
12	样品编号	PKHFB			C	11	0
13	备注	MDBTZ			C	12	0
14	数据入库日期	MDBTKI		月日年	D	8	0
	* * 总计 * *	TOTAL				102	

(MDCEAA)、煤的发热量(MDCDF)、热稳定性(MDCHM)和透光率(MDDCAR)等数据项，在“煤样分析要求(设计)”这一记录型式中，都属于逻辑型(L)字段，而在“精煤煤质抽测成果”等记录型式中，则都属于数值型字段，其单位和取值范围都有相应的变化。保留这种不一致性，是因为在未来的数据库操作中，用户不会将这两个文件联结起来查询、检索和批处理，用不着担心出现数据的不一致性问题。岩层一般参数的关系模式(表4-10)可以作为数据完整性约束的实例。

数据库的全部数据结构完整性约束和一般完整性约束，在煤炭资源勘查区点源数据库的数据文件结构描述报告中通常都有详细说明。各种数据入库时，其完整性约束将由系统的输入子模式来完成。金属、非金属、石油和天然气勘查、水文和工程勘察以及区域地质测量数据库的结构设计，方法与此相同，但数据模式和数据文件结构的设计必须具体进行。

第三节　地矿资源属性数据库数据字典设计

数据库系统是一个复杂的系统，其中所包含的信息除了用户数据外，还有很多非用户数据信息。例如，模式和子模式的内容、文件间的联系、数据项的长度、类型、用户标识符、口令、索引等等。这些非用户数据是整个数据库系统的情报系统，如果没有它们或它们遭到了破坏，则整个系统将陷入瘫痪状态，即使数据库本身完好无损，也将无济于事。为了使数据库的设计、实现、运行、维护、扩充有一个共同遵循的标准和依据，并且也为了保证数据库的共享性、安全性、完整性、一致性、有效性、可恢复性以及可扩充性，人们在数据库中设置了数据字典，来集中保存这些信息。由于数据字典是描述数据库中各数据属性与组成的数据集合，因此有人把它看作是关于数据库中数据的数据库。

一、数据字典的基本概念

数据字典是整个数据库环境的重要组成部分，是一个管理有关数据库设计、实现、运行和扩充阶段的各种信息的工具，在数据库的生命周期内起着不可或缺的作用。

1. 数据字典的功能和作用

归纳起来，数据字典的功能包括以下几方面：①描述数据库系统的所有对象，如属性、实体、记录类型、数据项、用户标识、口令、物理文件名及其位置、文件组织方法等；②描述数据库系统各种对象之间的交叉联系，如哪个用户使用哪个子模式，哪个记录分配在哪个区域，存贮在哪个物理设备上；③登记所有对象在不同场合、不同视图中的名称对照表；④描述模式、子模式和物理模式的改动情况。

数据字典的作用是：①管理数据资源——数据字典提供了管理和收集数据的方法；②实现数据标准化——在数据库中，数据的名称、格式和涵义等在不同的场合下容易混淆，数据字典提供使之标准化的工具，它可以给这些内容予以统一的名称、格式和涵义；③使系统的描述文体化——所有和系统有关的描述，都可以对数据字典中的信息进行查询、插入、删除和修改；④作为设计的工具——由于数据字典中存放着与数据库有关的各种信息和原始资料，就成为数据库设计的有力工具；⑤控制和管理数据库的存取——数据库在接受每一个对数据库的存取请求时，都要通过检查数据字典用户标识、口令、子模式、模式和物理模式等；⑥是数据库管理员(DBA)观察数据库的眼睛和窗口——利用数据字典进

行各种查询，以便了解系统性能、空间使用状况和各种统计信息，及时掌握数据库的动态。

2．数据字典的类型与存放内容

凡是与数据库系统有关的信息都可以保存在数据字典中。所以，数据字典是数据库本身的信息系统。数据字典本身也是一个内容丰富、功能强劲的庞大软件系统。数据字典依据自身的特点和使用形式，可分为代码数据字典（简称代码字典）、模型数据字典（简称模型字典）、方法数据字典（简称方法字典）、综合数据字典（简称综合字典）和 CASE 技术数据字典（简称 CASE 字典）几种。

二、模型数据字典

模型数据字典包括模型管理字典、关系子模式字典和属性字典等 3 种。

1．模型管理字典（数据文件字典）

模型管理字典用于存放所有关系子模型的项目号、名称、关系名、关键字、关键词等有关信息，同时还作为调用各关系子模式的菜单（用户视图），并且具有模式之联接、选择及参数调用的功能。概念模型管理字典的建立，使整个系统的关系子模型便于修改、扩充和自动联接（吴冲龙等，1996；汪新庆等，1998）。概念模型管理字典结构如表 4-11 所示。

表 4-11　概念模型管理字典结构

索引标识名	索引类型	索引关键字	关键字说明
TABLE _ NAME	普通索引	TABLE _ NAME	表名

字段说明	字段名	类型	长度	小数	单位
表文件名	TABLE _ NAME	C	8	0	
表编号	TABLE _ CODE	C	3	0	
表名	TABLE _ CAPT	C	40	0	
主关键字	PRIMARYKEY	C	24	0	
	主关键字的标识名				
索引标识名	TAG _ NAME	M	4	0	
索引类型	TAG _ TYPE	M	4	0	
索引关键字	ND _ NAME	M	4	0	
关键字汉字名	NC _ NAME	M	4	0	
索引是否为降序	TAG _ DESC	M	4	0	
索引筛选条件	TAG _ FOR	M	4	0	
表路径	TABLE _ PATH	C	40	0	
表有效性文本	TABLE _ ERR	M	4	0	
表有效性规则	TABLE _ RULE	M	4	0	
插入	INS _ TRIG	M	4	0	
	插入触发器表达式				
修改	UPD _ TRIG	M	4	0	
	修改触发器表达式				

续表

索引标识名	索引类型	索引关键字	关键字说明
TABLE _ NAME	普通索引	TABLE _ NAME	表名

字段说明	字段名	类型	长度	小数	单位
删除	DEL _ TRIG	M	4	0	
	删除触发器表达式				
表注释	TABLE _ CMT	M	4	0	
表结构报告	TABLE _ STRU	M	4	0	
密级	P CLASS	N	1	0	
页号	PAGE	N	3	0	
共享属性	EXCL	C	4	0	
备份次数	BACKUP	N	1	0	次
数据备份日期	BACKUD	D	8	0	
总字段数	TOTFIELD	N	3	0	
数据记录总数	TOTREC	N	10	0	
数据更新日期	DATA _ RD	D	8	0	年．月．日
工作区	WORKAREA	N	2	0	
其他相关文件	OTHER _ FILE	M	4	0	
关键字菜单表达式	KEYMENU1	M	4	0	
关键字菜单表达式	KEYMENU2	M	4	0	
关键字菜单表达式	KEYMENU3	M	4	0	
关键字菜单表达式	KEYMENU4	M	4	0	
历史记录备案时间	REFERENCE	D	8	0	
条件表达式	CONDITION	M	4	0	
条件表达式名	CONDITION	M	4	0	
处理公式表达式	WORKEXP	M	4	0	
处理公式表达式名	WORKEXPN	M	4	0	
网络数据接收天数	NETDATAD	N	2	0	
上次接收时间	NETDATAT	D	8	0	
网络数据处理表达式	NETWORK	M	4	0	
备注	MSTRU	M	4	0	
用户字典控制	ZXDC	C	8	0	
字段名	FIELD _ NAME	C	10	0	
其他字典	OTHER DICT	C	8	0	
总计	TOTAL		296		

2. 关系子模式字典

关系子模式字典对数据的关系子模式进行详细描述，其内容有：属性名（字段名）、属性说明（字段说明）、属性类型（字段类型）、字段长度、小数位、单位、数据项约束、属性域、用户界面参数，等等（表 4-12）。对已有的属性域参数，系统将自动从属性域字典中传输过来，若属性域字典没有相应的属性域参数，系统将提示输入其相关参数，然后将其自动存放到属性域字典中，并且传送到关系模式结构描述字典中（汪新庆等，1998a）。

表 4-12　关系子模式描述字典

字段说明	字段名	类型	长度	小数	空值	单位	约束	存储
字段名	FIELD_NAME	C	10	0				
类型	FIELD_TYPE	C	1	0				
长度	FIELD_LEN	N	3	0				
小数位	FIELD_DEC	N	3	0				
空值	FIELD_NULL	L	1	0				
代码转换	FIELD_NOCP	L	1	0				
字段说明	FIELD_CAPT	C	32	0				
单位	U_NAME	C	12	0				
用户字典控制	ZXDC	C	8	0				
存储方式	STOREDMODE	C	1	0				
指定合法性检查	LEGALCHECK	L	1	0				
相关字段名	LINKNAME	C	20	0				
	LINKNUNBER	N	3	0				
总计	TOTAL		97					

3. 属性字典

每一个地矿勘查区的点源数据库通常都有百余个数据子模式，其中所含的属性可达数千余项。有关信息包括项目名称、字段名、字段类型、字段长度、小数位、单位，以及有关数据存贮方式、术语字典控制等数据完整性、一致性、安全性的约束条件，可以全部存放在标准属性字典（表 4-13）中。定性描述的属性，可按标准方式描述和存放约束条件；定量描述的属性，只需以属性域的范围作为约束条件来存放。在关系子模式维护、修改、扩充、重组过程中，属性字典起着恢复和传输各关系子模式属性参数的作用。

在表 4-13 中，ZXDC、STOREDMODE、ZXDCCODE 项参数用于沟通数据与代码字典的相互关系，同时解决数据一致性、数据存贮形式、数据存储和输出之间转换控制的各种问题。

表 4-13　属性字典

索引标识名	索引类型	索引关键字	关键字说明
FIELD _ CAPT	普通索引	FIELD _ CAPT	长度
FIELD _ NAME	普通索引	FIELD _ NAME	长度

字段说明	字段名	类型	长度	小数	空值	单位	约束	存储
字段名	FIELD _ NAME	C	10	0				
类型	FIELD _ TYPE	C	1	0				
长度	FIELD _ LEN	N	3	0				
小数位	FIELD _ DEC	N	3	0				
空值	FIELD _ NULL	L	1	0				
代码转换	FIELD _ NOCP	L	1	0				
缺省值	FIELD _ DEFA	C	16	0				
字段有效性规则	FIELD _ RULE	M	4	0				
字段有效性文本	FIELD _ ERR	M	4	0				
字段说明	FIELD _ CAPT	C	32	0				
单位	U _ NAME	C	12	0				
字段注释	FIELD _ CMT	M	4	0				
用户字典控制	ZXDC	C	8	0				
存储方式	STOREDMODE	C	1	0				
提示控制前导码	ZXDCCODE	C	6	0				
提示控制总长	ZXDCCOUNT	N	1	0				
提示控制步长	ZXDCLEN	N	1	0				
字段显示格式	FIELD _ FORM	C	6	0				
字段输入格式	FIELD _ INPU	C	32	0				
描述前加字	PRE _ DEPICT	C	12	0				
描述后加字	SUF _ DEPICT	C	12	0				
描述方式	DEPICTMODE	C	1	0				
指定合法性检查	LEGALCHECK	L	1	0				
代码	CODE	C	10	0				
相关字段名	LINKNAME	C	10	0				
字段映射的类名	DISPCLASS	C	16	0				
类库路径	CLASSLIB	M	4	0				
提示信息	LIEBIA	M	4	0				
提示信息频率	LIEBIA _ S	M	4	0				
表文件名	TABLE _ NAME	M	4	0				
数据传输	D _ INPUT	M	4	0				
数据效验	V _ INPUT	M	4	0				
数据效验错误信息	V _ MESSAGE	M	4	0				
计算汇总方式	COUNT _ MODE	N	1	0				
计算公式	COUNT _ EXP	M	4	0				
英译名	E _ NAME	C	80	0				
总计	TOTAL		322					

三、代码数据字典

在信息系统中，代码是人和机器的共同语言，是便于进行信息分类、校对、统计和检索的关键，是数据库规范化、标准化的依据。进行代码字典设计，就是具体地为数据库系统设计出可供系统各部分共享的代码系统对比、转换的标准字典。

1. 国家标准术语代码字典的编码库

要使国家标准术语代码直接应用于数据库中，所建立的术语代码库不但要能面对学科类别不同、字段长短不一、选词检索交叉的实际情况；还要在维持一词一意的原则基础上，保证维持术语之间、属性与值之间、术语与代码之间的有机联系，支持 DBMS 对全部术语和代码进行存贮、检索和传输（吴冲龙等，1996），保证属性名、关系结构和属性域的充分共享，解决标准术语代码库与各类、各级地矿数据库的联接问题。

为此，在建立国家标准术语代码字典之前，需要建立一个合理的编码库，统一存储管理各学科的全部术语代码（汪新庆等，1999）；同时还应当采用对数字值赋属性代码的办法，使之单值化（表 4-14）。另外，对少量术语，例如颜色、岩石名称、矿物名称等，在不打破其编码规则的基础上应进行局部调整，使其长度符合数据库系统的编码原则。为使代码简洁，便于应用，代码库中的代码字段长度通常不要超过十位。

表 4-14　地学学科代码

学科代号	学科名	学科代号	学科名
CH	测绘	KC	矿床学
DD	大地构造学	KS	矿山地质与采矿
DH	地球化学	KW	结晶学及矿物学
DM	地貌学	MD	煤地质学
DR	地热地质	PK	固体矿产普查与勘探
DS	地史学及地层学	QD	区域地质调查
DW	地球物理学	SD	数学地质
DZ	地震地质	SW	水文地质学
GC	工程地质学	SY	石油及天然气地质学
GD	古地理学	TK	探矿工程
GZ	构造地质学	WD	外动力地质学
HJ	环境地质	WT	地球物理勘探
HS	火山地质	XY	选矿与冶金
HT	勘查地球化学	YG	遥感地质
HX	化学分析	YK	岩矿鉴定
HY	海洋地质学	YS	岩石学
JJ	地质经济学	YZ	宇宙地质学

地矿术语分类代码国家标准（GB 9649 - 88）所收入的词汇量庞大，不便于装载在一个数据字典中。为了方便用户，需要分级建立代码字典。其中，一级标准术语代码字典用于储存全部地矿术语代码；各个二级标准术语代码字典分别用于储存基础地质调查、各矿

种勘探或各类工程勘察所需的有关地矿术语代码。

2. 一级地矿标准术语代码字典

一级地矿术语代码字典容纳了地矿学科35个专业（表4-14）的120000条国家标准术语代码，其结构见表4-15。国家标准地质矿产术语代码（除岩石地层单位标准术语代码以外）采用混合编码，即前两位采用专业分类助记码（前两个汉字的拼音的第一个字母，见表4-14）第三位至第六位采用字母顺序码，代码长度2位至6位不等，作为数据项属性代码。属性值域代码采用在数据项属性代码的基础上，增加1位至8位数字顺序码或十进制码。

表4-15　一级国家标准地质矿产术语代码字典表

索引标识名	索引类型	索引关键字	关键字说明
C_NAME	普通索引	C_NAME	汉字说明
CODE	普通索引	CODE	代码
GBXK	普通索引	GBXK + CODE	学科分类 + 代码

序	字段说明	字段名	类型	长度	小数	空值	单位	约束	存储
1	代码	CODE	C	10	0				
2	汉字说明	C_NAME	C	40	0				
3	英译名	E_NAME	C	80	0				
4	备注	C_MEMO	C	20	0				
5	类型	C_TYPE	C	1	0				
6	原代码	O_CODE	C	10	0				
7	学科分类	GBXK	C	2	0				
8	控制类	GB_CONT	C	1	0				
9	注解	GB_MEMO	C	20	0				
	总计	TOTAL		185					

3. 二级地矿标准术语代码字典

二级字典存放各矿种、各工程勘探涉及的国家标准地质矿产术语代码，常用词大约都在10000个左右。该类字典所描述的字段，只有术语本身和相应的代码（表4-16）。

表4-16　二级国家标准地质矿产术语代码字典表

索引标识名	索引类型	索引关键字	关键字说明
FIELD_NAME	普通索引	FIELD_NAME	长度
NAME	普通索引	NAME	字段说明

序	字段说明	字段名	类型	长度	小数	空值	单位	约束	存储
1	字段名	FIELD_NAME	C	10	0				
2	术语（说明）	NAME	C	32	0				
3	英译名	E_NAME	C	80	0				
4	类型	FIELD_TYPE	C	1	0				
5	原代码	O_CODE	C	10	0				
	总计	TOTAL		134					

4. 岩石地层单位标准术语代码字典

岩石地层单位标准术语代码采用分类助记码。其中，界系采用习惯字母，统采用 1 位数字顺序码，组采用小写拼音助记码。在不重码的原则下，编码尽量短（如“徐庄组”编码为“Cm^2x”，而不是“Cm^3xz”，因“Cm^2x”已经不重码了）。各二级代码字典所选的岩石地层单位词语，都在 3000 个左右。该字典结构与一级标准术语代码字典相同。

5. 用户定义标准术语代码字典

由于用户对国家标准地质矿产术语代码的熟悉程度有限，在使用时往往需要有提示和帮助，因此有必要设计一种由用户定义的标准术语代码字典（表 4-17）。

表 4-17　用户定义的标准地质术语代码字典表

索引标识名	索引类型	索引关键字	关键字说明
ZXDC	普通索引	ZXDC + FIELD_NAME	用户字典控制 + 代码

序	字段说明	字段名	类型	长度	小数	空值	单位	约束	存储
1	用户字典控制	ZXDC	C	8	0				
2	代码	FIELD_NAME	C	10	0				
3	术语（说明）	NAME	C	24	0				
4	英译名	E_NAME	C	80	0				
5	类型	FIELD_TYPE	C	1	0				
6	次数	NUMBER	I	4	0				
7	选择控制	XZ	L	1	0				
	总计	TOTAL		129					

6. 行政区代码字典

行政区划代码字典中纳入全国所有省（直辖市、自治区）、地区（市）、县（市）的国家标准代码，共计 3384 条。全国行政区划代码采用分类码，即前两位代表省（直辖市、自治区），第 3 位至第 4 位代表地区（市），第 5 位至第 6 位代表县（市）。

7. 地质矿产术语代码字典的应用

在数据库关系数据模式设计与物理模式设计中，标准地矿术语代码字典技术的应用，主要是解决数据存贮形式和数据输入输出与代码字典连接处理问题。这是实现数据库的安全性、完整性、一致性、可恢复性、有效性、可修改性以及可扩充性的重要保证。

(1) 地矿标准术语代码字典与数据库结构描述的转换

通过标准代码字典和标准数据模式字典，可以顺利地实现地矿标准术语代码与数据库结构描述之间的转换。它们之间的相互关系如图 4-9 所示。

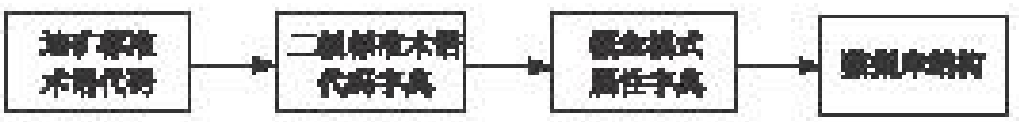


图 4-9　数据结构、代码数据字典与地矿标准术语代码的关系

利用各类代码数据字典对数据进行存取控制和管理，同样需要经过一系列复杂的转换过程。该过程可以用图 4-10 来表示。

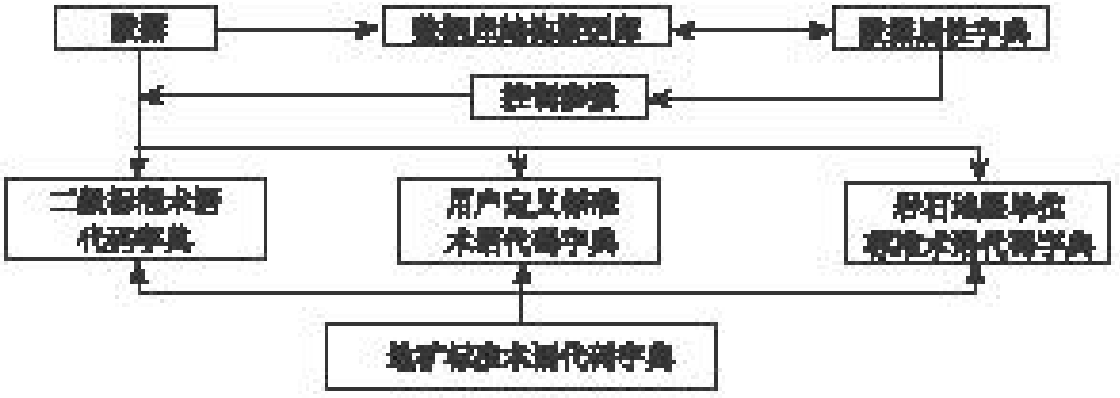


图 4-10　利用各类代码数据字典对数据进行存取控制和管理

(2) 解决数据存贮形式问题

为了实现地矿调查与勘探数据的计算机存贮和处理，野外描述要求尽可能采用标准化和定量方式；而为了方便用户，地矿数据库系统应当能够同时用代码和术语两种形式存贮数据。此外，数据子模式应具备将标准地矿术语自动转换为标准代码，然后存入物理数据库的功能。只有这样做才能既减少计算机的存贮空间，又便于查询检索和实现数据充分共享，方便数据管理员及地质技术人员利用。

在数据模式字典中，描述型数据可采用多数据或单数据代码存贮和术语存贮方式；数字型数据中的准确数据可直接按数值方式存贮。对于数字型数据中的一些不确定值或只给出值域的数据，也可借用描述型数据的存贮方式，将它们当成字符来存贮。在地矿数据库子系统中，规定了不同类型数据的处理方式代号：①“　”代表描述型、单项数据代码存贮；②“，”代表描述型、多项离散性数据代码存贮；③“～”代表描述型、多项连续性或范围值数据代码存贮；④“＊”代表描述型、单项数据术语存贮；⑤“/■”代表描述型、多项离散性数据术语存贮；⑥“-”代表描述型、多项连续性数范围值据术语存贮；⑦“%”代表数字型、不确定值或值域数据当作字符存贮。这 7 种数据存贮方式的代号均存放在模式字典的 STOREDMODE 属性中。

(3) 数据输入输出与代码字典连接处理

有了标准术语分类代码字典和数据模式字典，数据输入输出的标准化就有了保证。但在具体工作中，还需逐一解决输入约束算法及输出的术语与代码的双向转换算法问题。

1) 数据输入约束算法处理。在一般情况下，调用数据模式字典的数据项约束控制参数，可直接对数据输入、输出和数据修改进行数据一致性控制。对于描述性数据项的一致性约束，采用标准化描述方法，统一描述口径，利用标准术语代码字典提示输入。对于部分选词范围比较广，但具区域局限性又无规律可循的数据项，可利用用户定义标准字典来连接。主要目标是实现数据的标准化、规范化、一致性和完整性，其次是简化用户操作，实现面向当前对象的术语输入、代码输入或提示选择输入等自由操作。数据输入约束算法的主要功能体现，是通过数据项属性关联标准数据模式字典，调用数据项的有关参数（包括约束控制参数、数据对象参数等），对不同的数据项类型及不同的参数，确定不同的数据传输处理方法，解决输入、查询、修改等操作过程中数据项约束、输入方式和存贮方式问题，以及实现该方法自身与术语代码、特殊代码字典、用户定义代码字典、数据对象的连接。

2) 数据输出转换算法处理。系统的数据可以有多种存贮方式，但查询结果只能以标

准地矿术语输出。为此，要求在输出之前，各种数据项都必须进行标准术语转换。数据输出的转换依据，是标准数据模式字典中的第 ZXDC 和第 STOREDMODE 项参数。转换算法是通过数据存贮形式参数（STOREDMODE），识别数据值与数据定义的存贮形式是否一致？是标准码还是标准术语？是数据项码还是文字值码？是单项数据还是多项数据？不同的数据形式采用不同的转换方法，ZXDC 参数用于控制术语代码字典库的转换方式。

对于存贮形式为“ * ”、“ /■ ”、“ - ”、“ % ”等 4 种数据格式的数据，可以不经转换而直接输出；对于存贮形式为“ ”、“ , ”、“ ～”等 3 种数据，在转换过程中首先要将数据和存贮形式代号分离，然后识别是简码、全码还是术语。若是简码，就添加属性码并进行转换；若是全码，则直接转换；若是术语，则直接输出。当数据转换输出时，用户定义标准术语字典库优先。对于无法转换的数据，则以原码形式输出。

3）数据合法性检查处理。数据合法性检查处理方法与数据输出转换处理基本上相同。其要领是：将无法转换的数据和数据格式，以及数据存贮形式不一致的数据记录下来，生成数据错误报告，用于检查是数据本身错误还是术语代码字典选词不够。然后，还须对有范围值的数据进行检查。

四、方法数据字典

方法字典就是存放方法模型、参数及方法代码的数据字典。按方法面向的对象不同，可把方法字典分为两类，一类是面向数据项的，称为属性方法字典，一类是面向数据表的，称为方法模型字典。

对于数据库中完整性、一致性保护，一般是建立完整性约束定义模块，在面向对象编程中设置完整性约束定义方法，但是每一个属性的属性域值在不同专业、不同区域、不同部门可能是不同的，它所涉及的数据效验的方法也会不同。例如，在区域地质填图中，区域范围和地质点坐标（经度、纬度）值域是不同的。在中国经度值域为［73，136］，纬度值域为［5，55］，但在不同地区不同图幅的值域完全不同，不可能在模块或方法中定义完整性约束。通过方法字典就可以很好地解决这一问题。

五、综合数据字典

系统综合数据字典包括系统参数字典、系统功能字典、系统功能模块使用记录字典、数据操作记录字典、最后记录状态字典、在线错误处理字典和用户字典。

1. 系统参数字典

系统参数字典用于存放系统的版本信息、系统用户信息及系统环境参数信息。该字典与系统主控模块设置（见第四节）相应的参数，可使系统处于最佳状态。

2. 系统功能字典

系统功能字典的建立，使整个系统的分析、设计和实现等三阶段工作，既相对独立，又紧密连接，真正实现了结构化程序设计的方法和思想。系统功能模型库的结构中包含的信息的涵义已经超出库本身，不仅保存了整个系统运行过程中功能菜单和用户界面信息，而且把系统实现阶段的模块层次划分、层次间关系处理和接口参数设置分离出来，使型式的划分、实施和升级变得非常容易。

3. 系统功能模块使用记录字典

该记录字典用于记录系统运行和调试过程中的各功能模块的使用情况。

4. 数据操作记录字典

数据操作记录字典，是数据库操作的日志，也是数据库系统的“黑匣子”。它用于记录在系统运行过程中，用户对数据所做的各种操作的真实过程。这是数据库系统管理员维护数据的依据，因而对于数据库系统数据的安全是非常重要的。

5. 系统最后记录状态字典

最后记录状态字典用于记录系统运行的最终状态。当用户再次对运行记录进行查询、修改或添加时，系统可以通过该字典自动定位到上次操作的位置，既可以确保工作的连续性，又可以减少不必要的移动记录操作和查找操作。

6. 在线错误处理字典

一个数据库辅助设计系统的优劣不在于是否出错误，而在于能否及时发现错误并准确迅速地将错误情况、错误标志，用适当的方式告知使用者，以便采取措施及时纠正。为此，在系统中需要专门设置出错信息库。

7. 用户字典

用户字典在地矿数据库管理系统中主要起着安全保护作用，存放用户姓名、使用权限、权号、口令及使用情况记载。地矿数据库管理系统只允许在用户字典中有登录的合法用户使用，而不允许非法用户使用。不同的用户由于使用权限不同，其使用范围也受到限制。

六、数据字典在数据库程序设计中的应用

计算机辅助软件工程（Computer Aided soft Engineering）是数据库程序设计中的一项重要技术。随着地矿勘查科学技术水平的提高，地质矿产勘查总是不断地改进仪器设备和工作方式，所获取的数据资料类型和数量随之变化，所采用的数据处理方法、项目、内容也经常发生变化，因此，数据库系统具有显著的不稳定性。换言之，数据库在自己的生命周期中是可变的，主要表现在：①信息定义可变；②字段数量与长短可变；③数据表示方式可变；④访问方法可变；⑤次要关系性质可变；⑥关系系统的闭合性质可变；⑦关系的划分可变；⑧关系的合并可变。这些可变性可以概括为：数据库结构是可变的；数据库应用模式也是可变的；由此而导致所开发的程序也应当是可变的。为了适应客观环境的变化，必须寻找一种系统自适应控制方法，以便实现系统的动态稳定性。只要将数据字典改造成为程序的一部分，利用它在数据库中的数据转换作用，就可以充当数据库自动变更的重要手段（汪新庆等，1999）。

这种经过改造的、可充当数据库自动变更手段的数据字典，就称为计算机辅助数据库软件设计字典（简称为辅助设计字典或 CASE 字典）。CASE 字典是一个内容丰富、功能强劲的庞大软件系统。具有 CASE 字典特征的代码字典、模型字典、方法字典和综合字典都可以作为 CASE 字典使用。CASE 字典一般都由软件平台提供，如 Visual FoxPro 中 SCX（表单），VCX（类库），PRX（报表），PJX（项目），DBC（数据库），MNX（菜单）等等都属于 CASE 字典。

1. CASE 字典的特征

CASE 字典是软件生存期的集成器和通讯机构，能把系统分析和设计阶段的原型建立、检查规格描述，与系统实现和维护阶段的编码、测试、数据库生成、数据规范化及效果分析等任务连接起来，并且能把信息从一个阶段传递到下一阶段，实现系统信息的自动化管理和多用户共享。此外，它还能把项目组每个成员所做的成果进行汇总、分析、归

并，成为统一而完善的软件系统。总之，CASE 字典是数据库辅助设计的核心，是实现下列功能的基础：数据库辅助设计工具的集成；系统规格说明的一致性和完整性控制；系统信息的共享；文档标准化；系统文档的生成；软件的可复用性；项目的管理与控制。

CASE 字典在逻辑上可分为项目库和模型库（系统模型、用户模型），在物理上则分成若干个对应于数据库辅助设计系统软硬件平台的层。在 PC 级微机上开发单用户数据库辅助设计系统，不需要做物理设计，只需要做逻辑设计；若在工作站级微机上，可用一个局部的字典支持单个的开发人员，而用基于主机的字典保存所有的系统信息。相应地，可以在项目级上，用一个中型的字典保存所有的项目信息。

2. CASE 字典的内容

数据库辅助设计系统的主要功能是对系统分析和系统设计的规格说明信息，进行控制和分析，把规格说明信息转换为程序代码和文档。辅助设计字典的功能，则是具体执行规格说明信息的存贮、访问、更新、分析和报告。为此，辅助设计字典保存了所有用于建立、修改、展开和管理整个软件系统的全部信息。这些信息包括：所要解决的问题；问题的定义域；形成的解系统；所用的软件过程；项目的资源和历史；组织的环境。

具体地说，结构化图形、屏幕与菜单的定义、报告的格式、记录的说明、处理逻辑、数据模型、处理模型、源代码、事务规则、项目管理形式、数据元素以及各模块之间的关系，等等，都可以存放在辅助设计字典中。存放在辅助设计字典中的每个信息目标是以其属性来描述的，即：标示符（ID）；代名（号）；类型；文字描述；模块部分；该目标所在的文件；范围值；编辑与推导规则。

此外，每个目标同其他目标的关系及审计信息也存在辅助设计字典中，例如，包括了某个目标所有扩展层的关系信息，所有使用了该目标的目标和交叉引用信息，等等。审计信息与目标所有权认定和变更控制规则有关。辅助设计字典中的内容会随时间推移而不断地增加。随着更多的信息被存贮，建立一个新的软件系统所需的新信息就会越来越少，因为对同一组织或同一应用领域来说，许多项目管理信息、组织模式、数据模型、数据实体，甚至某些程序的处理模型都可能是相同的。

3. 辅助设计字典的窗口

辅助设计字典能自动地跟踪并维护所有存贮在库中的信息，并通过一组标准的报告，让开发人员了解辅助设计字典中的内容。例如，每个数据元素在系统中的什么地方用到？是谁存贮了这些信息？什么时候修改了它们的定义？一共修改了多少次？等等，都可以出现在交叉引用报告中，供开发人员随时使用。辅助设计字典报告的基本类型包括：内容/列表报告；交叉引用/何处使用报告；分析报告；扩展报告；错误报告。

4. 辅助设计字典的信息管理功能

存贮信息的主要目的是为了对所有的系统信息实现自动化管理。在数据库应用软件开发过程中，如果课题组有许多人员协同工作，使用非形式化的手工方法来维护系统信息的完整性是不可能的。为了保证开发人员能以安全的、有序的方式共享系统的信息，就需要有一种形式化和自动化的信息管理设施。

GDCASE 辅助设计字典的信息管理工具为开发人员提供了以下功能：

（1）安全性控制。用口令的方式来限制对辅助设计字典中的内容，或其中的某个部分的访问，用权限的方式来限制访问内容的范围和深度。

(2) 访问优先级。可赋给不同开发人员只读或全面更新的权利。

(3) 型式控制。能跟踪某个系统展开的全过程，并允许多个当前的型式存在，如生产、测试以及开发人员个人的工作集。

(4) 多用户访问。允许多个开发人员同时访问辅助设计字典中的同一个信息，但不能同时更新一个信息。

(5) 变更控制。自动跟踪辅助设计字典中信息修改情况，并能自动地综合多次的变更。

计算机辅助设计字典中存贮着许多类型的信息目标及目标之间的关系，同时也存贮着强大的信息管理工具。因此，我们也可以将计算机辅助设计字典看成是一种由数据库管理系统支持的关系数据库。辅助设计字典与外部词典和数据库之间需要双向的信息流，数据库辅助设计系统必须为进出辅助设计字典的系统信息，提供简捷而恰当的途径。利用数据库管理系统的功能，可以设法让外部词典和数据库联合起来，自动生成辅助设计字典。

第四节　地矿资源属性数据库系统的应用

地矿资源属性数据库系统主要用于存储、管理和处理除空间数据之外的各种描述性数据和测试数据。地矿资源属性数据的“多源、多维、多类、多量、多主题”特征，使得其存储管理工作显得格外复杂。目前，在针对地矿资源属性数据研发的数据库系统中，包含了大量性能较强的地矿资源数据管理和处理的应用系统。

一、地矿资源属性数据库系统的启动和系统管理

1. 系统的启动

地矿资源属性数据库系统安装就绪，即可启动运行。系统的启动通常有三种方式：第一种是直接用鼠标双击系统图标。安装系统时，系统的图标会自动在桌面上显示；第二种是在“运行”窗口中键入本系统的执行文件（图4-11）；■第三种是通过资源管理器找到本系统的执行文件，用鼠标双击后即进入系统。



图 4-11　在“运行”窗口中启动系统

2. 用户权限的查验

地矿资源属性数据库的服务对象，主要是基层勘查单位的各级技术人员、科学研究人员，以及友邻勘查单位、科研单位和上级综合数据库。不同身份的用户对数据的要求各不相同，为了体现内外有别、上下有别，保证数据库的安全与完整性，需要按情况赋予用户相应的查询权限，即采取限制数据项值和给定专用子模式进行存取控制。

为了有效地鉴别用户和核对权限，在定义地矿资源数据库系统子模式时，通常给每个用户制定一个用户标识符——用户代号，并且将这个代号放在数据库用户字典中。地矿资源数据库系统将所有用户分为：数据库管理员（1级用户），应用程序员（2级用户）和终端用户（3级用户）。数据库管理员（1级用户）的代号定义为10～19，共有10个名额，其中10为特等用户管理员。数据库管理员可以使用地矿资源数据库系统中的所有功能模块和调用该数据库系统的所有关系模式，但不能查询、更改其他用户信息（特等用户管理员除外）。应用程序员（2级用户）的代号定义为20～49，共有30个名额。应用

程序员可以使用地矿资源数据库系统中的部分功能模块，也可以查询该数据库中所有关系子模式。并允许产生自己的用户视图模式及数据，但不能进行数据的修改、添加、删除等更新操作。终端用户（3 级用户）的权限定义为 50 ～99（共 50 个）。若终端用户超过 50 个，其代号可用字母扩展。终端用户只允许查询所有关系子模式的数据，不允许添加、修改、删除等更新操作。

用户权限应当在主程序中设置，其过程如下:

```
PUBLIC PCUSERLIMITS, PCPASSWORD, PCUSERNAME, PCPASSWORDMODE
  &&  使用权限,          口令,          用户名,          加密方式
DO FORM &DRW. RS _ P     && 用户身份验证
IF EMPTY (PCUSERNAME)
   QUIT
   ENDIF                 && 未授权用户，系统自动退出
```

用户权限一旦设定，系统启动后便会自动要求用户输入自己的姓名（或代号），查验通过后还要求输入密码口令做进一步证实。如果不能同时通过代号查验和口令证实，用户将被拒绝使用本系统。

3. 系统的管理

地矿资源勘查信息系统的数据库子系统采用多级系统菜单来组织各大功能模块（图 4-12）。各大功能模块内部采用表单控件方式和动态数据模型，来实现丰富的数据操作功能。为了保证系统的正常运行，该系统有一个主要用于系统路径设置、数据路径设置、系统参数设置、系统用户设置、系统日志管理、系统维护以及系统初始化等的管理模块。系统管理模块的总控菜单如图 4-12 左上方所示。

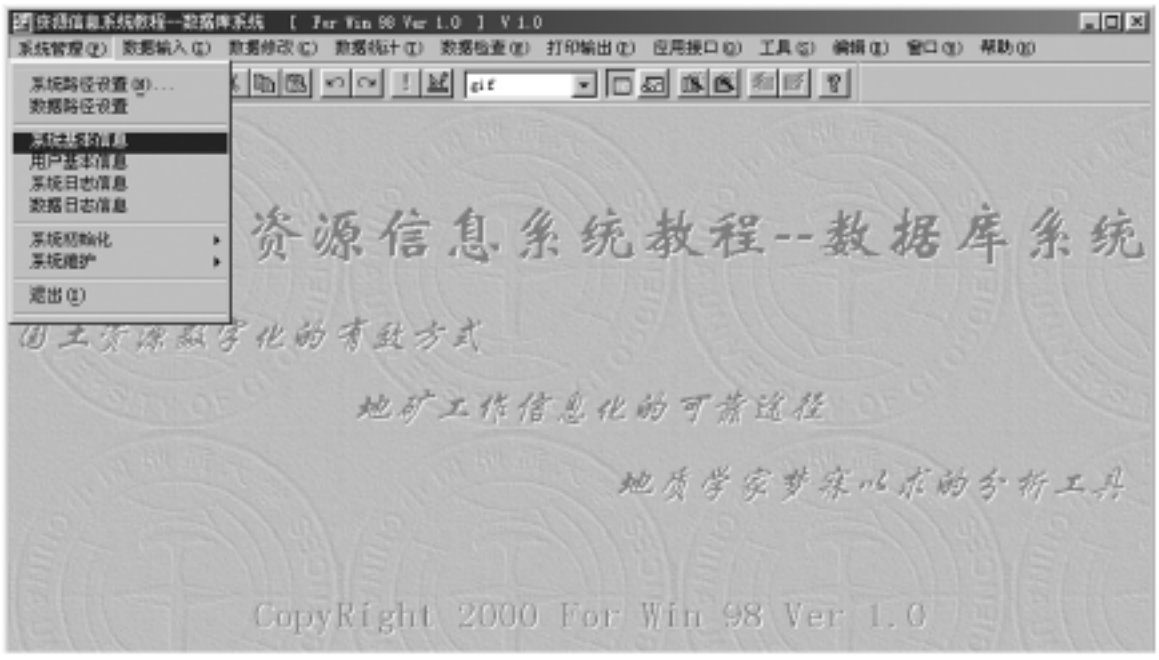


图 4-12　系统管理模块的总控菜单

(1) 环境设定

1) 系统主窗口参数设置。设计具有自己独特风格的系统主界面:

```
SCREEN. CLOSABLE = . T. ;  主窗口可以关闭
SCREEN. MAXBUTTON = . T. ;  主窗口可以最大化
```

SCREEN. MINBUTTON = . T. ; 主窗口可以最小化

SCREEN. MOVABLE = . T. ; 主窗口可以移动

SCREEN. CONTROLBOX = . T. ; 主窗口可以控制

SCREEN. ICON = ‘ MAIN. ICO ’; 主窗口图标

SCREEN. WINDOWSTATE = 2; 主窗口状态正常

SCREEN. VISIBLE = . T. ; 主窗口可见

SCREEN. FONTNAME = ‘ 宋体 ’; 主窗口字体

SCREEN. CAPTION = ‘ 资源数据库系统 ’; 主窗口标题

2) 设计系统运行环境。大多数显示在“ 选项 ”对话框选项卡上的选项，都可以通过编程方式 SET 命令或给系统内存变量指定值来进行修改。例如，在命令窗口输入 SET DATE TO ANSI，则将日期格式设为年．月．日。如果使用 SET 命令配置环境，设置仅在本次 VFP 运行期间有效，退出后设置自动失效。例如:

SET ANSI OFF && 确定 SQL 字符串比较方式 OFF 指定较短的字符串不要填充空格

SET AUTOSAVE ON && 无论什么时候退出 READ 命令或返回到命令窗口，指定缓冲区的数据都要保存到磁盘上

SET BELL OFF&& 关掉铃声

SET BROWSEIME ON && 指定在“ 浏览 ”窗口中定位到一个文本框时，显示输入法

* SET CARRY ON && 当使用 INSERT APPEND 和 BROWSE 命令时创建新记录时，便将当前记录数据复制到新记录中

SET CARRY OFF && 当使用 INSERT APPEND 和 BROWSE 命令时创建新记录时，不将当前记录数据复制到新记录中

SET CENTURY ON && 指定在包含 10 个字符的日期格式中，用四位数字表示年

SET CENTURY TO 19 ROLLOVER 60 && 大于等于 60 年的年份被认为是当前世纪，小于该年的年份被认为是下个世纪

SET CLOCK STATUS && 在图形状态栏显示时钟。执行 SET STATUS BAR ON 命令可显示图形状态栏

SET COMPATIBLE OFF && 允许运行 FOXBASE + 创建的程序而无须作任何操作

SET CONFIRM OFF && 通过键入最后一个字符来退出文本框

* SET CONSOLE OFF && 不将输出送到 VF 主窗口或活动的用户窗口

SET CONSOLE ON && 将输出送到 VF 主窗口或活动的用户窗口

SET DATE TO ANSI && 指定日期表达式和日期时间表达式的显示格式为 YY. MM. DD

SET DECIMALS TO 2 && 指定数值表达式的小数位数

SET DELETED ON

……

(2) 文件组织

即使是一个很小的应用程序，也可能涉及多种类型的文件，如 DBC 数据库、DBF 表以及菜单、表单、报表、位图等。如果把这些文件都放在一个目录下，将会给以后的修改、维护工作带来很大的不便。因此，需要建立一个如图 4-13 的层次清晰的目录结构，让不同类型的文件各归其所。

(3) 路径设置

数据库子系统规模庞大，为了便于组织、管理和维护，必须把各种模型分类存放，并且建立其联系和查找路径。在本系统中，路径变量以 DR 字母开头，存放在 DRV. MEM 变量文件中（表 4-18）。

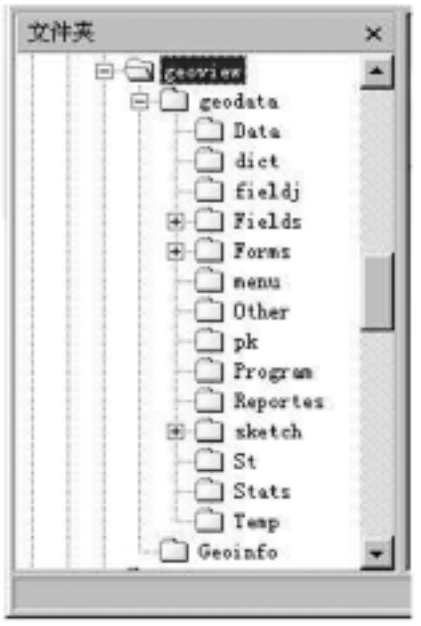


图 4-13 目录结构

1）系统工作路径设置。系统的工作路径可以根据工作需要事先设置，也可以根据自己的习惯、计算机条件和工作对象的变化来调整路径设置，甚至可以将资源数据库子系统安装到指定目录下，并配置自己的工作路径（详见后面的系统管理部分）。具体方法是：选中文件目录中（图 4-13）的路径设置项，调出“系统路径设置单表”（图 4-14），直接查看或修改原有的路径设置。

表 4-18 路径变量及其说明

路径名称	路径变量	缺省值
系统路径	DR	[D: \] GEOVIEW\ GEODATA\
系统名称	DRMC	资源数据库系统教程 - 实习系统
数据路径	DRM	[D: \] GEOVIEW\ GEODATA\ DATA\
网络数据路径	DRN	[D: \] GEOVIEW\ GEODATA\
结构模型库路径	DRS	[D: \] GEOVIEW\ GEODATA\ ST\
专题路径	DRJ	[D: \] GEOVIEW\ GEODATA\ JT\
报告路径	DRK	[D: \] GEOVIEW\ GEODATA\ PK\
源文件路径	DRP	[D: \] GEOVIEW\ GEODATA\ PRG\
标准字典路径	DRG	[D: \] GEOVIEW\ GEODATA\ DICT\
用户视图模型路径	DRZ	[D: \] GEOVIEW\ GEODATA\ MODEL\
绘图接口路径	DRT	[D: \] SXPM\
用户临时处理路径	DRY	[D: \] GEOVIEW\ GEODATA\ TEMP\
图像信息路径	DRTX	[D: \] GEOVIEW\ GEODATA\ DATA\ PCX\
二次开发系统路径	DRW	[D: \] SYS\
一级国标字典路径	DRGB	[D: \] GB\ DICT\
GIS 数据路径	DRGIS	C: \ ESRI\
统计路径	DRTJ	[D: \] GEOVIEW\ GEODATA\ STATS\
系统类库路径	DRCLASS	[D: \] SYS\ CLASS\

系统工作的路径通常定义为全局变量，并设置为缺省值和自动搜索。系统使用者可视硬盘分区情况对“用户视图路径、系统类库路径、临时路径、系统报告路径”等做灵活的设置，其他方面建议使用系统缺省设置。对本系统不太熟悉的用户最好全部使用系统缺省设置。设置完毕，用鼠标单击“存盘退出”，修改结果即被保存了。在系统第一次运行时会自动执行路径设置功能，依系统安装所在盘符设置好各个路径的缺省值。

2）数据路径设置。一个资源数据库系统往往管理着不只一套数据。为了便于管理和查询检索，同样需要准确地设置数据的存取路径。其方法如图 4-15 所示。



图 4-14　系统路径设置



图 4-15　数据对象及其数据路径设置

(4) 设置全局控制变量

全局控制变量的作用在于有效地组织、传递和控制主控模块与过程之间、模块与模型之间的信息。全局变量是整个系统所公用的，它的值可以被任何过程所引用或改变，因此，能够直接将信息带入所调用的过程，并将调用结果带回调用程序。例如，在多个过程中对同一指定的数据库进行操作时，所使用的数据库名及其相应的结构文件名、索引文件名、关键字等内存变量，就是整个系统的公共变量。将它们定义为全局变量，就不用在各个过程中再定义成实体参数，也不需要有对应的形式参数及对其赋值。这样做可以简化程序设计，并使程序易于阅读和调试。在地矿资源属性数据库系统中，除了所有模块和过程共同使用的信息或参数（表 4-19）之外，系统路径、工作区和用户权限等，都应统一定义并设置为全局变量。

表 4-19　系统全局变量设置

WFILE	数据库菜单文件名	PSYDZ	单位地址
DFILE	数据库文件名	PSYDH	电话号码
SFILE	数据库结构文件名	FIRSTKEY	主关键字
NFILE	数据库索引文件名	FIRSTVALUE	主关键字值
DBYE	数据库名称	NFORMSET	表单集数
NCNAME	索引关键字说明	MAINFILE	主表名
NDNAME	索引关键字说明	SUBFILE	子表名
YER	年月日	TABLE _ SET	表模型字典
TT	条件说明	TABLE _ SUB	表单集表
TTT	条件值	CTPARA	参数项数
TQH	用户权号	FIRSTKEYNA	主关键字名称
PSYDW	填报或使用单位	FIRSTKEYLEN	主关键字长度
PSYFZ	单位负责人	KEYTOT	关键字数

(5) 内部环境参数设置

系统内部环境参数可以为系统提供多种功能设置和开关设置。本系统的环境参数存放在系统参数字典 WMIS 中，其说明见表 4-20。

表 4-20　系统环境参数说明

变量名	值地址	设置说明	
		ON	OFF
S_DMTX	WMIS.SDMTX	混合编码体系	纯码编码体系
S_ENGLISH	WMIS.SENGLISH	英文版	中文版
S_AUTO	WMIS.SAUTO	自动处理	人工处理
S_WRITE	WMIS.SWRITE	修改知识库	不修改知识库
S_CHECK	WMIS.SCHECK	自动检查数据	不自动检查
S_ATTMODE	WMIS.SATTMODE	指定属性描述	全部属性描述
S_TSXS	WMIS.STSXS	指定提示显示	全部提示显示
S_CHILD	WMIS.SCHILD	父表中包含子表	父表中不包含子表
S_PARENT	WMIS.SPARENT	子表中带父表	子表中不带父表
S_DELETE	WMIS.SDELETE	显示删除记录	不显示删除记录
S_DOUBLE	WMIS.SDOUBLE	双层描述	单层描述
S_SHORTCUT	WMIS.SSHORTCUT	快捷键开启	快捷键关闭

(6) 工作区设置

地矿资源勘查信息系统所涉及的模型种类和数量很多，而多数模型都以数据库方式体现，因此每打开一个文件，就必须给它指定一个工作区。在系统比较庞大、数据又比较复杂时，往往需要同时打开多个数据文件，并对数据进行综合处理。一般数据库管理系统都设置足够的工作区供使用，例如，Foxpro For Windows 就有 32767 个工作区。为了提高系统的运行速度，避免出现紊乱，应当对工作区进行统一设置（表 4-21）。

表 4-21　系统工作区设置

工作区	数据库名称	文件名	索引标识
1 - 100	基础数据表调用	XXXXXXX	
101 - 180	数据菜单及辅助表调用	XXXXXXX	
201	基础数据表属性字典	TSXXT.DBF	TABLE_NAME
202	统计数据表属性字典		TABLE_NAME
203	区辅助数据表属性字典	TSYSDBF.DBF	TABLE_NAME
204	自由数据表属性字典	TABLEPRO.DBF	TABLE_NAME
231	系统用户库调用	USER.DBF	
232	系统使用日志	RS_PZ.DBF	
239	错误处理库调用	ERRHELP.DBF	ERRNUM
241	数据结构属性字典	DICT.DBF	FIELD_NAME
242	地质术语代码（二级）标准字典	DICTMD.DBF	FIELD_NAME
243	用户字典库调用	ZXDICT.DBF	ZXDC
244	岩石地层单位标准字典	DSB.DBF	CODE
245	地质术语代码（一级）标准字典	GB.DBF	CODE
246	中国行政区代码字典	GB2260.DBF	CODE
247	地质学科代码字典	GBXK.DBF	CODE

(7) 出错处理

即使进行了大量的测试和调试，也不能保证程序不出错。比如在处理一个数据表时，发生了停电，就会毁坏这个表。那么以后当程序中用到这个表时，就会出现错误。当过程中的代码发生错误时，VFP 将检查与 ON ERROR 例程相关的错误处理代码。如果 ON ERROR 例程不存在，VFP 就显示默认的错误信息。在 ON ERROR 后面，可以包含任意有效的 VFP 命令或者表达式，但一般情况是调用一个错误处理过程或程序:

ON ERROR DO &DRW. ERRHAND WITH ERROR ( ), MESSAGE ( ), MESSAGE (1), ;

SYS (16), LINENO (1), PROG (1)

(8) 其他设置和操作

1) 搜索路径设置:

SET PATH TO &DRD; &DRZ, &DRG; &DRW; &DRCLASS; &DRFY && 设置自动搜索路径

2) 类库调用:

SET CLASSLIB TO &DRCLASS \ WIZSTYLE, &DRCLASS \ MYCLASS ADDI && 打开包含类定义的 . VCX

3) 过程文件打开和外部库文件引用:

SET PROC TO &DRW. MAINPROC, &DRW. SPECIALTY 打开系统内部过程文件和专业函数库

SET LIBRARY TO && 打开一个外部的 API (应用程序接口) 库文件

4) 打开系统缺省数据库:

```
IF ! EMPTY (DATAFILE)
    IF FILE (DATAFILE + '.DBC')
        OPEN DATABASE (DATAFILE)
        SET DATABASE TO (DATAFILE)
    ENDIF
ENDIF
```

5) 打开系统字典

```
= LOADDICT ( )
```

6) 打开系统主菜单

```
DO MAINMNUE. MPR
```

7) 事件驱动

```
READ EVENTS
```

8) 关闭系统

```
CLOSE DATABASE
SET SYSMENU TO DEFAULT
RETURN
```

二、数据输入、查错与修改

数据输入、查询、修改是数据库系统最基本的操作功能。其中，提高数据输入质量，

是提高系统运行与服务质量的基础。地矿资源数据库系统的数据输入模块充分考虑了地矿资源数据的多源、多类、多量特征，以及多主题综合利用的特征，可以同时存贮并管理同一地区的不同资源数据，其功能设定以满足用户需求为原则。

1．数据输入模块的应用

(1) 通用输入模块

所谓通用模块，是指适合于整个数据库的全部数据子模式，并适用于调整或扩展后的数据模式。输入模块与数据子模式是一对多的关系，并且不受数据子模式变化的影响。它采用命令文件方式，但具有按数据文件的格式输入数据的功能。

按照数据形式，数据输入方式可分为术语输入、代码输入和选择输入三种；按照数据组织状况，数据输入方式可分为用户模式输入和规范模式输入两种；而按照操作方法，数据输入方式可分为直接输入、预处理输入和多机输入三种。它们之间的关系是互补并且交叉的。要按照传统方式逐一地完成数百个数据子模式，即数百个数据库文件的八种数据输入方式的编程工作，而且还要满足数据库动态变化的要求，这在实际上是不可能做到的。因此，需要开发出一种具有强大自适应功能、能够面对不同资源数据采集和入库者的数据输入模块。

该模块的命令文件有很强的自学习和自适应功能，一个命令文件能够对所有数据文件输入数据，不因数据文件及其结构变动而改变。选用该模块输入数据时，系统将提供数据文件选择菜单（图 4-16），并且询问是输入全部数据项还是部分数据项。若选择部分数据项输入，系统将提供数据文件的数据项菜单，随后自动计算并显示屏幕输入格式。



图 4-16　选择数据表

(2) 用户视图形式输入模块

如何用最快速最有效的方法实现数据入库，是地矿资源数据库系统能否顺利建立的“瓶颈”问题。解决这个问题的最佳方案是设法让用户按地矿勘查数据的原始格式——用

户视图入库，而将数据整理工作交给计算机来完成。

在设计数据输入模块时，通常考虑用户和数据两方面的特点。从用户角度看，初级用户通常希望有完善的输入提示和数据自动检验功能，以便降低出错率；高级用户则要求有简洁的输入格式和快速的输入响应，以便提高工作效率。从数据角度看，地质勘查工作所获取的数据往往以综合的表格形式填写。例如，岩心鉴定表、化学成分分析表、选矿试验表和各种力学测试表，与数据子模式规定的格式差别很大，属于用户模式，不便直接输入。如果进行人工转换再输入，既会增加大量人力、物力的投入，又要浪费时间，而且可能会扩大数据出错率并破坏数据一致性、完整性。为此，需要设计出与用户模式数据构成一一对应关系的专用数据子模式，设置计算机辅助数据规范化功能模块，还要求采用“动态数据模型”和“逆向规范技术”来开发数据采集模块。地矿资源勘查数据库管理系统的用户视图格式（原始资料格式）输入模块，能根据用户选择输入的数据项，自动进行分析处理并完成相应的数据分配设计。用户不但可以很方便地进行原始地矿数据输入，还可以随时修改或调整数据项或有关参数，而且仍由模块自动跟踪调整。

(3) 与《地质矿产分类术语代码（GB 9649—88)》的连接

能否使地矿勘查数据库管理系统与国家标准《地质矿产分类术语代码（GB 9649 - 88)》字典实现有效连接，是提高本系统共享性的重要环节。

为了使数据输入输出标准化，在地矿数据库管理系统中，分别为各专业和各矿种设计了地矿分类术语代码的二级库和用户定义的标准术语代码库。每个二级库选入该专业或该矿种常用词 10000 个左右。在输入、输出的提示和转换过程中，各功能模块都力求体现用户定义优先准则。利用数据输入模块进行数据字典连接及数据存贮时，既允许使用国标全码和文字值码，又允许使用简码和标准术语。例如：数据项煤层稳定性（MDAGCO）的文字值及其全码为：稳定（MDAGCO1)、较稳定（MDAGCO2)、不稳定（MDAGCO3)、极不稳定（MDAGCO4)。当输入的描述内容是“煤层为较稳定”时，可以输入“MDAG-CO2”，也可以仅输入“2”或“较稳定”。若描述内容为复合值时，输入功能模块将自动以文字值的隶属简码表示。例如：当输入的描述内容是“煤层稳定到不稳定”时，系统将提示输入“稳定”和“不稳定”，计算机随后存贮下来的则是“1 - 3”；输出时系统将自动转换为“稳定—不稳定”。

在地矿资源数据库管理系统中，全码和简码是混合使用的。通常在单值时使用全码，多值时使用简码。这样，既可节省存贮空间，又可提高输出速度，数据值还可以参与计算处理。仍以煤层稳定性为例：煤层稳定性这一数据项，在设计数据模式时定义为字段名 = "MDAGCO"，类型 = 字符型，长度 = 7 位，二级国标控制码 = "MDAGCO"，输单值存贮全码正好占据 7 位；输多值时，7 位长可以表示四个值，如"1 - 2 - 3 - 4"。当数据为单值时输出速度较高，为多值时输出速度降低但节省空间。为了提高输出速度，输出处理被设计成可控制转换输出；为了节约存贮空间，规定以汉字表达的标准术语尽量不直接使用，仅下列两种情况例外：①此数据项不会参与数据计算和处理；②标准术语很短且不会产生误解，隶属的文字值术语量很少，不会有多值或复合值出现，比如：有、无、是、非、强、弱，等等。若数据项是字符型并设置了约束，屏幕下方将显示出该项数据的常用术语和代码，作为输入提示。

在本模块支配下，输入的数据在系统内以代码方式存贮，而以标准术语方式显示及输

出。进入代码输入状态时，以代码形式输入和存储，而以标准术语形式显示及输出。

2．数据输入模块应用实例

(1) 露头地质点数据录入模块的应用

为了便于使用，露头地质点数据录入采用单一界面（图 4-17），内含四个部分：①系统菜单；②系统工具栏；③观测点概况表单；④点上详细描述表单。一些辅助功能隐藏在菜单和工具栏中，用法已在前面介绍过。该模块可由 ArcView 调用，录入该点的属性数据。

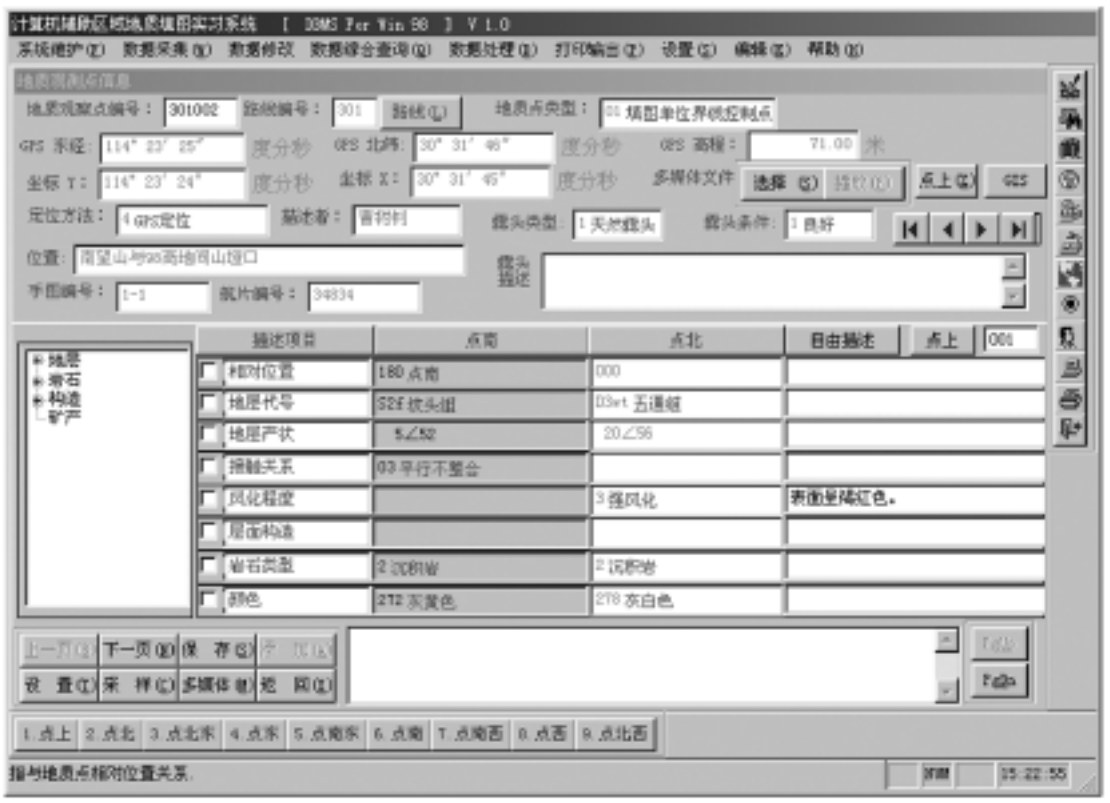


图 4-17　区域地质调查中的露头地质点数据采集界面

(2) 路线数据录入模块的应用

基本的路线数据通常在绘制或设计路线时录入。在进入系统主界面后，单击“路线(L)”按钮可打开野外路线描述表（图 4-18）。确定路线信息无误后，便可在地质观测点表单中填入点上的基本数据。其中地质观察点编号、路线编号、经纬度、描述者、手图编号、航片编号等，由系统根据“每天工作设置”表中的信息自动填入，作业者可根据实际情况修改。其他地质点类型、海拔高程、定位方法、位置、露头条件、露头特征和小结要具体填写。

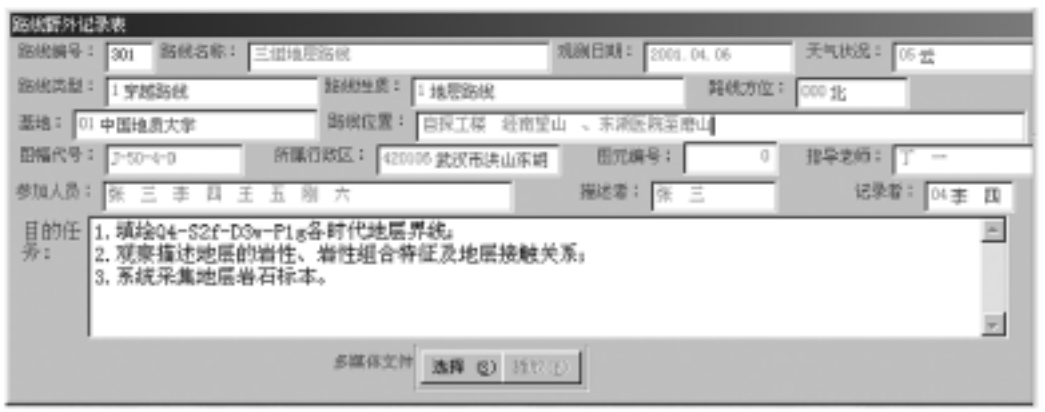


图 4-18　野外地质填图的路线数据采集界面

为了尽量减少野外鼠标操作，系统提供了大量快捷键：①操作系统原有的快捷键，如 Ctrl + C —复制，Ctrl + V（粘贴），Ctrl + F（查找），Alt + F4（关机）等，其含义和用法照旧；②系统界面上的命令按钮的快捷键，由 Alt + 按钮上标明的一个字母组合而成。这个字母一般用括号括住并带下划线；③提示录入工具条上的快捷键，在数据录入时光标移入某些文本框时，在屏幕下方即出现类似于汉字输入法的提示条，这时只需用提示术语前面所标的数字键选择即可录入标准的代码，如图 4-19，敲“3”，“手图定位”术语即被录入。



图 4-19　野外快捷输入的选择提示

(3) 钻孔数据录入模块

该模块可用于钻孔数据、边坡数据、平硐数据的录入、查询和浏览，其界面如图 4-20。

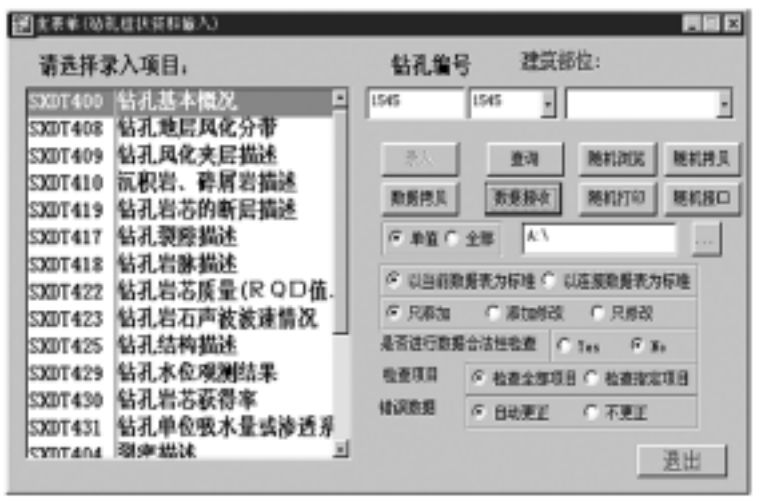


图 4-20　钻孔（井）数据采集界面

在“请选择录入项目”列表框中列出了钻孔中需要录入的项目，作业者先在里面选定要操作的项目，然后在表单右上方的钻孔编号文本框中键入欲录入的孔号（已有记录的孔号也可通过钻孔编号组合框选择孔号）。最后单击“录入”可进入相应项目的录入界面；单击“查询”可进入相应项目的查询界面；单击“浏览”，弹出字段选择表单，用鼠标在其中选定欲浏览字段或全部字段，再单击“确定”按钮，即可对相应字段进行浏览。

3. 数据检查模块及其应用

为了保证所录入数据的准确性，提高数据的可信度，地矿资源数据库系统设置了数据检查功能模块。该模块可以对由用户录入的数据进行各种检查，然后提示出错的信息。根据这些信息，采用系统提供的各种修改功能子模块，可以方便地修改数据错误。这些功能子模块还能查出经计算处理后与原始数不一致的数值型数据项，并显示输出。

(1) 数据检查的基本类型及检查目的

地矿资源数据库对勘查数据的检查类型包括：数据完备性检查、数据规范性检查及数据的特殊检查等。

1) 数据完整性检查。数据完整性检查是指系统对录入的数据项是否完整、有无缺失进行检查并给出提示。又可分为：①缺失父编号的检查。造成父编号缺失起因于录入人员疏忽或失误，只录入内部数据项而未在其主关键字对应的父表中录入相应的信息。若缺失此编号，系统将无法检索和提取该区的内部数据。②缺少整库内部数据的检查。由于录入人员失误或者原始数据表有缺省，录入系统的数据虽有预测区或勘探区的封面编号及部分内部数据，但还有些数据整库地缺失（表现为库中无数据）。③单表中缺少编号的检查。目的是查出哪些库中虽有数据却缺少作为关键字的预测区或勘探区编号。④单库内数据项缺失的检查。在某些预测区或勘探区中，部分数据项的对应库虽已录入数据，但不完整或有残缺。

2) 数据的规范性检查。数据的规范性检查是指系统对录入的数据（主要是字符型数

据）的合法性进行检查，目的是通过与系统中标准库数据进行对比，找出在形式上不合法的数据项，并指出这些数据所在的预测区或勘探区及其对应库名和库中项目序号，重点是地层单位和矿层、矿体编号。

3）数据的特殊检查。该项检查针对数值型数据，主要是对各预测区或勘探区的各级储量进行快速核查，并做出统计，提示出各级储量或总量的计算值和原始值不一致的预测区或勘探区。

（2）数据检查和处理方法

数据检查的一般顺序是：数据完整性检查、数据规范性检查、数据的特殊检查。数据检查后的处理，即是指对各类检查结果的分析和应用，亦即利用检查后提示的信息进行缺失数据的补充、错误数据的修改，以及诸如储量等重要数据的核查。

1）数据完整性检查和处理方法。先通过主菜单进入数据检查模块，再选定“3 - 数据完整性检查”子模块。打开数据完整性检查子模块后，依照界面右上方的提示，可以逐一检查相应的内容，并且显示和打印出检查结果，提供操作者对照修改。

2）数据规范性检查。这项检查是运用主菜单中的数据录入模块和数据检查模块进行。其中，数据检查模块有“全部数据检查”和“单库数据检查”等两项功能。

4. 数据修改模块及其应用

在数据库系统中，数据修改功能模块起着维护数据的作用，主要用于处理以下几种情况：①同类数据在大范围内出现同样错误或数据不一致；②数据子模式之间的数据不一致；③随机性数据错误及数据不一致；④不同数据项以及不同数据库文件之间的数据处理与传输；⑤数据项及数据与模式的分解和合并过程中的数据传输。

（1）选择修改模块

选择修改模块可以面对任意数量的数据项和任意范围的记录，能对有约束的数据实现标准化控制修改，而且能对同类错误进行人工快速修改。

数据修改模块应当允许采用给定条件索引和随机条件索引两种方式进行作业。前者是指用户事先给出一个索引关键字，修改模块以此为依据进行连续检索并修改，因此，又称为索引关键字修改；后者是指用户不必事先给出索引关键字，修改模块以作业中随机给出的一个或几个索引关键字为依据，进行循环检索并修改。

按照上述流程设计的模块，能根据数据模式的关键字自动进行判断处理，并设置默认的关键字段个数和关键字。关键字在这里的含义只针对全屏幕处理过程，当用户选择全部数据项时，模块将会自动处理数据项参数及屏幕输出关系；若选择部分数据，模块则在用户选完所要处理的数据项之后，再处理数据项参数及屏幕输出关系。用户所选择的关键字，可以是数据模式的索引关键字，也可以不是数据模式的索引关键字。

（2）多功能修改模块

多功能修改适用于单数据模式，实现过程主要使用运行变换技术，修改方式为多记录、多数据项横竖双向移动修改。实际上，修改通过标准化约束控制手段来进行，而修改范围的控制采用索引约束控制或自由条件控制。多功能修改的数据模式选择、关键字的确定、修改数据项的确定、修改方式选择、修改范围的确定等操作过程，与选择修改相同。假定用户在钻孔编号菜单中，选定 16104 钻孔，屏幕显示如图 4-21 所示。

多功能修改模块的使用准则是：若遇到相同数据表中的相同字段有不同的值，以本



图 4-21　多功能修改模块的界面

机当前库中值为准；而如果选择“以连接数据表为标准”，则当遇到相同数据表中的相同字段有不同的值，就以他机源数据为准修改本机库；如果选择“只添加”，则遇到本机没有的记录，就自动从他机调入；如果选择“添加修改”，则遇到相同数据表中有不同记录，就自动将添加缺少部分，若遇到相同数据表中的同一条记录有不同字段值，系统也会根据上面选定的以本机或他机为标准来修改；如果选择“只修改”，则当遇到相同数据表中同一条记录有不同字段值，系统就根据上述约定分别以本机或他机为准来修改，不同记录则不做处理。

三、专题数据的提取与组织

地矿资源数据库系统可以直接向功能模块提供数据，也可以通过专门的接口模块来为一些专题应用模块，如各种地矿图件编绘（图 4-22）、资料综合整理、储量计算和资源统计预测模块，以及为某个特定的 GIS 提供数据。

1. 钻孔柱状图数据接口文件的生成

首先从图 4-22 所示绘图接口模块的总控菜单中选中“钻孔柱状图”，进入如图 4-23 所示的界面。然后按照下列步骤生成数据接口文件:

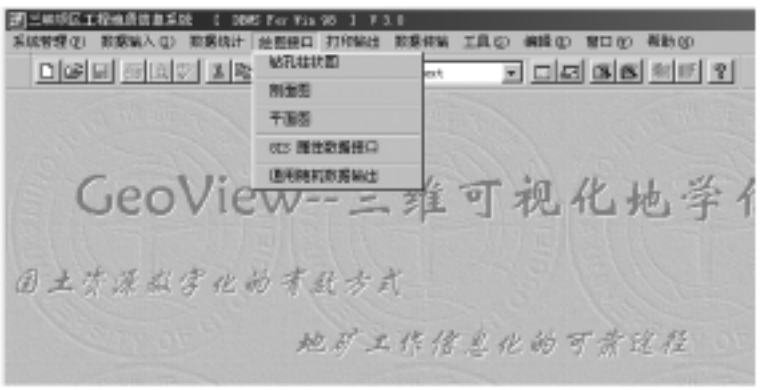


图 4-22　绘图接口模块的总控菜单

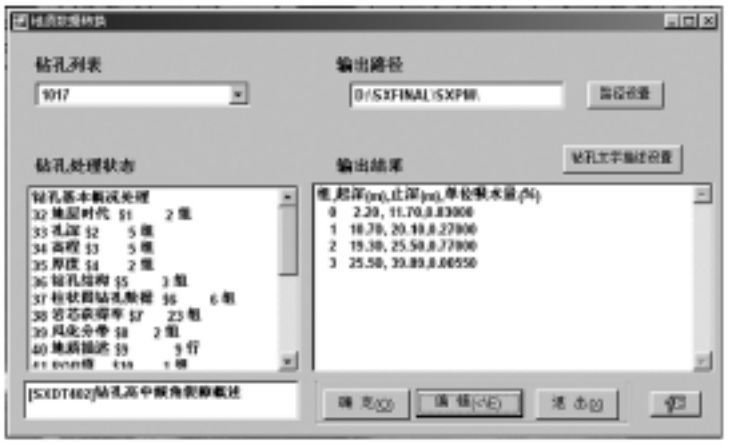


图 4-23　钻孔柱状图数据接口文件生成界面

(1) 在“钻孔列表”组合框中键入或选择一个钻孔号;

(2) 在“输出路径”文本框中键入或选择接口文件的路径名;

(3) 单击“确定”按钮，系统开始组织接口数据，处理情况在钻孔处理状态窗口显示；一旦处理完毕，该窗口提示“生成完毕!”，并在输出窗口显示处理结果;

(4) 单击“编辑”按钮进入编辑窗口，编辑生成的接口数据;

(5) 完成后单击“退出”按钮，选择接口文件，保存在“输出路径”文本框的路径名下。

一般地说，初次生成钻孔柱状图数据接口文件时，要先单击“钻孔文字描述设置”按钮，进入如图 4-24 所示界面，设置钻孔文字描述。钻孔文字描述设置内容包括：钻孔描述项目设置、钻孔描述数据表设置、文字描述方式设置。一次设置可多次使用。

2. 地质剖面图数据接口文件的生成

先从图 4-22 所示的绘图接口模块的总控菜单中选中“剖面图”，随即进入如图 4-25 所示的界面。然后按下列步骤生成勘探剖面图数据接口文件:



图 4-24　钻孔文字描述设置



图 4-25　工程勘察基岩剖面图数据接口文件设置界面

(1) 对于一条新剖面，先选中“基岩剖面设置”页（图 4-25），按上面要求键入剖面编号、投影范围、坐标系、剖面方位等剖面的总体信息;

(2) 选中“基岩剖面数据接口处理”页（图 4-26），依次键入剖面编号、接口数据文件名、投影范围、绘图接口路径，再单击“剖面数据接口处理”按钮，提取数据并组织接口文件;

(3) 生成数据接口文件后，单击“退出”，回到系统主界面。

整个数据接口文件的生成过程及结果，都自动显示于列表框中。

3. 平面等值线图数据接口文件的生成

先从图 4-22 所示的总控菜单中选中“平面图”，进入如图 4-27 所示的平面等值线图数据接口文件生成界面。然后按下列步骤生成平面等值线图数据接口文件:



图 4-26　工程勘察基岩剖面图数据接口处理



图 4-27　平面等值线图数据接口文件生成功能模块菜单项

(1) 单击下方工具条中“编辑”按钮，依表单顺序先键入平面图对应的建筑物编号、建筑物类型、工程位置、建筑物名称、边界坐标系、平面图坐标系、建筑物边界点坐标;

(2) 单击“自动搜寻钻孔”按钮，搜索落入边界范围内的钻孔。注意：搜寻钻孔的数目与边界方式有关，采用精确边界（即在右侧选择按钮中选定“边界内部”）搜寻到的孔确定在边界内；采用带缓冲边界（即在右侧选择按钮中选定“边界内外”）搜寻到的孔在边界的一定缓冲范围内。缓冲范围大小在下方的“边界点搜索范围”文本框中键入，单位为 m。搜寻结果显示在第二个列表框中，其中带“*”的孔号为缓冲区内的钻孔，可参与成图。

(3) 单击“浏览”按钮，屏幕弹出“字段选择表单”后，浏览将进入接口文件的数据项。可以根据按需要浏览全部字段或者部分字段值。

(4) 经过浏览证实数据无误后，在“选择输出路径”文本框中键入接口文件输出路径（可缺省），用鼠标左键单击“接口文件生成”按钮，即完成了选定区域完整数据接口文件的生成。用鼠标左键单击“随机数据接口”按钮，弹出字段选择表单，选定平面图中要用的数据项按“确定”，即完成了随机数据接口文件的生成。随机数据接口文件是指在边界坐标点范围内，有选择的挑出一些钻孔，并且从全部接口文件数据项中挑选部分数据项用于平面图中。

(5) 用鼠标右键单击“接口文件生成”按钮，可以调出已生成的选定区域数据接口文件进行编辑；用鼠标右键单击“随机数据接口”按钮，可调出已生成的随机数据接口文件进行编辑。除了自动搜索钻孔和生成平面图数据接口文件，界面表单还有如下基本功能：按输入的坐标计算面积，浏览接口文件中将出现的数据，移动、查找、添加、编辑和删除记录等。表单中的一条记录对应一幅平面图。

四、数据的打印输出

数据输出是资源数据库系统的重要功能，数据输出的形式及其灵活性是体现系统功能的重要方面。本系统数据输出形式是多样化的，例如图形从绘图仪输出，文字、图像从屏幕显示，也可以以报表形式从打印机输出。一般地说，打印输出是最为常用的输出形式。

打印输出模块的总控菜单如图 4-28 所示，功能分为基础报表输出和专题报表输出两种。

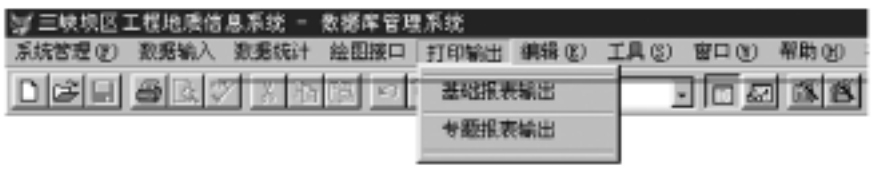


图 4-28　打印输出模块的总控菜单

基础报表指原始库中数据对应的报表，专题报表指原始库中数据经重组、统计、加工后生成的对应报表。基础报表与专题报表打印方式相似，在此仅以基础报表打印为例说明。

打印输出功能由一个通用表单执行。该通用表单汇集了欲打印数据的录入、数据查询、结果打印、打印项目操作等功能于一体。打印项目类型包括钻孔、边坡和平硐的信息。

在打印报表前，应当先查看报表对应库中有无数据。其方法是先单击“数据查询”键，利用条件构造器设置查询条件（例如查询某些钻孔的某些字段），然后在浏览窗口中浏览数据。条件构造器、浏览窗口的使用方法参见本章第一节。如果报表对应库中无数据，则单击“数据录入”调用录入表单录入数据。最后单击“结果打印”，利用条件构造器设置打印条件，再单击“打印预览”，满意后即可单击“打印”，令其打印输出。

五、系统工具模块的使用

系统工具模块是实现本系统强大功能的基础，其总控菜单如图 4-29 所示。



图 4-29　系统工具模块的总控菜单

系统工具模块的主要功能子模块有：数据字典、结构修改、函数管理、编辑器、数据库管理、数据表管理器等。该模块是一个集新建数据表、数据表结构修改、浏览表结构、数据表普通浏览、数据表条件浏览、数据库表和自由表间的无信息损失转换、数据库改名、字典操作等功能为一体的操作工具。其界面如图 4-30 所示。系统工具模块的应用较为复杂，这里不准备作详细介绍，必要时可以阅读有关的专著和系统操作说明书。



图 4-30　数据表浏览工具界面

思　考　题

1. 什么是数据库？什么是数据库管理系统？什么是数据库系统？什么是主题式数据库系统？什么是数据仓库？

2. 数据库与数据文件的本质区别是什么？什么叫数据的独立性？

3. 关系数据库的数据结构采用什么模型？这种模型的表达方式是什么样的？

4. 简述数据库系统在地矿资源勘查信息系统中的作用和地位。

5. 关键字有哪些种类？它们是根据什么原则来划分的？谈谈各种关键字在数据库关系模式表达中的作用。

6. 你能否用图解法简单地说明数据库的结构与工作原理？

7. 数据库应用设计中的基本问题是什么？简述数据文件的结构设计过程与方法？

8. 什么是数据字典？数据字典有哪些种类，各有什么用途？简述数据字典在计算机辅助软件工程（CASE）技术中的应用。

9. 代码在数据库中的作用是什么？编制术语代码应当遵循什么原则？常用的术语代码有哪些类型？试分析地质矿产分类术语的国家标准代码的结构特征。

10. 简述数据库系统管理模块、输入输出模块、修改模块和系统工具模块的功能特点。

11. 什么是专题数据提取？请用一种图形文件的生成过程来说明其操作方法要领。

12. 简述资源数据库子系统的资源管理模块的功能及操作要领。

第五章　地矿资源空间信息子系统设计与应用

空间信息子系统是地矿信息系统的重要组成部分，其主要任务是：①采集、存贮和管理地矿勘查所获取的空间数据；②实现各种空间指标的量算并绘制各种图件；③进行地矿及其控制因素的空间分析与综合评价；④开展与地矿形成、分布和开发有关的空间过程模拟仿真。

第一节　地矿资源空间信息子系统的结构与功能

地矿资源空间信息子系统的研发平台是具有强大空间信息管理和处理功能的 GIS，目前有些地矿资源空间信息子系统实际上就是 GIS 的直接应用。

一、地矿资源空间信息子系统的结构

地矿资源空间信息子系统由对各类空间数据进行采集、存储、管理、分析、处理、输出的计算机硬件、软件所组成，可综合表示为图 5-1。

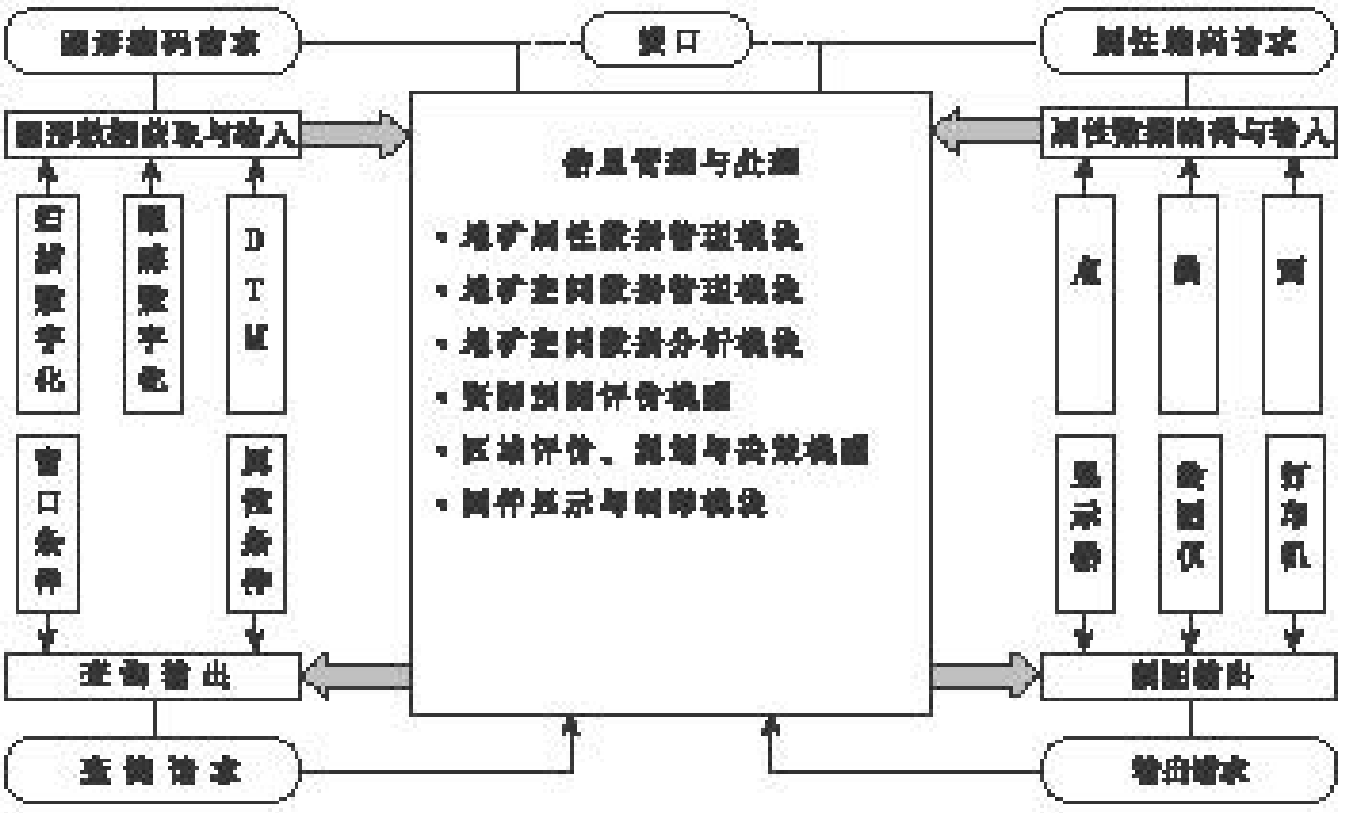


图 5-1　地矿资源空间信息子系统的组成

二、硬件组成

地矿勘查区空间信息子系统的特色在于管理和处理空间（图形）数据。由于所涉及

的信息类型特殊且数量巨大，除了要求有性能较强的微型计算机外，还需要一些专用的输入、输出设备和数据转存设备。各种输入、输出设备和数据转存设备通过总线系统与计算机联接起来，组成 GIS 的物理外壳（图 5-2)。

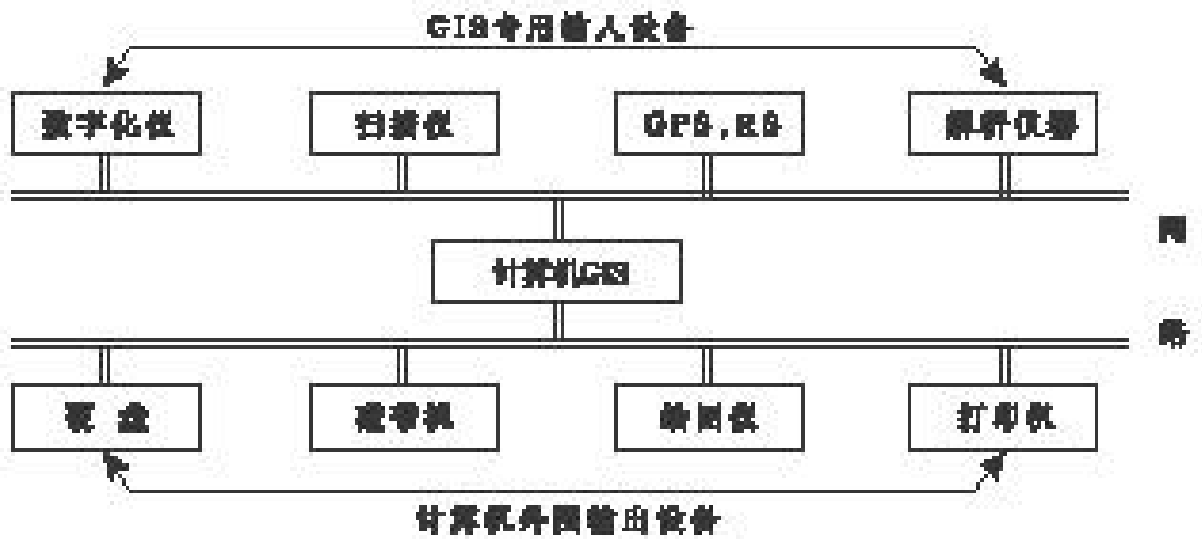


图 5-2　地矿资源空间信息子系统的硬件系统组成

1. 数据管理、加工、分析设备

由于地矿资源空间信息子系统所面临的是海量的空间数据和复杂的分析与编图工作，只有内存大、速度快的计算机方能胜任。工作站是大型 GIS 系统的主流机型，通常配有大屏幕高分辨率的图形终端和专用的图形加速卡，特别适宜于 GIS 的图形图像处理。工作站还可以通过网络将各种微机联成一个网络系统，达到地矿数据和地矿信息共享。但对于局部小区域或特定地矿勘查区而言，微型机的适用性更强一些。现行的 PENTEM—Ⅲ以上的 PC 机所具有内存储量（≥128 MB)、速度（≥600 MHz) 和外存（≥10 GB)，足以应付地矿勘查区的各种空间信息的存储管理和处理。虽然微机在内存储器与硬盘之间的传输速率方面较工作站稍慢，但在当前的数据量和处理方式下其影响并不明显，相反却比工作站更为方便和经济。

2. 外存储器

地矿信息系统面对着空间数据的海洋，例如，一幅 1∶5 万的地质图，有 50 ~100 MB 的数据量；一幅遥感图像一般有 10 ~30 MB 的数据量，而一幅 1∶5 万地形图或地理底图，通常也有 50 ~80 MB 的数据量。为了存贮、管理和处理这些数据，需要有大容量和高速度的外存储器。一般地说，为了达到快速响应，要在主机中配备大容量软盘和硬盘；而为了减轻主机硬盘的负担和保证数据的安全，还要配备大容量的光盘机或磁带机。

从目前情况看，磁带的存储量一般为 150 MB 以上，但易于发霉变质导致数据破坏，不便长期保存空间数据；硬盘的存储量已经达到 40 GB 以上，其体积很小，但也易于遭受破坏，只能存贮和处理日常的空间数据，同样不便长期保存空间数据；光盘的体积最小且每盘的存储量已达 600 M 以上，易于保存而不霉变，而且配备光盘机的成本低廉，基层勘查单位负担得起，因此可作为庞大的地矿资源空间数据长期存储的良好介质。

3. 输入设备

(1) 手扶跟踪数字化仪

手扶跟踪数字化仪是 GIS 最基本的输入设备。它将地矿图件转换成数字形式（矢量数据），主要针对地矿工作的前期资料和某些须预处理的图件。地图数字化的精度受数字化仪本身的分辨率、数字化方式、操作者的经验和技术等多种因素影响。

这种数字化仪由一块布满栅格状电子坐标网络的平板、一个图形坐标检测器（类似于鼠标器）和一根连接导线构成，用来将图形转换成矢量形式的数据。选择手扶跟踪数字化仪时要注意，它的实际分辨率一般比出厂时标定的分辨率要低。例如标定分辨率为 0.025 mm，实际分辨率可能在 0.07 mm 至 0.15 mm 之间。手扶跟踪数字化仪的图板幅面有 A0、A1、A3 等多种规格，在地矿信息系统中采用 A1 幅面，可以得到最佳的性能/价格比。一般地说，手扶跟踪数字化仪的分辨率愈高，图板的幅面愈大，则价格愈高。

（2）扫描仪

这是一种通过光电效应将图形直接转换成栅格化数据的设备。扫描仪既可用于输入由空间图形构成的线条图件，也可以用于输入由空间像素构成的栅格图像。

栅格数据经过识别与转换，可以转换成矢量形式的数据。购买扫描仪时，要根据实际需要和性能/价格比最优化准则，选择色型（单色还是彩色）、幅面及分辨率。扫描仪的分辨率是用像素大小（10 μm 至 100 μm 之间）或每英寸的网点数（dpi）来表示的。两者的转换关系是：400 dpi 大约相当于 64 μm× 64 μm 像素大小，每平方厘米约有 25000 个像素。1000 dpi 相当于 25 μm× 25 μm 像素大小，每平方厘米约有 155000 个像素。

（3）解析测图仪

为了适应空间信息系统发展的需求，一些技术开发单位和厂家对原有的模拟测图仪（例如 Technocart、Topocart-B）进行改造，增加了接口和相应的驱动软件，使之变成了新一代的解析测图仪。解析测图仪与计算机相连接，便可作为航空照片等图形数据的采集站。近年来，随着 GIS 的推广和应用，国内外又基于解析测图技术，相继开发出了新一代数字摄影测量工作站——机助测图系统和各种遥感图像处理系统，用于从航空相片和航天照片上采集所拍摄的影像数据。有了这些系统，这类地矿资源空间数据的采集和更新就更为方便、快捷和准确了。地矿资源空间信息子系统所管理和处理的这类空间信息，可以来自这些输入设备的直接联机采集与输入，也可以来自有关数字测图部门提供的软盘、磁带或光盘。

（4）其他仪器

如 GPS（全球定位系统）接收机，可通过卫星定位，将地面的移动目标如野外测量者、测量仪器车等的空间坐标实时地记录甚至输入计算机。

全站型速测仪，可以自动地以数字形式记录所测量的地矿资源空间数据。

4. 输出设备

由于数据分析处理结果主要以图形方式来表达，空间信息子系统的主要输出设备是各种绘图仪和图形显示器。从画板的样式和画笔的驱动方式看，绘图仪可分为平板式绘图仪和滚筒式绘图仪；从画笔的样式及其着墨方式看，有笔式绘图仪、静电绘图仪、喷墨绘图仪和激光绘图仪。但就其绘图原理而言，上述绘图仪均可归为矢量绘图仪和栅格绘图仪两大类。

此外，为了存储数据量大、不经常调用的二维、三维数据体和图形编绘程序，在数据输出时也经常用到磁带机和刻盘机。

1）矢量绘图仪

笔式绘图仪是矢量绘图仪的代表。这是一种传统的图形输出设备，其图板样式和驱动方式既有平板式也有滚筒式。矢量绘图仪的输出精度主要取决于控笔马达的步进量，根据地矿与地质勘查图件绘制的精度要求，步进量不应大于 0．054 mm。目前高精度绘图仪的分辨率可达 0.0125 mm，精度可达 0.125 mm，幅面宽达 1 m，长度可达几十米。矢量绘图仪绘制简单线条图形的精度较高而成本较低，但是绘制复杂图件费时较长且累积误差大；它的另一个缺点是色彩比较简单，也不便于进行区域（面）着色，因而不能输出图像。

2）栅格式绘图仪

针式点阵打印机、静电绘图仪、喷墨绘图仪和激光绘图仪都属于栅格式绘图仪。其中针式点阵打印机是最简单的栅格式绘图仪。它的精度较低但经济、快捷，可以用来输出草图。彩色激光打印机的精度和感官效果最好，质量也最高，但价格昂贵，适合于区域地质图或印刷版型的输出。彩色喷墨绘图仪在功能、制图质量、感官效果和价格上都介于二者之间，较适合于一般地矿与地质勘查成果图件或区域地质图草图的输出。

3）图形显示器

图形显示器在空间数据输入、检索分析和处理等阶段都要用到。地矿资源空间信息子系统应该配有高分辨率的图形显示器，其分辨率至少要达到 1024× 768 dpi 或 1152× 864 dpi。此外，为了加快图形图像的处理速度，最好配有图形加速卡。立体图形显示器在制作地矿勘查原始图件和分析图件方面起着重要作用。国外某些公司推出的数字摄影测量工作站都配有立体显示器，可以把摄像匹配的结果叠置在显示屏幕上，使用户一目了然。

三、软件系统

地矿资源空间信息子系统的软件包括一个 GIS 基础平台和一批应用程序。

1．GIS 基础平台

GIS 因其具有管理和处理空间信息的能力，被选做地矿资源空间信息子系统的应用软件开发和运行的基础平台。目前，国内外已经开发出来的 GIS 基础软件数以百计，在国内市场上流通的主要有 Arc/Info、Arcview、Mapinfo、Intergraph、Genamap 和 MapGis、Geostar 等。其中，Intergraph 和 Arc/Info 的功能最为强大，可用于开发大型地矿资源空间信息子系统，但主要的软件版本是工作站版的，微机版本的功能尚有待进一步开发，其价格也较为昂贵。Genamap 的情况与 Arc/Info 相似。从目前情况看，基层地矿与地质勘查单位都不具备使用这三种软件的条件。Mapinfo 和 Arcview 的微机版在国内较为流行，前者图元的拓扑关系描述和空间分析功能较差，但图形管理功能较强，适合于大量扫描入库的栅格化图件管理和处理；后者图形管理功能稍弱，但有一定的拓扑关系描述和空间分析功能，经二次开发，可以作为基层勘查单位野外空间信息采集系统和一般二维信息管理与处理系统。MapGis 和 Geostar 为国产软件，均具有较强的拓扑关系描述、空间分析和“三 S”（GIS、GPS、RS）集成功能。其中 MapGis 为微机版软件，通用性较强而专业化程度较低；Geostar 既有微机版也有工作站版，专业化程度较高而通用性较弱。二者均有网络版，可以作为基层勘查单位的二维信息管理与处理系统，但不太适合作为野外空间数据采集系统。

一般地说，GIS 的软件构成是有一定层次的（图 5-3）。最下面两层为操作系统和系统开发工具库。这两层与硬件有关，统称为系统软件。往上一层为 GIS 本身的基础软件库，

由此保证各种图形库、空间数据库、窗口系统等的正常运行，并且支持查询检索模块、功能处理模块、空间分析模块的作业以及对 GIS 的二次开发。最上面三层是 GIS 的功能应用软件，包括处理地矿空间信息的基本功能软件、特定地矿信息采集与处理软件以及针对地矿勘查与评价专业人员的用户界面。这三层代表了地矿资源空间信息子系统的能力和用途。其中的空间信息的基本功能软件通常由 GIS 开发者提供，而特定地矿信息处理软件以及专业用户界面属于 GIS 的二次开发的内容，由应用开发单位或用户根据实际需要自行开发。



图 5-3　GIS 的软件层次

2. 功能模块的组成

按照地矿资源空间数据的采集、加工、管理、分析和表达的需要，地矿资源空间信息子系统主要由 5 个模块组成：空间数据输入模块、空间图形编辑模块、空间数据库管理模块、空间信息查询与分析模块以及成果输出模块。它们之间的联系如图 5-4 所示。

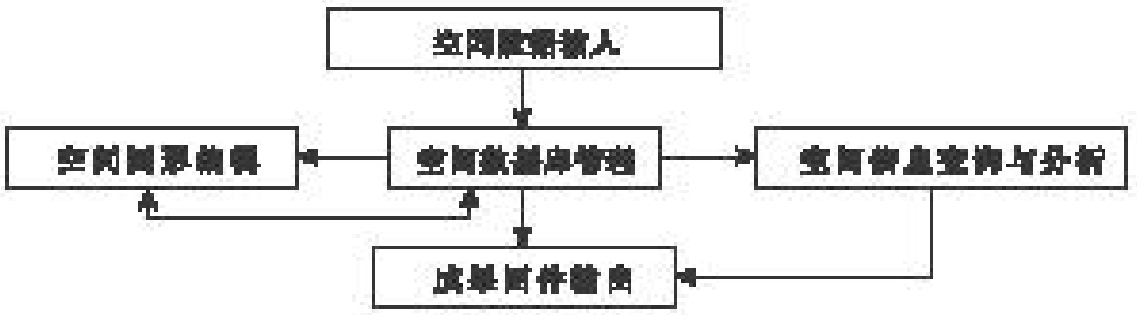


图 5-4　地矿资源空间信息子系统各功能模块间的联系

(1) 空间数据输入与转换模块

空间数据输入与转换模块，是一组相对独立的功能软件模块。它所面对的空间数据包括：原有的各种地矿勘查成果图件、航空相片、遥感数据和图像、地形地物测量数据、野外现象照片和素描、文本资料等多种类型。许多计算机控制的外部设备，如键盘、手扶跟踪数字化仪、扫描仪、磁带机、软盘和软盘驱动器等都可以作为输入设备。针对不同的输入设备和不同类型的地矿资源空间数据源，应当配备不同的软件，以便保证各种来源的空间数据都能转换成 GIS 可以接收的数字形式，并进入地矿资源空间数据库（图 5-5）。

(2) 图形与文本编辑模块

图形与文本编辑模块主要用于对输入数据的质量检查与纠正。一般地说，经扫描仪和数字化仪输入的空间（图形）数据，还需要进行图形编辑。图形编辑模块除了提供基本的空间、属性数据交互编辑、错误检查、修饰成图、图幅拼接功能外，还提供对图形进行修饰并设计线形、颜色和符号注记功能，提供建立拓扑关系、组合复杂地物、输入某些属性数据的功能，还提供文件数据转换、坐标转换、比例尺放缩、投影形式转换等功能，有的甚至可以提供多达 38 种地图投影方法（如 Arc / Info）。总之，地矿资源空间信息子系统

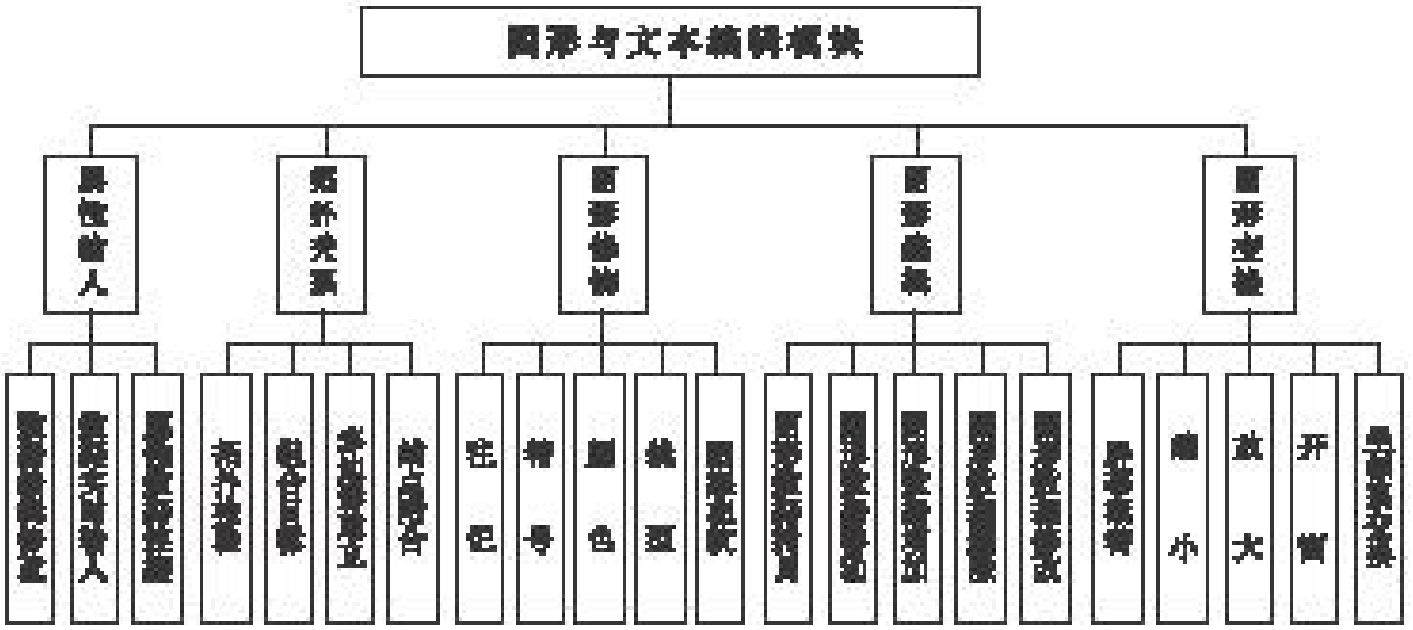


图 5-5　数据输入与转换模块示意图

将提供一套完善的、自动化的空间数据编辑功能（图 5-6)。

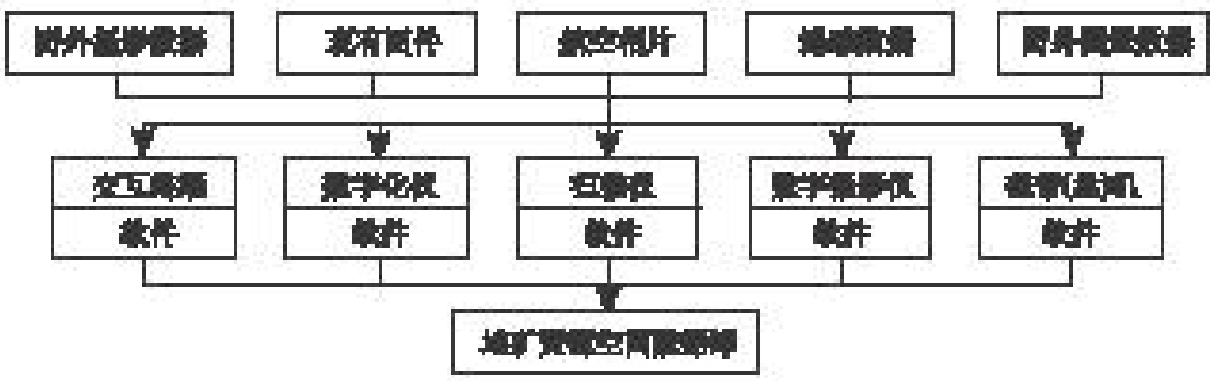


图 5-6　图形与文本编辑功能组成模块

对于描述地矿空间实体的属性数据，虽然也可以在前述的数据输入模块中输入，但也可以在图形和文本编辑模块中，直接参照空间图形输入，即在图形输入的同时输入相应的属性数据。这样做有利于实现图形数据与属性数据的连接。至于哪些属性数据在前述的数据输入模块中输入，哪些属性数据在图形编辑模块中输入，必须明确规定，避免重复。

(3) 空间数据库存储与管理模块

地矿资源空间信息子系统的数据存储与管理，涉及地矿资源实体的位置、空间关系以及地矿资源属性数据的结构和组织。地矿资源属性数据的存储、管理和处理通常利用地矿资源属性数据库子系统，而地矿资源空间（图形）数据的存储、管理和处理则使用专门设计的地矿资源空间数据库管理系统，少量与空间数据匹配输入的属性数据也采用空间数据库管理系统进行存储、管理和处理。

由于目前的空间数据库管理系统管理属性数据的功能较低，人们正在研究一种面向对象的空间数据模型，它既可以表示地矿属性数据又可有效地表示地矿空间图形数据。地矿资源空间信息子系统的空间数据库组成、存储与管理功能以及数据流向，如图 5-7 所示。

(4) 空间查询与空间分析模块

空间查询是指按一定的条件从 GIS 目标之间的属性、空间位置、范围和空间关系中派生出信息，从而形成一个新的地矿数据子集，并以此实现有关地矿空间信息的查询与分析；空间分析则是通过对各类地矿实体的空间特征进行一系列检索、运算、分析、综合，来研究地质体和地矿资源的时空演化规律。该模块的结构和内部关系如图 5-8 所示。

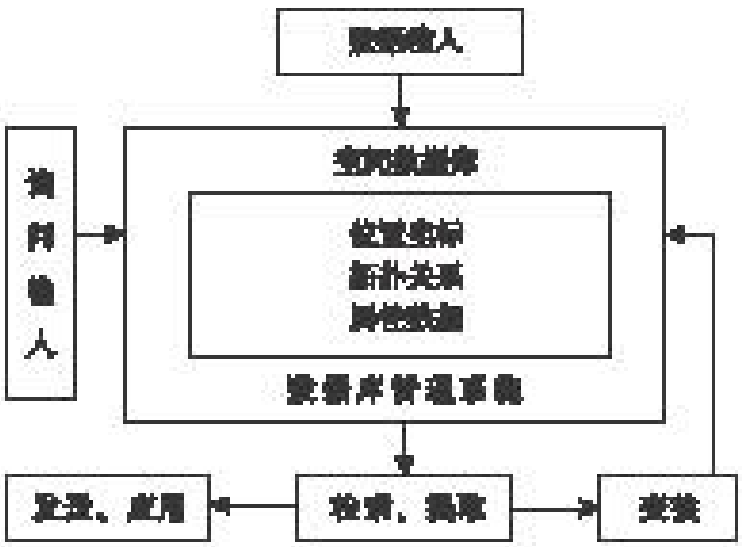


图 5-7　空间数据库的组成与功能

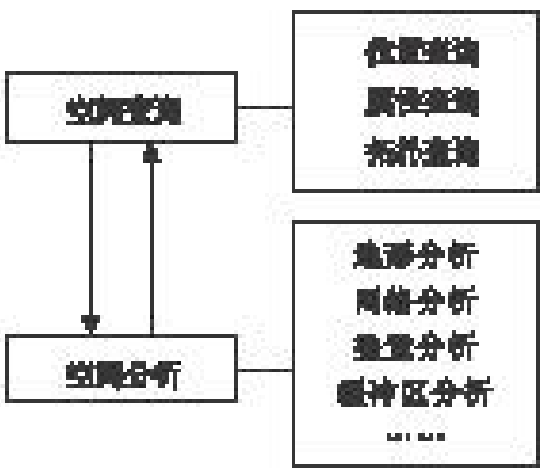


图 5-8　空间查询与分析模块

空间分析的基础是：地矿空间实体都可用点、线、面等三类图素的空间特征来抽象表示，其间的拓扑关系反映实体间的相互关系；三类图素可产生多种组合关系，表达相互间的邻接性、闭合性、包含性、一致性等。GIS 拥有自动建立上述拓扑关系的整套算法，一方面可以验证拓扑关系的正确性，修正地矿资源空间数据录入及存贮的错误；另一方面提供了对空间信息进行操作和分析运算的各种功能。但是，由于不同地矿类型的空间分析内容、方式和过程都不相同，为了提高空间分析的质量和效率，需要建立不同的地矿空间分析模型。地矿空间分析模型是对具体地矿空间对象及其演化过程的抽象，它们通过一系列典型的运算与查询命令来表达，可以用来导引并帮助用户完成相关地矿勘查和资源评价业务的各项空间分析任务。

空间查询与空间分析是空间信息系统最重要的功能。研究人员不仅可以就某个地矿实体本身的信息进行双向查询，即根据地矿空间图形查询相应的属性信息，或按地矿属性的特征查找相应的地矿空间实体，而且可以从整体上把握地矿体的时空结构，获取很多新信息和新知识，实现地矿、环境调查和国民经济建设中的综合评价、预测、规划和决策。因此，该项功能大大增加了各种图形信息和属性信息的利用深度和广度。

(5) 数据输出模块

地矿资源空间数据的处理与分析结果，通常要以用户最容易理解的图件形式来表达。在地矿资源空间信息子系统中，所有的地矿图形图像信息都以数字形式存储，而且都以有效的数据结构与各种地矿专题信息有机地联系在一起，因而在最终成果的表达与输出方面有显著的优越性。地矿资源空间信息子系统的数据输出模块通常具有如下特点:

1) 输出内容的可选性。该模块不仅可为特定地矿研究人员输出全要素地矿图件，也可以根据用户的需要，分层输出各种地矿专题图件，如地矿分布图、地矿特征参数图、地矿区划图、地质图、勘探剖面图、地理底图、地形等高线图等，还可以通过空间分析得到

一些特殊的地矿分析用图，如坡度图、坡向图、地矿预测图、地矿评价图，等等。此外，该输出模块还能根据用户的特殊需要，随时调用和叠置专题信息，并迅速形成各种综合图件，如地球化学综合异常图和地球物理综合异常图，等等。

2）图式图例的可变性。在成果输出之前，可以利用图形编辑模块，设计出适合于特定应用专题的图形符号、填充花纹和复杂线型，并且制作出相关的图例、责任表、拼图表、指北针、比例尺等各种图饰；可以自由选择各种字型字体进行标注，并实现图形与描述文字混合排版；甚至可以处理地矿图形的复杂压盖、避让等特殊图形问题。由于所有的图形都已数字化，在输出前可以反复处理，随时修改，满足不同用户对图件形式的要求。

3）产品类型的多样性。该模块能够以图形、图像、图表、表格、数字和文字等多种形式向用户提供中间和最后结果。例如，能够输出表示地矿空间分布特征和相互关系的专题地质图和专题影像地图；能够输出表达统计信息、决策信息的饼状图、直方图和空间分布图；能够输出支持地矿属性数据查询、统计和运算的统计表格。信息产品可以装载于纸介质（图纸）上，也可以装载于磁介质（磁盘或磁带）和光介质（光盘）上。

4）图形设备的普适性。地矿资源空间信息子系统的图形输出模块与设备独立。它能生成各种标准格式的图形图像数据文件，能够在不同型号的在荧光屏、磁盘驱动器、光盘机、打印机、笔式绘图仪、静电绘图仪、喷墨绘图仪和激光照排机等设备上输出。

四、国内外主要 GIS 平台简介

目前，GIS 基础软件品牌繁多，由于这些软件产生背景不同，数据存储方式不同，处理手段和功能也不同，导致它们在地矿勘查应用领域的适应性也不同。

下面将对几种国内常用的 GIS 作概要的介绍。

1．ARC/INFO

这是一种功能较强、目前在我国应用较多的国外大型 GIS 软件产品。

开发公司：美国环境系统研究所（ESRI）公司

推出日期：1982 年

硬件环境：小型机、工作站、PC 机

操作系统：UNIX、VMS、DOS

主模块：包括 ARC、INFO、ARCEDIT、ARCPLOT、ADS、MAP LIBRARIAN 等功能模块；内中包含了处理地理信息系统的各种高级功能，如：数据输入和编辑、数据转换和集成、基本 GIS 功能——地图投影及投影转换和多边形叠加分析等、完整的数据查询和显示工具、地理数据库、生成图形用户界面的工具。

附加模块：NETWORK——线状网络分析模型、TIN——地表模型和地形分析模型等、COGO——交互式坐标几何输入、管理大比例尺成图、ARCSCAN——扫描图预处理及矢量化模块、GRID——栅格数据模型分析和显示模块、ARCSTORM——基于现代计算机技术客户/服务器结构的空间数据管理软件模块（它能使 GIS 用户管理大型地理数据，能对地理数据进行特征管理，支持多用户存取连续空间数据库，还能对分布数据库进行有效组织与管理）。

数据交换格式：DLG、DXF、HPGL、SIF、TIGER、DEM、DIME、ERDAS、LANDSAT、SPOT

网络：支持 NFS

数据类型：矢量、栅格

数据管理特点：按图幅存放，但能合并图幅；采用 QUADTREE 方法空间索引；图形信息与属性信息的关系较为紧密，连接码由系统指定。

Internet 网址为：http：//www. esri. com

2. ARCVIEW

ARCVIEW，起初是 ARC/INFO 的一个空间信息浏览模块，随着 PC 机硬软件的发展，ARCVIEW 逐步成为一个独立的易学、易用的 GIS 工具和界面友好的桌面地理信息系统。它能让用户对组合数据进行查询与显示；直接对 ARC/INFO 数据进行操作。ARCVIEW 提供完整的桌面制图、属性表管理、统计分析、连接多媒体等功能，支持多类型数据和多种数据库，有完善的基于面向对象的二次开发功能。

开发公司：美国环境系统研究所（ESRI）公司

推出日期：1982 年

硬件环境：PC 机、工作站、小型机

操作系统：Windows X、DOS、UNIX、VMS

主模块：ARCVIEW——面向目标的产品、AVENUE——面向目标的二次开发语言

附加模块：各种空间分析模块、网络分析模块等

数据交换格式：SHP、DXF、MIF、DLG、TIF、ERDAS、LANDSAT、SPOT 等

网络：NFS

数据类型：矢量

ARCVIEW 特征：具有简洁、直观的用户界面；先进的显示和观察工具；支持复杂的空间和属性双向查询及 ODBC 支持；能与其他桌面地理信息系统（如 MapInfo）和不同类型数据进行热联结“HOT LINKS”；商业制图功能——较强的图形编辑和裁剪功能，能制作统计图形、柱状图和饼状图；地图结构设计与符号化处理；地理网络地址匹配与分析；完善的面向对象的二次开发工具；适用于大量的数据查询和显示及 SQL 结构化查询。

其 Internet 网址为：http：//www. esri. com

3. MapInfo

开发公司：美国 MapInfo 公司

推出日期：1992 年

硬件环境：PC 机

操作系统：DOS、MS Windows X

主模块：MapInfo——面向地图的目标产品、MapBasis——二次开发语言

附加模块：各种空间分析模块、网络分析模块等

数据交换格式：WOR、DBF、DXF、MIF、TIF、TXT、TAB、LANDSAT、SPOT 等

数据类型：矢量、栅格

MapInfo 特征：简洁、直观的用户界面和方便的显示和观察工具；支持 OLE 拖放和多种数据格式转换；支持空间和属性双向查询和 ODBC 支持；商业制图功能——能制作统计图形、柱状图和饼状图等；地图结构设计与符号化处理及标注；完善的自动跟踪数字化和灵活的副本地图窗口创建；地理网络地址匹配与分析；完善的面向对象的二次开发工具（MapBasic）；较强的图层分层管理和统计报表编制。

Internet 网址为: http: //www. mapinfo. com

4. MAPGIS

MAPGIS 是由中国地质大学(武汉)自行开发的有自主知识产权的全汉化国产优秀桌面地理信息系统软件。

开发公司: 中国地质大学中地公司

推出日期: 1995 年

硬件环境: PC 机

操作系统: MS Windows X

主模块: 输入子系统——MAPGIS 提供的数字化仪、扫描矢量化输入、GPS 输入和其他数据源的直接转换模块; 图形整饰子系统——编辑修改矢量结构的点、线、面的空间位置及其图形属性、增加或删除点线面等, 并适时自动校正拓扑关系; 输出子系统——将编排好的图形显示在屏幕上或在指定的设备上硬拷贝输出等; 空间分析子系统——对空间与非空间数据进行分析和查询, 包括矢量空间分析、图像分析、数字高程模型等; 库管理子系统——通过图形库管理模块和专业属性库管理模块来实现空间和属性数据的管理; 实用服务子系统——用来实现对拓扑关系等方面错误的检查和不同图形系统间的数据格式交换, 以及进行系统库的编辑等。

数据交换格式: WT、WP、DXF、TIF、EOO 等

数据类型: 矢量

5. GEOSTAR

GeoStar 是由原武汉测绘科技大学自行开发的又一个有自主知识产权的全汉化国产优秀地理信息系统软件。这是一种较为先进的组建式 GIS 软件平台。

开发公司: 原武汉测绘科技大学吉奥公司

推出日期: 1995 年

硬件环境: 工作站、PC 机

操作系统: UNIX、MS Windows X

主模块: 数据采集——提供手扶跟踪数字化、图纸扫描屏幕数字化、直接从测量仪器获取数据及外部数据文件等多种手段; 空间数据管理——可以使用文件系统或关系数据库来存储和管理空间几何数据; 图形编辑——可以自由地增加、删改、旋转、分割、捕捉和整饰点、线、面, 能方便地进行属性操作并自动构建和解除拓扑线、面, 能自动进行图幅接边处理; 空间查询——可以根据属性查图形、根据图形查属性, 提供空间关系查询、几何量算和统计分析功能, 可以进行多种查询选择的任意组合, 可以一次点选多个目标并浏览任意对象的属性, 查询结果可作为进一步分析报表的选择集; 空间分析——运用空间算子思想, 采用面向对象的方法与技术, 提供单目算子和双目算子, 用户可以通过这些算子的组合或二次开发, 完成各种空间分析功能。二次开发——提供数据库读/写、数据查询、数据更新、空间分析、数据转换、影像处理、界面设计等一系列应用程序接口(API)函数; 此外, 还有普通地图制图和符号设计模块、专题制图模块、数据转换模块、地图投影和坐标变换模块、元数据管理模块、报表设计与属性数据处理模块。

数据交换格式: WT、WP、DXF、TIF、EOO 等

数据类型: 矢量、栅格

第二节　地矿资源空间数据库的应用设计

空间数据库设计是地矿资源空间信息子系统应用设计的核心任务。为了方便地存储、管理和处理各种空间数据和属性数据，GIS 内部同时设置有关系数据库（称为内置关系数据库）和空间数据库。其中，内置的关系数据库主要存放一些与空间位置密切相关的简单属性数据，例如点、线、面等空间实体的名称、性质等。这种 GIS 内置关系数据库的原理和方法，与第四章所介绍的关系数据库基本上是一致的。

本章将着重介绍有关内置空间数据库的设计方法和设计步骤。

一、空间数据库设计的概念

空间数据库是合理存放于计算机存储设备上的相互关联的空间数据集合。与属性关系数据库相同，空间数据的存储也要求满足以下 4 个基本原则：①最小冗余度原则即数据尽可能不重复；②资源共享性原则即以最优的方式服务于一个或多个应用程序；③数据独立性原则数据的存放尽可能地独立于使用它的应用程序；④管理统一性原则：能够用一个软件统一管理这些数据，例如对数据的维护、更新、增删和检索等一系列操作。

空间数据库的设计需要考虑数据结构模式，因为空间数据有矢量和栅格之分，各种数据又同时具有空间和属性的特征，有的还带有时间信息特征。信息的特征不同，所采用的表达结构模式就不同。而且，在不同地点、不同单位，各类地矿资源空间数据库的开发可能基于不同的 GIS 平台，所存放数据的格式也各不相同，但在地矿调查与勘查工作中，往往需要统一存储、管理和处理这些数据。如何做到兼容各种各样的数据类型和格式，是一个空间数据库设计成功与否的关键因素之一。空间数据库的设计还应该兼顾应用目的。

空间数据库的系统设计原则上也可以分为总体设计（或称概念化设计）和详细设计两个亚阶段，下面将分别叙述。

二、空间数据库的总体设计

空间数据库的总体设计也称为概念化设计，其内容主要包括数据库的宏观地理定义和数据模型设计两大部分。此外，还要考虑整个数据库的冗余度、一致性和完整性等问题。

1. 数据库的宏观地理定义

空间数据库的宏观地理定义主要是指对数据库比例尺、地图投影和坐标系统的定义。

(1) 比例尺定义

比例尺是指地图、地质图和各种资源勘查图件与地表上长度的对应关系，它是图件或数据精度和详细程度的标志。数据库的比例尺通常取决于用户对数据精度的要求。尽管被输入到数据库中的地理与地矿资源空间数据，可以用任何比例尺进行显示，但由于产生这些数字化数据的原始图件是有比例尺的，而原始图件比例尺的大小决定着一个数据库的精度。因此，在进行空间数据库总体设计时，必须定义比例尺。一般地说，小比例尺的图件通常要比大比例尺图件的精度低，因为幅面太小的小比例尺图件所覆盖的面积要比大比例尺图件大，而大比例尺图件所包含的信息则比小比例尺的图件详细。例如，一幅 1∶ 25 万的地质图所覆盖的面积可以用 25 幅同等大小的 1∶ 5 万的地质图来覆盖（图 5-9）。

比例尺的大小也直接影响图件地理特征的几何分辨率。大比例尺图件上的地理特征在小比例尺图件上可能会被简化、被综合，甚至被省略。这在地图学上称为地图综合。地

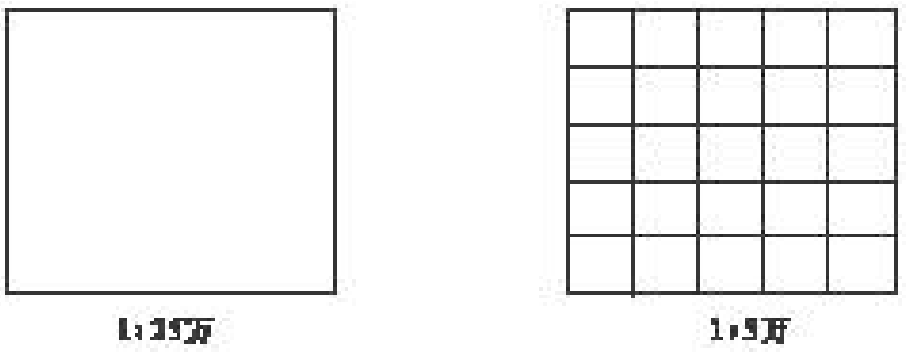


图 5-9　不同图件比例尺的关系

图综合有一整套原则，各个国家的原则不尽相同，但无论哪种原则，究其核心，都取决于主观判断，因为地图不仅是一种技术，也是一种艺术；不仅要求技术上的合情合理，又要求艺术上的美观易读（Robinson，1985）。

当然，一个空间数据库可以有多种比例尺，根据数据的种类及用户要求，不同的图件可以采用不同的比例尺。用户在对多种图件数据进行转换、叠加和空间分析的时候，应该了解原始数据的比例尺，因为分析结果的精度总是与最小比例尺数据的精度一致的。此外，不同比例尺图件上的数据有不同的分辨率，在空间数据库进行管理和处理往往会造成某些技术上的困难。例如，在 1∶ 1 万地形图中等高线的间距是 10 m，而 1∶ 5 万地形图中等高线的间距是 50 m。当对这两种数字化地图进行图幅拼接时，由于不同比例尺的精度不同，很难找到对应的等高线；即便是找到了，相邻等高线通常也不会自然地连接起来，需要凭经验用手工方式进行编辑，才有可能使之实现连接。

多种不同比例尺原始图件数据的同时使用常常引起的另一种困难，即冗余多边形的问题。同一个地理区域中的某些内容在不同比例尺图件中可能均有表达，但因为精度和分辨率的不同，使得该区域内地质地理特征在两种比例尺图中不相吻合。例如 1∶ 20 万化探异常图和 1∶ 5 万地质图都有地层界线，当我们对这两种图形数据做叠加分析时，因为地层界线在两种比例尺图上不相吻合，会产生出无数细小的多边形。这种小多边形的存在将大大增加数据库的数据量，影响各种处理的速度，需要加以删除。

从上面分析的结果可以知道，图件的幅面数总是随着比例尺的减小而成指数增长关系的。这种关系也适用于图形数字化输入和图件生产的费用预测，所以对比例尺的选择，是空间数据库系统建设造价和运行成本的最有影响因素，应当引起重视。

(2) 地图投影和地理坐标系统

地图投影是利用某种坐标系统来测量和表达地球上某一位置的方法。空间数据通常可以用两种系统来表达：一个是地理坐标系统，另一个是笛卡儿平面坐标系统。球面上的各个位置，都可以通过一定的数学模型转换到由这两种坐标系统构成的二维平面上。地图投影坐标系统的选择会影响整个空间数据库的数据形态，从而制约空间数据库使用的有效性。因此，空间数据库中地图投影方式的选择是一个关键环节。

地图投影所引起的变形主要表现在形状、面积、距离和方向等 4 个方面。一种投影方式可能对其中某一方面没有任何影响，但是在另外几个方面则可能造成不小的变形。在选择地图投影方式时，应当首先了解用户对空间数据库将如何使用，再决定应该在哪个方面保持无变形或少变形，然后决定采用哪一种地图投影方式。

表 5-1 列出了常用的地图投影方式及其特点。这些投影方式的类型，是根据几何特性的差异、适应地区和通用范围划分的。不同的地图投影方式尽管各有特色，但就其整体而言，以下几条是共同的，在选择投影方式时应当注意。

表 5-1　常用地图投影一览表

投影方式	类型	等形	等面积	等距离	等方位角	等罗盘方向
全球	球面	全部	全部	全部	全部	
墨卡托	柱状	全部			部分	全部
横轴墨卡托	柱状	全部				
倾斜式墨卡托	柱状	全部				
星际倾斜式墨卡托	柱状	全部				
米尔特柱状	柱状					
鲁滨逊	伪柱状					
正弦等面积	伪柱状		全部	部分		
正射	正射方位				部分	
球极平面	正射方位	全部			部分	
兰勃特方位等面积	正射方位	全部			部分	
阿尔勃斯等面积锥形	锥形			部分	部分	
兰勃特等形圆锥	锥形	全部	全部		部分	
等距离圆锥	锥形		全部			
多元圆锥	锥形					
双极倾斜等形圆锥	锥形	全部				

①任何地图投影均有变形；②等形投影保持小区域内形状不变，但不能保证大范围内的形状不变；③等面积投影仅保持同比例的情况下面积不变；④等距离投影保持某种距离上不变，但不能保持一幅地图上任何两点间的距离不变；⑤等方向投影保持地图上的方位角与真实方位角相同。其中，墨卡托投影（UTM）是一种国际通用的地图投影方式，适用于 ±80 °纬度范围内的地区。它以每 6 °经度和 8 °纬度为一个分割带。中国、前苏联和美国均使用 UTM 作为地图投影，从 180 °起算向东分 1 到 60 个分割带，南纬 80 °向北编号从 E 到 X，北半球以赤道起算为 0。南半球从 10000000 m 起算，即所有该区的值均要加上 10000000 m。

面积较大的地区常常跨越几个投影带，在设计空间数据库时可以有下列几种选择：①将所有数据放入一个带。例如手头的数据覆盖两个带，那么可以选择其中之一作为数据库的带；②将各带分别存放，这样做的好处是变形较小，但该空间数据库在使用时不够方便，没有好的整体感；③选择适当的地图投影方式，即可以选择一个能够将整个地区放在一个带中的地图投影方式，但该种地图投影方式应当是常用的；④采用地理坐标系统，此

种方法适合于大区域的空间数据库建设。另外，假若该数据库将服务于多种用途时，也可以直接采用地理经纬度进行存储。只是在各具体应用时，需要将其转换成所需要的平面坐标系。

综上所述，空间数据库的地图投影和坐标系统选择的一般原则是：①在经常需要投影变换而且覆盖面积较大的情况下，应该使用地理坐标系统；②笛卡儿坐标系统对于面积小和坐标系固定的区域最为适合；③根据研究区的形状来选择变形最小的投影；④如果有地区标准的话，应该使用地区标准；⑤如果研究区的面积很重要时，可以考虑使用一种等面积投影方式进行面积计算，但数据存储仍可使用另外一种投影方式。

2．空间数据的结构模式

空间数据库的数据结构分为 3 种，即矢量数据结构、栅格数据结构和矢量—栅格混合数据结构。下面将分别加以简单介绍。

(1) 栅格数据结构及其表达

栅格数据以规则阵列来表示地矿实体的分布，其中每个数据表示地矿实体的属性。栅格数据更适合描述面状地质体的分布，用于地质体位置、面积等的计算，但其数据量庞大，一般均需压缩存储。有许多地矿资源空间数据，如遥感与数字摄影测量数据，扫描图像以及网格数字地面模型都是栅格形式，不必要都经矢量化后才进行空间分析，可以直接采用栅格数据结构进行分析处理。

$$\begin{bmatrix} A & A & A & A \\ A & B & B & B \\ A & A & B & B \\ A & A & A & B \end{bmatrix}$$

图 5-10　二维矩

1) 栅格矩阵。表示地矿空间信息的栅格数据实际上是地矿资源空间数据二维的离散量化值，每一层的像元值组成二维信息数组的像元阵列，行、列表示它的位置。故此，栅格数据可以表示为一个二维矩阵。例如，图 5-10 的影像，当扫描输入时，一般从左上角开始，逐行逐列进行。矩阵的每个元素，代表了 1 个像元的值，像元值可以是 1 个字节，2 个字节或 4 个字节。矩阵元素的行、列号即是像元的位置，在计算机中，矩阵数据是顺序存放的。例如，1 个 4× 4 的影像，在机器中存储的顺序为：

A　A　A　A　A　B　B　B　A　A　B　B　A　A　A　B

通常一个文件存储一层的信息，也可以在一个文件中存储多层信息。每个像元可能是占一个字节的整型数。如果地矿图像比较大，这种逐点存储的方式，所占的空间是十分巨大的。所以在实际存储时，多采用压缩存储的方式。最简单的是行程编码方式。另外，还有四叉树编码方式、二维行程编码方式等压缩倍数更多的方法。但数据压缩所获得的空间节省与数据处理的时间花费往往是成正比例，即所节省的空间越多，数据压缩与解压所需的时间往往也多，因此所采取的方法要根据实际的情况而定。

2) 行程编码。地矿资源空间数据一般具有相关性，可采用某种编码方法进行压缩和合并。行程编码即为一种压缩方法，其要领是按行扫描，将相邻等值的像元合并，并记录行程长度及它的值。对于图 5-10 影像矩阵，根据空间数据的相关性，可以将等值的相邻像元合并，得到 6 组行程编码。如果按行合并，则有：

4A　1A　3B　2A　2B　3A　1B

如果不按行，连续合并，则有：

5A　3B　2A　2B　3A　1B

显然，行程编码节省空间。区域越大，数据相关性越强，节省的空间越多。

3）四叉树编码。根据栅格数据的二维空间分布特点，将地矿实体所在空间区域按四个象限进行递归分割（$2^n \times 2^n$，且$n \geq 1$），直到子象限的值单调为止，最后得到一棵四分叉的倒树（图 5-11）。

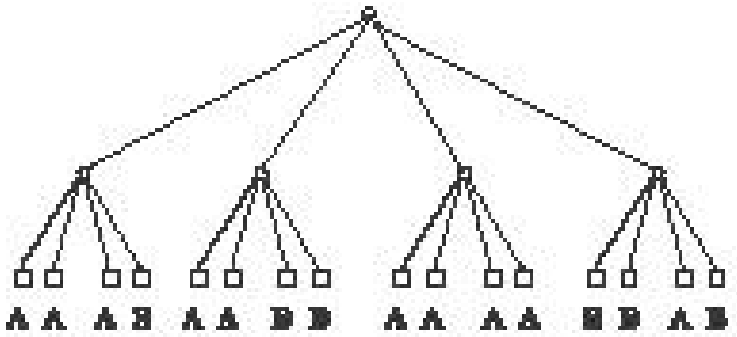


图 5-11　四叉树结构

四叉树有两种，一种是指针四叉树，另一种是线性四叉树。前者在叶节点与根节点之间设置指针，由于指针占用大的存储空间，达不到压缩数据的目的；后者只记录叶节点，用地址码表示叶节点的位置，能达到压缩数据的目的。

为了得到线性四叉树的十进制地址码，其方法是把二维栅格的行列号转成二进制数，然后交叉放入十进制地址码变量中。例如：I = 5，二进制数为 0101；J = 7，二进制数为 0111；行列交叉的二进制数为 01010111。

该地址码转为十进制地址码为 55 如表 5-2 所示。

表 5-2　线性四叉树的地址码

十进制地址码 列号 / 行号	J	0	1	2	3	4	5	6	7
I									
0		0	1	4	5	16	17	20	21
1		2	3	6	7	18	19	22	23
2		8	9	12	13	24	25	28	29
3		10	11	14	15	26	27	30	31
4		32	33	36	37	48	49	52	53
5		34	35	38	39	50	51	54	55
6		40	41	44	45	56	57	60	61
7		42	43	46	47	58	59	62	63

这样就将二维数组转成一维数组的十进制地址码。对一维数组扫描，依次检查每四个相邻格网的属性值，如若相同，则合并成一个节点，并记录地址和格网值，否则不合并。第一轮检查与合并后，再依次检查每四个大块的值，如若相同，就合并成更大的节点，否则不合并，依次递归循环，直到没有合并的叶节点为止。

（2）矢量数据结构及其表达

矢量数据结构是一种最常见的图形数据结构，即通过记录坐标的方式，精确地表示点、线、面等空间实体的位置。其坐标空间假定为连续空间，不必像栅格数据结构那样进行量化处理。因此矢量数据能精确地定义地矿实体位置、长度和大小。当然，其精确度受

如下因素影响：①表示坐标的计算机字长；②数字化仪分辨率；③输入时实体边界曲线上的选点密度；④人工输图时的定位操作误差；⑤输出设备的分辨率。在空间数据库中，矢量数据是按隐式关系以最小空间存储的。

1）矢量数据概念。地矿空间目标可划分为点、线、面、体四种基本几何实体类型。在地图中的各种要素都可分别用这四种几何实体来表达。这些几何实体在计算机中以数字形式存储和处理。

点　图上的居民点、矿点、水井、高程控制水准点以及文字标注等实体，都可看作点。其空间位置通常采用一对坐标（X，Y）来表示。这对坐标（X，Y）既可以是地表的经纬度坐标，也可以是数字化输入过程中的平面坐标系提供的相对位置。其表达式即为（X，Y）。

线　图上表示河流、道路、断层、地质体界线、行政界线等几何实体均可看作线。线是一组有序的点的集合，具有起始点、终止点和明确的方向性。因此，这类地矿实体的位置可以采用一组有序的（X，Y）坐标来表示。这些坐标是在曲线上按一定的间隔获取的，称为矢量数据。其表达式为：（X1，Y1），（X2，Y2），……（Xn，Yn）。

面　是指有封闭边界和面积的实体，是由一组有序线段包围而成的区域。例如湖泊、矿区、宗地、行政区划、地质单元等图斑。地矿空间实体的边界位置可采用一组有序的线段来表示。这组有序的线段首尾位置必须重合。其表达式为（X1，Y1），（X2，Y2），……（Xn，Yn）。最末一点的坐标与第一点的坐标是相等的，即 Xn = X1，Yn = Y1。

为了满足空间查询和空间分析的需要，在存储和管理某图形元素的空间坐标的同时，还需要记录各图形元素之间的相互关系以及与之相应的图形参数——属性数据。对于线来说，它的图形参数有线型、线宽和线色等；对于面来说，它的参数为填充图案类型、填充颜色和填充方式等。对于点来说，如果是文字，它的参数为字高、字宽、字间距、字型、字体、颜色、旋转角；如果是图符，它的参数为符号代码、符号高、符号宽、颜色、旋转角等。计算机根据这些图形元素的坐标以及它们的图形参数，可以精美地重现地矿实体的原形。

在空间数据库中，同一种类型的地矿实体，一般具有相同的属性项目，属性数据通常以表格的方式存储。如果只是点状的地物，可以将空间目标的几何数据与属性数据结合在一起，将点的坐标对直接作为目标的两个属性项，与点状地物的属性信息一起用表的方式存储。线状和面状地物就不可能这样，因为它们的坐标点个数是变化的，无法建立统一的数据表。可先为几何图形数据建立一个或几个文件，并对每一个目标设置一个唯一的标识号，然后为属性数据单独建一个表结构文件，用标识号将两者联系起来（表 5-3）。

表 5-3　GIS 中几何图形数据与属性数据的连接

ID	几何图形数据
	属性数据

2）矢量拓扑结构。在点、线、面等表征地矿实体的图形元素之间，还存在着某种空间拓扑关系，例如“包容”、“在里面”、“在外面”、“相连”和“相邻”，等等。一个面由若干条弧线组成，每条弧又由若干个节点组成，而在一个面内部，又可能存在若干个小的面（称为“岛”）。除了这些组合关系外，点—弧—面之间，还有连接关系。例如，一个节点上连接若干条弧，一条弧连接两个点，甚至连接两个面。所有这些拓扑关系，在空间数据库中都用结构表的形式来表示。从图 5-13 可以看出，除了从上到下完全地表达

面块、弧段、节点的关系外，还可以利用关系表列出面块、弧段、节点之间关系（表 5-4、表 5-5、表 5-6、表 5-7）。

表 5-4　面块与弧段的拓扑关系 P = P（A）

面	弧　　段
P1	A1　A5　A6
P2	A4　A3　A6
P3	A2　A4　A5

表 5-5　弧段与节点的拓扑关系 A = A（N）

弧段	节　　点
A1	N1　N3
A2	N1　N2
A3	N2　N3
A4	N2　N4
A5	N1　N4
A6	N4　N3

表 5-6　弧段与面块的拓扑关系 A = A（P）

弧段	左面块	右面块
A1	0	P1
A2	0	P3
A3	0	P2
A4	P2	P3
A5	P1	P3
A6	P1	P2

表 5-7　节点与弧段的拓扑关系 N = N（A）

节点	弧　　段
N1	A1　A2　A5
N2	A2　A3　A4
N3	A1　A6　A3
N4	A4　A5　A6

表 5-4 和表 5-5 表达了从上到下的拓扑关系，表 5-6 和表 5-7 表达了从左到右的拓扑关系。在不同的空间信息系统中，空间数据库可能采用不同的拓扑关系结构表，或者以弧段为纲，或者以节点为纲，或者以面块共享公共弧段、弧段共享公共节点的原则来从上到下建立几何拓扑关系，但所建立的结构表都能完整地表达节点、弧段和面块之间的拓扑关系。有的 GIS，例如，ARC/INFO 将与弧段有关的表 5-5 和表 5-6 结合在一起，但其中的节点坐标并不单独存储，而是分别存放在有关的弧段文件中。

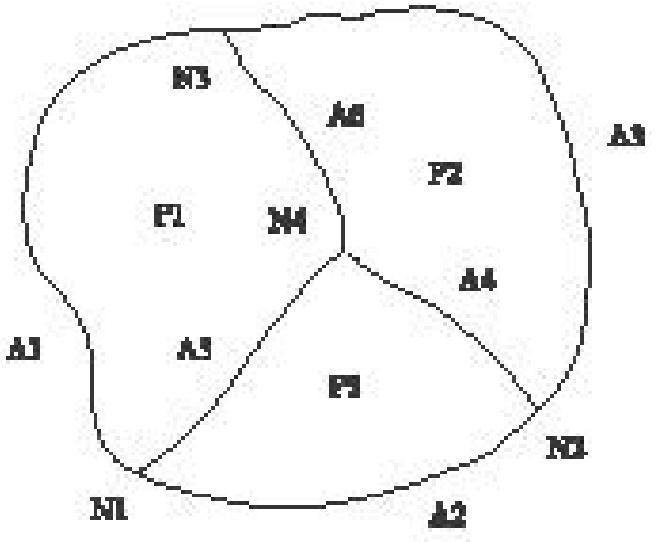


图 5-12　空间拓扑关系示意

利用这些结构表，很容易判别图 5-12 中反映地矿实体的多边形 P1 的相邻多边形是哪些，从节点 1 到节点 3 经过哪些弧段。

（3）矢量栅格一体化数据结构及其表达

矢量栅格一体化数据结构是当前众多 GIS 经常采用的一种特殊的混合数据结构。在以这种数据结构为基础的空间数据库中，二者与属性数据并存，相互关系如表 5-8 所示。

表 5-8　混合数据结构

ID	矢量数据
	栅格数据
	属性数据

矢量栅格一体化数据结构的构想，是从使栅格数据矢量化的角度提出来的。首先是提高栅格的精度，将点和线通过的网格分成 256× 256 等份；其次是记录细分网格的空间坐

标，采用基本栅格和细分格网并用的线性四叉树地址编码，把一个点的位置用两个十进制地址码表示：第一个表示点在基本格网中的位置，第二个表示进行细分格网的十进制地址码，即点的地理坐标 X，Y 转换成 M1，M2。对于线状地物和面状地物的弧段，除记录原始采样点的位置外，还要记录每条弧段通过每个基本格网边的交叉点，便于解决线性目标与栅格数据交互叠置的问题。于是，按照这种方式所建立的数据结构，便同时具备了栅格数据和矢量数据的特征，可以进行混合结构处理。

点状实体的数据结构　点状地矿实体的数据结构与矢量拓扑关系结构相似，只是一对坐标转换成两个十进制地址码。即：

Mode　1D	M1	M2	关联的弧段

弧段的数据结构　弧段的数据结构与矢量拓扑结构相似，它除了采用十进制地址码表示地矿实体空间位置特征外，还记录弧段通过的所有格网。即：

ARC　1D	起节点	终节点	左多边形	右多边形	中间点（M1，M2）

面状地矿实体的数据结构　在矢量数据栅格数据一体化的结构中，面状地矿实体的数据结构既包含了组成该面状地物周边的弧段，又包括该面状地物中间的面域栅格。面域栅格由二维行程编码表组织，在执行时采用循环指针将每个面状地物所属的叶节点串起来，然后建立面状地物的数据结构。即：

地矿实体标识号	周边的弧段标识号	中间面域叶节点信息

3．空间数据模型的设计

空间数据库的数据模型设计主要解决以下四个方面的问题：①空间数据的结构模式选择；②空间数据的有效组织方式；③属性数据在空间数据库中的表达方式；④空间数据与属性数据的逻辑关系表达方式。

(1) 空间数据结构模式的选择

不同的空间数据结构模式有不同的应用范围。要正确地选择空间数据的结构模式，首先应当弄清各种数据结构模式的用途，同时考虑数据本身的特征和数据应用目的。在选择地矿资源空间信息子系统的 GIS 平台时，也应该将内在的数据结构模式配置作为重要指标。

为了执行对各项地矿资源空间数据的描述、管理与处理任务，空间信息子系统内通常同时设置了矢量、栅格和栅格-矢量混合等 3 种数据结构模式。用户可根据空间数据的应用目的及其处理方式来选择其中一种，或者以其中一种结构模式为主，而以另一种结构模式为辅。采用栅格模式时，空间实体以像元形式来描述和存储；采用矢量模式时，空间实体以点、线、面形式来描述和存储；采用栅格-矢量混合模式时，空间实体则同时以点、线、面形式和像元形式来描述和存储。一般地说，矢量数据结构模式比较适合于描述、存储、管理和处理点状和线状空间实体，栅格数据适合于描述、存储、管理和处理面状空间实体，而栅格-矢量混合数据模式适合于描述、存储、管理和处理点、线、面混合的空间实体。这 3 种数据作用是互补的，必要时相互之间还可以转换。

在地矿勘查与开发工作中，人们常用栅格数据结构模式来描述、存储、管理和处理航天航空遥感数据、遥感图像、数字录像和数字摄影数据，各类地球物理异常、地球化学异常的数据和图像，以及航空照片、野外露头照片和室内镜下照片的扫描数据。利用这些数

据，可以快速地进行矿化信息叠置、组合和空间分析；可以定量地进行多因素综合成矿预测；可以方便地计算各种矿体、矿床和油气藏的位置、分布、面积和储量；可以对各类地质体和矿体的表面特征进行仿真模拟；等等。

矢量数据结构模式常用来描述、存储、管理和处理道路、河流、沟渠、建筑物、地形等高线、地层界线和断层等线状实体，以及各种具有复杂结构特征的土地、林区、水体（湖泊、海湾）和地质体（岩体、矿体、成矿区带、矿化单元等）多边形边界数据。对于这些数据，我们可以分区、分段、分点地随机进行空间位置、拓扑关系和属性的关联查询、检索和操作处理；可以用来细致地建立各种地矿和环境空间分布的表象结构，方便地进行控矿构造网格、控矿构造期次、矿液（或油气）运聚通道分析等；可以快速地进行地矿开发、利用和环境保护以及矿山建设和运输线选择方案的决策。对二维矢量数据面和三维矢量地矿数据体，允许进行任意的剪切处理。二维矢量数据面被剪切后形成的面片，可以提供进行局部区域的细微空间结构分析；三维矢量数据体被剪切后，可以形成任意方向的剖面图，也可以形成任意的小方块和小岩柱，提供进行局部地质体的细微空间结构分析。此外，还可以利用矢量数据来绘制各种高质量的地矿勘查图件。

栅格-矢量混合数据结构模式可以用于上述栅格数据结构模式的常规应用领域，也可以用于上述矢量数据结构模式的常规应用领域，但其优势主要在于描述、存储、管理和处理那些具有复杂内部结构的地质体数据。例如，多期次侵入的大型岩浆岩体、多期次拼合的构造地层体、多期次叠置的构造推覆体，以及综合表示区域地质结构的区域地质图和地质剖面图数据。这方面的应用刚刚开始，由于地矿勘查工作面临的都是复杂的地质体，栅格-矢量混合数据结构模式有着广泛的应用前景。在未配备栅格-矢量混合数据结构模式的空间信息子系统中，为了信息处理和利用的方便，可以采用栅格和矢量数据结构模式联合应用方式。这种联合方式主要用于设计和编制各种高质量的资源预测和勘查图件。为此，需要先用矢量数据结构模式将主要地矿勘查图件数字化，再用栅格数据结构模式存储各类遥感图像、航空照片、物化探异常图等叠置层。如果有矢量化的矿区（或勘探区、预测区）等高线图，也可以通过与这些栅格化叠置图层结合来建立其数字影像地图（DEM），但这需要有专用的软件。在实际设计制作图件时，地质体的边界和内部结构线条可以用矢量数据形成和绘制，而封闭地质体及其内部结构区的颜色可以用栅格数据形成和填绘。

（2）空间数据库的数据组织

在总体设计中，空间数据库的数据组织涉及空间数据的分层和分幅存储、管理问题。

1）空间数据的分层存储管理。空间数据的特点是大量不同性质的数据集（图元和图素）重叠地分布于同一个空间中。为了有效地处理和利用这些数据，GIS 采取了分层存储和管理的数据组织策略。对于地矿调查与勘查空间数据而言，层的划分应当遵循以下几个准则：①根据数据处理和应用的需要，图形应当按图素内容划分成若干图层；②图层划分要适应 GIS 分析处理工作需要和软件的功能特点；③具有相同逻辑内容的空间数据尽量放在一个图层中；④点、线（弧段）、面（多边形）分开的原则；⑤具有共同边界的区域（多边形）数据应当存放在同一个层中；⑥特征（或属性）、使用目的和使用方式相同的数据可以放在同一个层中；⑦为了便于维护，不同部门的数据通常应该放入不同的层中；⑧安全保护级别不同的数据应该存放在不同的层中；⑨为了便于更新，数据源不同的数据

应当存放在不同的层中；⑩即使是同一类型的数据，属性特征不同时也应该存放在不同的层中；1 特性（或属性）相同但获取时间、尺度或结果不同的数据应当存放在不同的层中；2 由于软件限制不能存储于同一层中的数据应当存放在不同的层中。

在实际工作中，上述准则应当统筹兼顾而不可偏废。以区域地质图为例，根据上述准则，可将各种地理要素和地质要素划分为 13 个基本图层（表 5-9）。

表 5-9　一种区域地质图图形和空间数据的分层方案

图　　层	内　　容	拓扑类型	属性项	注记文字
基本信息图层	图幅框架	点、线矢量	角点坐标、图幅号、比例尺、调查单位及出版年代等	有
水系图层	河流、湖泊、水库、海岸线	线、面矢量	名称和分类	
交通图层	铁路、公路	线状矢量	类别和名称	有
居民点图层	乡镇及以上各级政府驻地和居民点	点、面矢量	名称	
境界图层	县以上（含县）各级行政区划界线和其他特殊地区、自然保护区界线	线状矢量	类别和名称	有
地形等高线图层	地形等高线及高程、各类测量控制点、山峰高程点	点状、线状矢量或栅格	高程	有
地层图层	地层界线、地层单位；侵入岩界线、侵入岩；各类脉岩；各种围岩蚀变界线；各种混合岩化带、变质相带；各种特殊成因岩石	线、面矢量	类别、名称及其主要特征	有
断层（带）图层	各类断层（带）线	线状矢量	类别、名称及主要特征	有
褶皱图层	各类褶皱线	线状矢量	类别、名称及主要特征	有
韧性变形带图层	各韧性变形带	线、面矢量	类别、名称及主要特征	有
产状符号图层	地层产状和片理、片麻理、劈理、线理、流面、裂隙及褶皱等产状	点状矢量	类别、名称及主要特征	有
矿产图层	矿产地（矿化点、矿点、矿床）	点、面矢量	矿种、矿床成因类型、矿床规模、矿石品位、成矿时代及地质工作程度等	有
其他图素图层	火山构造、化石采样点、同位素年龄采样点、钻孔、震中、泉及图切剖面线等	点、线矢量	类别、名称及主要特征	有

2）空间数据的分幅存储与管理。在一般情况下，地矿调查与勘查数据的存储和管理是按图幅分别进行的。但由于某些地矿普查与勘探工作涉及广大的区域，特别是当所采用的比例尺比较大时，空间数据库管理的数据量急剧膨胀。面对这样的海量数据，单一的图幅整体管理方式已经不能适应需求，应当采用多图幅组合管理的方式。所谓多图幅组合管

理，是按标准规格将大区域的空间数据进行图幅分割，然后统一进行存储和管理。对空间数据进行分幅存储、管理，势必造成一个空间实体分属于多个图幅。为了做到既能按分幅数字化录入、存储和管理空间数据，又能够将分属不同图幅的同一目标建立起正确的拓扑关系，以利于对整个空间数据进行正确的检索、分析和统计，需要着重解决数据库中图幅的组织方法和图幅间被分割目标的组织方法。

数据库中图幅的组织方法　在这种可进行多图幅组合管理的空间数据库中，数据是以图幅为单位进行管理的。目前，国际上大多采用经纬线网格或规则矩形网格的方式进行分幅（图 5-13）。用户在进行数字化录入和数据编辑等操作时，均是面对所分割出来的单一幅图，只是当完成对数据的编辑之后，再将这些图幅组织到一个大的对象空间中。

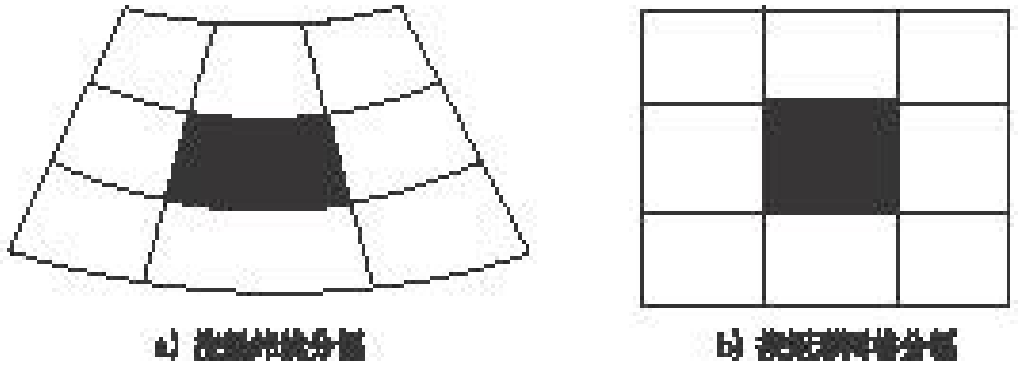


图 5-13　图幅的划分及其邻接关系示意

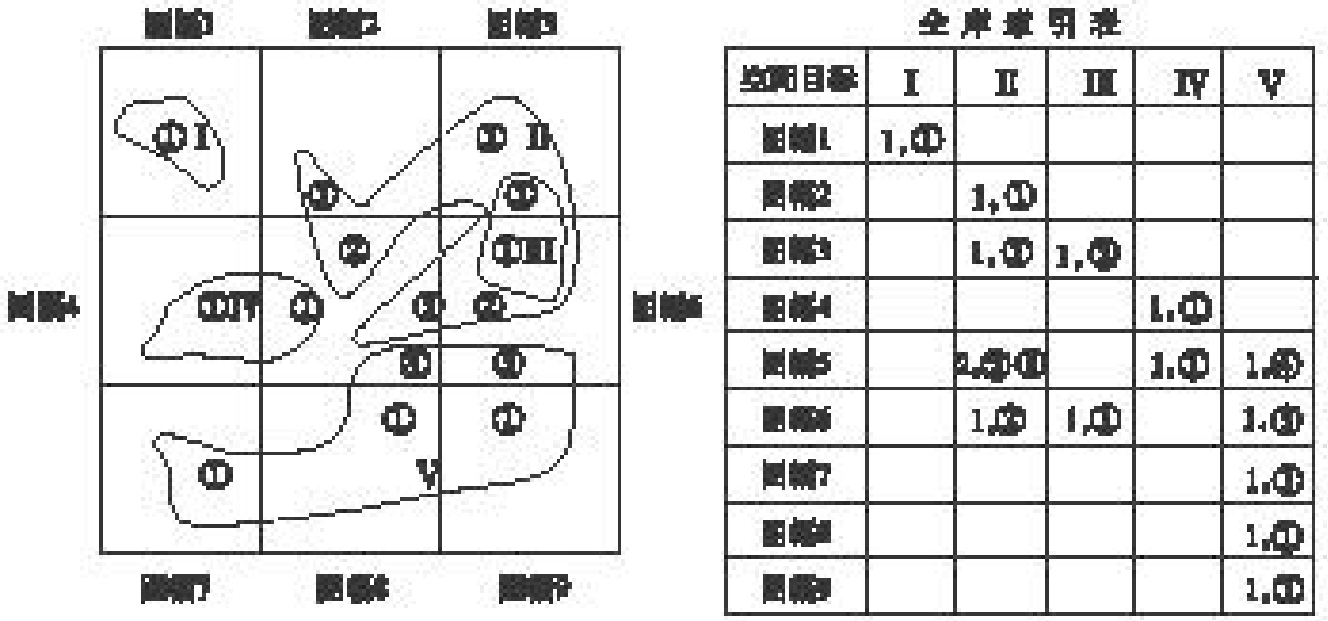
（据吴信才等，2000）

具体方法是：首先为对象空间的组织形式进行定义，其中包括：坐标单位、经度和纬度跨度（用经纬线网格分幅时）、比例尺、图幅的宽度和高度（用矩形网格分幅时）、地图投影类型、椭球体参数等。然后按照上述形式将对象空间内所涉及的图幅逐一录入到数据库中，同时指出该图幅的文件名（可能一个图幅含有多个文件）和用于确定图幅在库中的位置的横向纵向序号（矩形分幅）或左下角经纬度（经纬线分幅）。各种空间数据分幅录入后，系统可以针对用户按某局部范围或按分类要素进行查询的要求，自动检索与指定范围对应的图幅文件或者图幅之间邻接关系所需要的要素。

图幅间被分割目标的组织方法。按照上述方式将图幅输入数据库中时，并没有建立相邻图幅间同一目标的联系，为了有效地进行目标的跨幅拼接，还需要补充建立并对被分割目标加以组织。目前有两种常用的组织方式：一种是统一建立和管理对象空间的所有目标的联系；另一种是只建立和管理被分割目标的联系。

统一建立和管理对象空间的所有目标的联系，是指建一个全库索引表（刘纪平，1998），将整个对象空间的所有目标及其分属的图幅号均放入其中，不管目标是否被分割。如图 5-14 所示，Ⅰ、Ⅱ、Ⅲ、Ⅳ、Ⅴ分别表示全库中的目标编号，①～④分别表示各图幅中的对象编号。在全库索引表中，存放着同一目标在各图幅中的对象个数和编号。如Ⅱ号目标在图幅 3 中有一个对象，编号为①；在图幅 6 中也有一个对象，编号为②。除了全库索引表外，还应建立一个区域嵌套关系表，如Ⅰ号区与Ⅱ号区无嵌套关系，而Ⅲ号区嵌套在Ⅱ号区内。

用全库索引表来表示不同图幅间目标连接关系的方式，优点是结构清晰明了，缺点是索引表比较大，填写麻烦且无用的内容较多，因为大多数目标并没有被分割。



全库索引表

空间目标	Ⅰ	Ⅱ	Ⅲ	Ⅳ	Ⅴ
目标1	1,①				
图幅2		1,①			
目标3		1,②	1,③		
目标4				1,①	
目标5		2,③④		1,②	1,⑤
目标6		1,③	1,④		1,③
目标7					1,①
目标8					1,②
目标9					1,③

图 5-14　不同图幅间目标的连接及其全库索引表

(据吴信才等，2000)

只建立和管理被分割目标的联系，是指每拼接一个目标，就将该目标的信息（包括该目标在两个图幅中的编号）插入索引表中的适当位置。如果该目标还与第三个图幅有关，也照此办理，将两两之间的关系信息记入索引表中。由于这种方式只组织和管理因分幅存放而被分割的目标，其索引表比上述的全库索引表要小得多，填写简单且无用的内容也少。在检索目标时，可根据所要检索的图幅号来查询索引表，从中找出符合要求的完整的目标对象；对索引表中没有的目标，可通过直接检索该图幅中的目标信息来实现查询。

(3) 空间数据库的属性数据表达方式

在空间数据中所存放的属性数据，是一些对空间实体简明扼要的描述。各种地矿实体的属性数据在数据库中的表达方式，涉及属性表和属性关系的定义。大型空间数据库通常将属性数据与空间数据分别存储，其中，空间数据由各种 GIS 软件提供的数据模型方式存储，而属性数据则使用 RDBMS 存储。要在空间数据库中有效地存储和管理属性数据，需要对属性表与属性关系进行合理设计。这项工作通常包括下面五个步骤：①设计相互关联的表格；②对数据表进行规范化；③定义主关键字和外部关键字；④使用实体关系模型来定义关系；⑤对关系表进行再规范化。这方面的内容在数据库子系统的有关章节已经有详细阐述，下面主要结合实体—关系模型做些补充介绍。

实体—关系模型也称为 E - R 模型（Entity-Relation Model），是用实体关系来表示数据的模型。当实体很多、关系很复杂时，该模型能够帮助人们清楚地分析和表达出其中的关系，并设计出属性表。其分析设计过程如图 5-15 所示。

为了清理和识别各个实体之间关系，需要首先对有关实体进行编码，然后用一种简单的图表来表示，实体编码的唯一性是确定实体间关系的保证。实体间的关系主要是指一对一、一对多、多对一或多对多的关系。通常，一个成矿带可以有一个或多个矿床，一个矿床只属于一个成矿带；在特殊情况下，一个矿床也可能同时属于几个交叉的成矿带，而有些零散的矿床或矿点则可能暂时不归属于某个成矿带。因此，矿床与成矿带可能是多对一(M - 1)、一对多（1 - M）或多对零（M - 0）的关系，但总的看来可以称为多对多的关

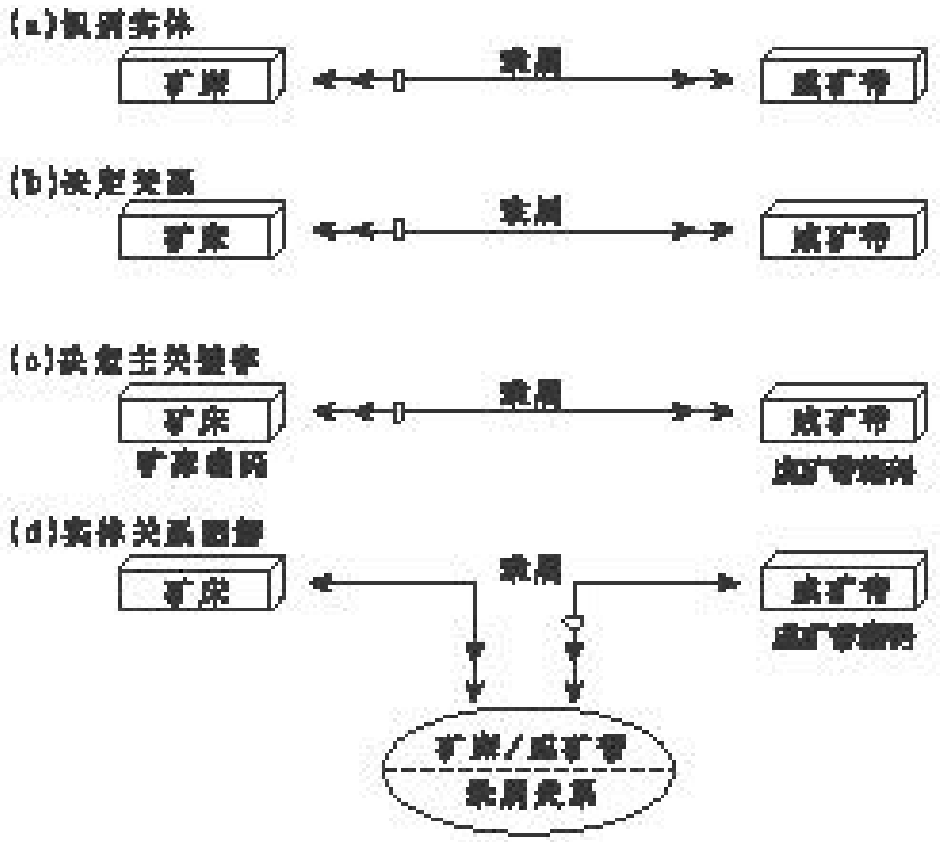


图 5-15　实体—关系模型设计过程示意

(注：单箭头一端代表单个实体，多箭头一端代表多个实体)

系。在空间数据库中，由于多对多的实体关系无法进行正确的数据查询（Date，1995），通常要求转换成几个一对多、多对一或一对一的关系。

(4) 空间数据与属性数据的逻辑关系表达

空间数据与属性数据的逻辑关系，可以用逻辑数据图来表达。以加拿大阿尔伯达省土地信息系统的空间数据库为例，与地籍有关的空间及属性数据库的逻辑关系如图 5-16 所示。其中空间数据包括点、线和面三种类型，是存放在 Arc/Info 中的。各种空间特征在相应的空间数据表中都有编码。这些特征的属性数据存储在 Oracle 关系数据库中。为了关联查询检索，各种特征的编码通常同时出现在空间表和属性表中，作为主关键字或外部关键字使用。显然，逻辑数据图是一种很有用的数据库逻辑设计的表达工具。

在总体设计接近尾声时，通常要求设计人员将整个设计和规划起草成正式文件。这样做，一方面可以为最终用户提供正式报告，作为监督管理和成果验收的参考依据；另一方面可以作为系统设计与开发者自身检测设计质量的依据。在编写总体设计报告时，应当着重注意以下几个问题：①将系统的结构—功能框架和目标加以概括和描述；②对空间数据库的数据结构模型和内容加以描述；③对空间数据的分层和分幅存储、管理问题加以描述；④在能够使用图表的情况下尽量使用图表；⑤详细的内容应当放入附录之中，以保持文本主体的清晰；⑥提供一个原型以便清楚地表示整个设计的概念。

4. 总体设计的书面总结

完成空间数据库的总体设计后，需要提交一份详尽的书面报告。其格式应当符合常规要求或预先约定的样式，其内容应包括上述数据库的宏观地理定义、空间数据的结构模式和空间数据模型设计等三个方面。

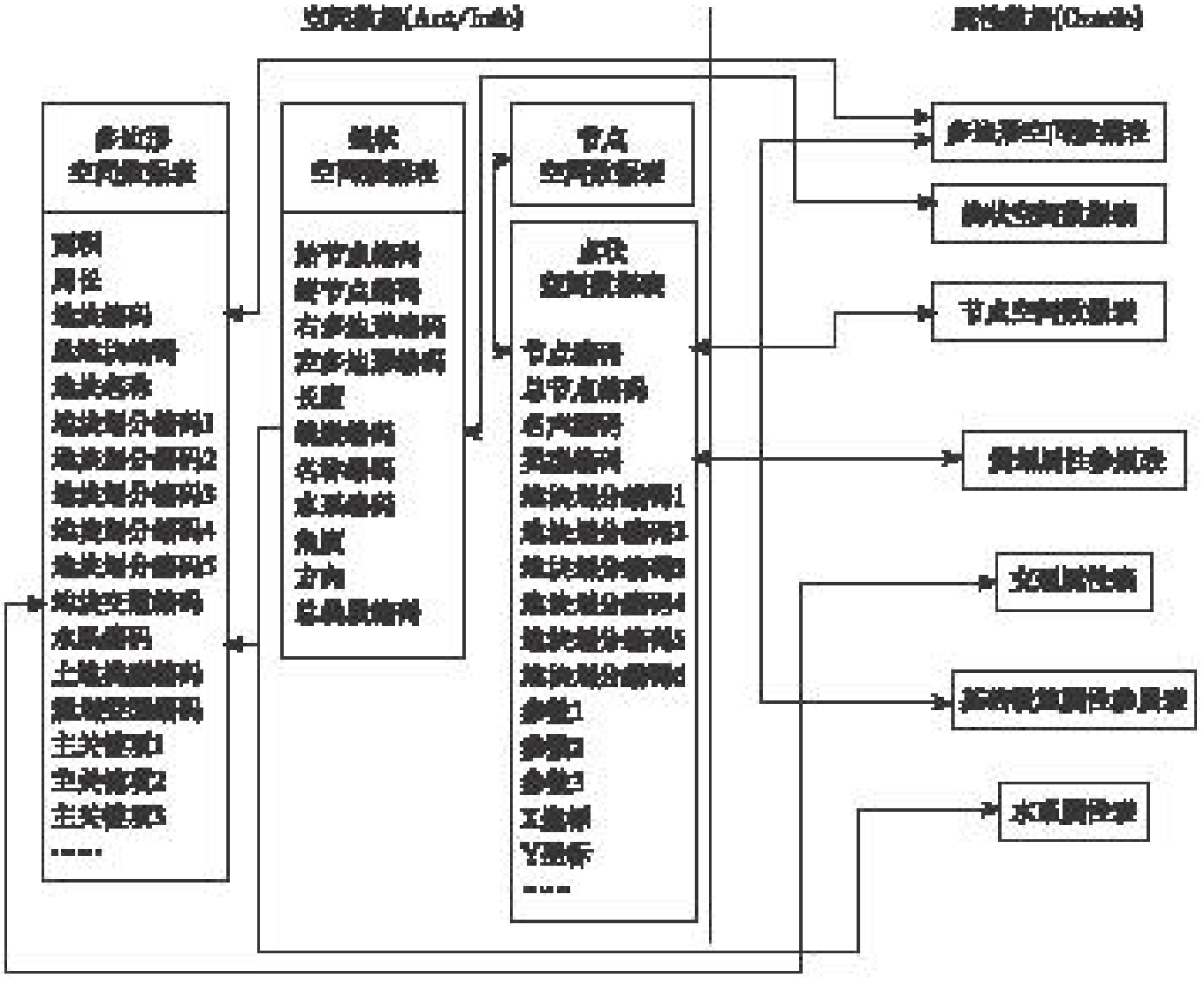


图 5-16　逻辑数据图举例

(据陈俊等，1999)

三、空间数据库的详细设计

空间数据库的详细设计主要包括数据质量评价和空间数据层设计等两方面内容（陈俊等，1999)。其中，各种数据集的评价包括数据一般状况评价、数据空间特征的评价和数据属性特征评价；空间数据层设计包括地质实体的空间数据结构、数据的定义与标识、编码标准的选择、数据字典的制作和空间数据的存储管理设计等内容。

1. 数据的质量评价

空间数据的来源多样化，可能造成内容丰富但质量参差不齐。这些数据是否可用，通常可以从三个方面进行评价：数据的一般状况、数据的空间特性和数据的属性特性。

(1) 数据一般状况评价

数据一般状况评价主要包括以下几点：

数据的目前状态　包括数据是否已有电子版，或是否有机构正在生产数据电子版。

数据形式是否标准化　主要指该类数据是否是在各政府机构或商业团体生产的数据的标准化之列，例如分类的标准、属性的标准等。是否符合本空间数据库的要求？点、线、面形式是否与本数据库的设计要求一致？

数据是否可以直接被 GIS 使用　某些数据需要经过一定的处理以后，才能与当前数据库中定义的数据相符合，从而造成当前数据库的实施困难。

数据的原始性　有些数据是由其他更原始的数据推导、综合而来的，不适合在数据库

中存储和管理。应该更注重使用更原始的数据，即第一手数据。

数据的可替代性　一种数据可能会有多种来源，有些容易获得，有些则较难。在决定使用哪一种时，应该先收集并仔细比较各种来源的数据，再做结论。

数据的一致性　即覆盖的地区是否一致，比例尺和精确度是否相同，数据的地理控制点是否符合本数据库的要求，投影方式是否与要求相符合等。

假若一个数据的电子版已经存在，那么可以使用下列的评价原则:

数据格式　数据格式是否与本数据库设计的要求相吻合？如果不吻合则需要进行数据转换，结果可能会是费力不讨好的，应该谨慎行事。

拓扑关系　CAD 来源的数据通常不具有所需要的拓扑关系。在数据转换过程中，则可能要求建立各种拓扑关系，而这也是一项费时费力的工作。

数据分辨率　分辨率是衡量数据精度的一个重要标志。对于栅格数据而言，数据的分辨率是指像元的大小，像元尺寸越小，分辨率越高；而对于矢量数据而言，分辨率指其原始数字化资料的比例尺，原始资料的比例尺越大，分辨率也越高。

数据覆盖面及一致性　该类数据是否能够覆盖整个工作区域？若能，全部数据是否具有同样的质量并符合有关标准？若不能，是否能够找到相匹配的原始资料？

数据可获得性　包括数据存储的媒介、适用的计算机平台、是否可以购买到或免费提供？能否在本数据库实施计划的时间内获得？价格和维护的费用如何？

样本数据　在数据评价过程中，如果能够有机会获得该数据的样本，一方面可以提高对数据的感性认识；另一方面将有助于对数据结构模式和组织方式的进一步选择。

加工的自动化程度　现有的数字化数据通常并不完全符合某一数据库设计的要求，可能需要做不同程度的加工才能入库。假若这种加工过程能很容易地自动实现，那么该数据可以接受；否则，应该比较数据加工和重新数字化的效益，再行定夺。

元数据信息是否完全　元数据是有关数据的重要信息，它的存在与否及质量直接影响到该数据的价值。元数据信息不全的数据，不要轻易使用。

(2) 数据空间特征评价

数据空间特征的评价主要有以下几点:

空间特征的表达方式　例如，城市既可以作为点，又可以作为多边形；地形数据既可是等高线式的矢量表达方式，又可以是栅格的数字高程模型。因此，应当比较数据的各种空间特征是否符合特定的要求。

空间特征的连续性和闭合性　在数字形式的 CAD 数据中，很多线性特征的表达是不连续的，例如铁路线的表达，而且有些面状的地理特征是不封闭的。对于不连续、不闭合的情况，需要自动或半自动地进行处理，以保证适合于 GIS 的要求。

表示规则比较　不同数据集在对同一类型的地理特征进行表示时，可能使用不同的规则。例如河流信息，有些是用双线表示，有些则用单线表示；又如油井，有时用多个点聚集表达其分布状况，有时则用多个多边形表达其覆盖范围。具体采用什么表达方式，在进行数据库详细设计过程中，要根据应用目的加以斟酌处理。

地理控制信息比较　不同数据集通常使用不同的大地控制系统：①GPS 点；②大地控制测量点；③人为划分的地理位置点，如图幅角点等；④道路等线性特征的交叉点。不同的方法代表不同的精度，在详细设计过程中应当加以比较和评价。

空间地理数据的系列性　当不同地区的同一类型数据比例尺不同，或覆盖有交叉重叠时，要决定不同地区的信息的衔接问题。这种决策经常是按主观经验而定的，不仅要考虑到整个数据库的质量，而且要兼顾实施的难易。

分类方法的比较和评价　不同数据集对同一类型的数据通常使用不同的分类方法。例如同样的盆地，石油地质界根据其与油气藏关系分为含油盆地和非含油盆地；煤田地质界根据其与煤层的关系分为含煤盆地和非含煤盆地；非金属地质界则可能根据其与盐类矿床的关系分为含盐盆地和非含盐盆地。有时两种数据的这种差异不论是从空间图形的角度还是从属性信息的角度，都可能会大到几乎无法匹配的地步，况且，即使同一种分类方法，随着时间的迁移也会起变化，因此分类问题在详细设计过程中应该引起重视。

地理参考系统的一致性　同一地区，不同地理特征的地理参考系统可能会由于比例尺、原始信息、年代的不同而出现不匹配的效应。例如当一条河流是行政边界时，它将会出现在水系层和行政边界层上，若两者的地理参考不匹配，进行图层叠合时将产生很多冗余信息，需要加以消除。消除冗余信息的办法通常有以下 5 种：①选择较为精确的一层的信息取代另外一层的相关信息；②根据冗余面积的大小，设定一个域值来消除（自动化方法）；③在数字化过程中自动地消除冗余多边形（半自动或全自动）；④在地图数字化之前，事先删除冗余多边形；⑤当差异很大的时候，采用拉伸法填充空隙。

(3) 数据属性特征评价

属性特征的评价主要从以下 4 个方面进行：

属性的存在与否　很多空间数据并不具有属性数据或不直接拥有所需要的属性数据。在详细设计过程中，对各数据层（图层）均要考察其属性数据存在与否。

属性数据与空间位置的匹配　很多属性数据以报表方式存在，没有图形信息与其直接匹配，需要使用编码的方法将属性数据的位置数据自动或半自动地产生出来。

属性数据的代码体系　不同来源的同一类数据的代码体系常常不同，需要查明并列出对比表，然后根据行业和部门的统一标准确定转换方案。

属性数据的适应性　各类属性数据的适应性总是随着时间的变化而变化的，在数据库详细设计过程中，对每层数据的属性数据的适应性应加以严格考察，以便制定数据更新策略，保持整个数据库的现实适应性。

2. 空间数据层的设计

空间数据层设计是将总体设计具体化的基本过程，主要内容包括：①空间数据的定义与标识；②属性数据类型设计；③编码标准设计；④各表格中主关键字选择；⑤数据字典的制定；⑥空间数据的存储管理设计。

(1) 地质实体的空间数据结构

根据面向对象技术，空间中的每个地质实体都是一个对象，每个对象都应当有一个唯一标识，它是地质实体空间对象操作的基础。在进行面向对象空间数据库设计时，把一个空间对象定义为一个 4 元组：[Oid，Aid，Gid，Pid]，其中：

Oid 为该空间地质实体对象的标识（Point、Polyline、Polygon），系统内定，不可修改；

Aid 是属性数据标识，可编辑修改，如地质点号，它是联系空间和属性数据的标识码；

Gid 是空间数据标识，为数字型字段，一般不要修改，为机内码，用来唯一标识每一空间图形实体对象;

Pid 是静态图像数据标识，为字符串字段，主要标识素描或照片存储部位。

这一 4 元组空间数据结构定义是空间数据库的核心，也是进行空间数据存储与操作的基本单元，任何数据的更新都反映在这一基本结构的更新上。

(2) 数据的定义与标识

1) 空间矢量数据。地矿资源勘查与开发领域的基本二维空间数据类型有地质实体抽象表示的点、线和面。为了减少数据冗余，采用以结点和弧段为基础的数据定义和标识，即只有结点和弧段才包含空间位置信息，其他空间对象都是由这些基本结点和弧段构成的。

一个结点 (Node) 是一个元组: [Point, Nid, X, Y]

其中，Point 是标识该结点为地质点类，Nid 是该地质点的数字标识; X, Y 为该点的唯一空间位置坐标信息。

一个弧段 (Arc) 是一个元组: [Polyline, Aid, (Xi, Yi)]

其中，弧段是一条连通的弧线，Polyline 是标识该弧段为地质界线类，Aid 为该弧段的唯一数字标识，(Xi, Yi) 是组成该弧段的点的空间位置数据。具备了基础空间数据(即点和弧段) 后，便可组织基本空间对象 (即点、线和面对象)。

一个地质点对象 (Point) 是一个元组: [Poid, Nid, X, Y, Ai]

其中，Poid 是该地质点实体对象类的标识 (Point); Nid 是标识该地质点实体的属性数字码; X, Y 是其对应的地质点实体的绝对地理坐标标识; Ai 是可选项，为该点所必要的描述性属性数据字段标识。

一个地质线对象 (Line) 是一个元组: [Lid, Nid, Aid]

其中，Lid 是该地质界线对象类的标识 (Polyline); Nid 为其标识该地质界线实体的机内标识码，由系统自动予以唯一赋值; Aid 是地质人员根据实体对象的类别给予的标识各地质线状类型的识别码。如同为正断层的地质界线属同一识别码。

一个空间地质面实体对象 (Region) 是一个元组: [Rid, Nid, Aid]

其中，Rid 是该地质面实体对象类的标识 (Polygon); Nid 为其标识该地质面实体的机内标识码，由系统自动予以唯一赋值; Aid 是野外地质人员根据面实体对象的类别给予的标识各地质面类型的识别码。如同类黑云花岗岩属同一识别码。

2) 属性数据。空间数据库中的属性数据是狭义的，专指与一个空间实体对象相对应的非空间、非多媒体数据，例如名称、类型以及各种特征的描述数据等。属性数据的类型有字符型、日期型、整型、实型、二元整型、大型二元数据对象 (BLOB-Binary Large Object) 等。

在地矿资源空间信息子系统中，这类属性数据采用与部分空间数据相应的组织形式来管理，以便提高空间数据管理和检索查询效率。

一个空间实体对象通常有多个属性数据项，可看成是一个元组:

[Aid, Tid, A1, A2, …, An]

其中，Aid 是该属性数据的标识，它与对应的地质空间实体唯一机内码对应; Tid 是相应空间地质实体对象类型标识，与标识相应地质空间类型的识别码一致，为数字型值; Ai ($1 \leq i \leq n$) 是第 i 个描述性属性值，可根据基础调查、地矿勘查或开发的用户需要加

以取舍，其类型可以是数值或文字。在空间数据库中只存储组合项（元组），需要单项属性值时再设法提取。组合项在多数情况下被用来做为主关键字（Primary Key）使用。

一个地矿资源空间实体对应的属性数据，可以通过 Aid 对所有的空间地矿实体对象进行分层管理，从而有助于实现空间数据和属性数据的双向查询检索和空间分析。对于准备入库的属性数据，应当根据所选 GIS 软件的要求具体地定义其类型和输出宽度。

3）静态图像数据。根据地矿勘查的实际需要，本系统内专门设置了以栅格结构为基础的静态图像数据存贮、管理模块。静态图像数据是指通过扫描、素描、照相、分析和处理（例如压缩）后以某种图像格式存储的图像数据，可以是一般的素描位图、照片图像，也可以是遥感影像图像。

一个地矿空间实体（主要是点实体）所具有的静态图像数据也可以看成是一个元组：

[Pid，Ft；B1，B2，…，Bn]

其中，Pid 是相应地质空间实体的静态图像数据标识（内定为 Image）；Ft 为其存储格式（系统支持的图像格式有 BMP、JPG、TIF）；B1，B2，…，Bn 为图像的二进制内容。

在系统对图像的实际管理中，只需地矿勘查人员把地质空间实体对应的静态图像如素描和照片等以给定的格式存入指定的工作目录（表 5-10），在空间数据库结构中，给所定义的图像字段赋以图像名字符串值，就能达到对静态图像数据的管理目的。

表 5-10　野外地质点实体空间数据组织表

Shape	ID	地质点号	图像名	X 坐标	Y 坐标	…	…
Point	1	0145	0145. bmp	135. 21	31. 25	…	…

（3）编码标准的选择

编码标准的选择是空间数据库详细设计的一个很重要的内容。长期以来，人们在开发各自目的数据库过程中均根据各自的要求进行编码，不同的人设计会有不同的编码。即使同一个设计者在不同时期也可能使用不同编码。结果使得各自开发的数据库不能与其他机构实行共享，更有甚者连自己先后开发的数据库也不能共享。

目前美国、加拿大、澳大利亚及其他西方发达国家均在政府的积极推动下，开始使用国内或国际的标准编码系统。我国政府也于近期颁布了一些行业标准，例如：《国土基础信息数据分类与代码（GB/T 13923 - 92）》、《国家基本比例尺地形图分幅编号（GB/T 13989 - 92）》、《地质矿产术语分类代码（GB 9649 - 88）》、《区域地质图图例（GB 958 - 89）》、《数字化地质图图层及属性文件格式标准（DDB 9701）》等等。在选择代码标准时，对于同一类型的数据而言，在有相应的国家标准或行业标准时，应首先采用国家标准或行业标准，特别是应该寻找已被普遍使用的国家标准代码；在没有相应的国家标准或行业标准时，要根据用户所在部门和应用目的，对类似的部门标准进行比较分析，从中选定最合适的标准来。假若现有标准代码不能完全满足要求，可在其基础上加以补充。

（4）数据字典的制定

关于数据字典的概念和设计方法，前面已经有专门介绍，这里结合空间数据库的特点再作一些补充说明。空间数据库的数据字典实际上是空间数据库的元数据，它描述数据库的整体结构、数据内容和定义等，是一种伴随数据库开发和维护而不断修正、不断更新的

动态文件。它既可以作为数据库建设、维护和更新的依据；也可以作为使用数据库的指南，并可以帮助用户开发和使用应用程序。为了方便使用，数据字典通常都有在线（On-line）版本。

1）数据字典的内容。一个数据字典通常包括以下几个内容：①数据库的总体组织结构；②数据库总体设计的框架；③各数据层的详细内容定义及结构；④数据命名的定义；⑤元数据内容等。其中，数据库总体设计框架部分包括：数据来源、整体命名方法、各特征的值域、有效值、地图投影方式、图幅匹配及精度、线和多边形的拓扑关系、线和多边形的连续性及封闭性、质量控制的过程和内容、数据的文件和表格等。

数据字典中各数据层的详细内容定义及结构包括：①标题类信息：名称、类型、数据质量；②各层的有关文件、表，各表的项及各项的定义、有效值范围等；③地理参考方面要求满足的情况；④其他有便于说明和理解的文字或图表等；⑤各层空间及属性的质量控制规范；⑥各层编号系统与其他各标准编号系统的关系；⑦各层数据的使用与各应用类型的关系等。

2）数据文件的命名方法。与数据有关的信息包括数据内容、所接受的操作、被使用同一类型操作的先后次序等。这些信息可通过数据文件的命名来反映。

如图 5-17 所示，文件名 SO01DG 表明该数据文件是土壤数据，是第一次通过数字化所获得的。不同 GIS 软件的数据文件命名规则和数据操作命令不尽相同，需要具体了解并严格遵循。

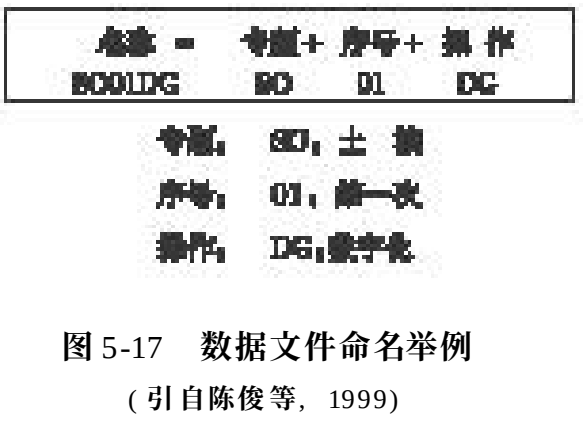


图 5-17　数据文件命名举例
（引自陈俊等，1999）

3）数据字典的元数据。元数据是“数据的数据”。它是对一个数据集的内容、质量条件及操作过程等的描述。元数据的规范化与标准化是实现地矿资源空间数据共享的一个重要条件。在经济发达的西方国家，例如美国，联邦政府的地理数据委员会（FGDC）已将元数据的内容规范化，并要求所有政府部门实施这个标准。目前不仅政府部门已经普遍地在使用这一标准，私营商业及学术机构也一并效法，使整个社会均使用统一的标准。

（5）空间数据的存储管理设计

1）数据精确性与精度标准的制定。通常使用精确性（Accuracy）和精度（Precision）来定义数据库坐标的可靠性，因此，数据精确性与精度标准的制定是数据库设计的基本内容之一。从地矿调查与勘查的角度看，空间数据的精确性是指在一个给定的误差范围内可以找到特定地质地理特征的能力。在制定数据库的数据精确性标准时，应当考虑以下基本事实：①数据的精确性不会因为采用自动化采集手段而增高；②经数字化过的地理数据特征，其精确性不会高于其原始地图；③数据库总体精确性是最低精确性部分的数据的精确性。

为了保证数据的质量，各国地图测绘部门、地质调查部门和土地管理部门都对与各有关的地图、地质图、矿产图、土地分布图等的精确性作了严格的规定，应当认真遵守。

精度是一种测度的可重复性，数据可以用单精度存放，也可以用双精度存放。单精度一般可以表达 6 到 7 位的数字，而双精度则是表达到 14 位。在进行数据库设计时，应当特别注意以下几个问题：①假若使用不充足的字位来存储数据，精度可能会降低；②使用双精度存储数据，数据库容量可能至少增加 20%；③精度存储方式不会明显地影响系统

的处理速度；④若不同层的数据必须要相互匹配的话，应当使用双精度存储。

2）空间数据库的管理模块设计。空间数据库的管理模块设计主要包括以下几个方面：

数据使用权限的设置　空间数据库应当允许许多用户同时使用同一种数据，但从数据库安全角度着想，各数据层均要由数据库管理员设置用户权限。权限设置的内容包括：全部可读可写，部分可读可写，全部可读不可写，部分可读、部分可写，或者部分可读不可写。用户只能按照所授权限范围使用数据，没有得到授权的用户则不能使用数据库。

数据库更新过程中的质量控制和安全控制　在更新数据时，除了要求操作者必须拥有可写性的权限外，为了保证数据质量，还应当建立一整套数据库维护规则。这种规则可以是关于整个数据库的，例如投影类型、大地坐标参照系统类型等，也可以是针对单一数据层的。同时，在进行数据库维护更新时，还需要考虑属性数据和空间数据之间的相互关系。因为当空间数据被修改后，与之有关的属性信息通常也要发生相应的变动。此外，如果采用网络版的 GIS 软件，为了保证所修改数据的安全性和完整性，所设置的安全控制应当能够在实施更新操作时将该数据层锁住，以免其他有权限的用户同时进行类似的处理。目前，国际上通用的网络型 GIS 软件系统均有类似的功能，设计者应当根据所使用的 GIS 软件的特点，设计出数据库更新过程中的质量控制和安全控制的操作模块。

数据库恢复的功能模块设计　数据库恢复是指允许数据库经过一系列修改更新之后，能够回到从前某一时刻的状态并重新运行的能力。这是评价数据库功能强弱的因素之一。许多商品化的数据库管理系统均提供重新运行（Roll-Back）的功能，但强弱有差别。设计者首先应当根据需求分析的结果选择适用的 GIS 软件，然后针对所选择的 GIS 软件的特点，设计数据库恢复功能。

合理管理单元的设定　一个空间数据库所覆盖的面积可能很大，为了管理和维护的方便，某些 GIS 软件将数据库的地理区域按某种方式——不规则的自然小区或规则的人为格网划分成更小的管理单元。例如一个省的土地资源数据库可以以县为单位划分成许多管理单元。在以不规则的自然小区为单位划分管理单元时，通常要求考虑下面三个因素：①该单元是相对较稳定的；②在该单元所覆盖的地区内地质地理特征相近；③重要的地质地理特征置放于同一个单元之中。在以规则的格网为单位划分管理单元时，则主要考虑格网的性质、坐标系、面积和比例尺。

数据库系统的网络功能设置　计算机网络是一种分布式处理设施，给更有效地结合各种技术来建立一个共享的用户环境提供了可能和工具。以城市 GIS 环境为例，一个城市可以由各有关单位来共同建立一个城市信息数据库，这个数据库共用一个文件服务器，使用商业化的关系数据库管理系统技术、计算机辅助设计技术和图像处理技术。各种数据采集设备，如数字化仪、扫描仪、立体测图仪等，都可以与之相连。一旦网络建立以后，它的资源可以不断地增加，处理应用的能力也相应地不断增大；城市规划、市政工程、社区发展、设施管理、土地利用、城市交通、公共教育等各单位的用户可以利用这个网络进行通讯，交换信息，最大限度地共享各种资源。在进行空间数据库的管理模块设计时，应重视网络技术的应用。

在数据库设计过程中应用网络技术应注意如下问题：

系统结构的扩充性　网络技术的发展日新月异，为了更好更快地利用新的技术成果，

设计者应该充分考虑其可扩充性，注意采用多种系统平台的连接和集成化技术。

系统功能的增长性 随着技术和用户需求的发展，用户对网络和 GIS 的使用次数将会不断增多，并且数据处理的难度也将不断增大。设计者应当正视这一趋势，重视多用户功能的设置，同时给用户提供再次开发平台，还应当允许挂接各种功能处理软件。

系统负载的均衡性 整个网络的设计应该尽量地做到负载均衡，避免任一部分负载过重，以延长网络的整个寿命，降低网络的出错率。

数据流量的合理性 数据与其用户应当尽量存放在一个网络（子网络）上，避免数据的迂回传输，以便减少网络中的流量，提高数据传输效率和速率。

网络连接的有效性 当必须在网络间进行数据传输时，设计者应当着重考虑网络连接的有效性问题，同时要注意对传输路径进行优化，尽量减少网络中不必要的传输。

设备名称的相关性 在给服务器、工作站、硬盘等设备命名时，其名称应具备一定的相关性，使用户能够方便地识别出其间的关系。

作为一般基层用户，为了降低成本，可以考虑采用 X 终端。X 终端是一种类似于工作站的设备，但不具有中央处理的能力，而只有自己的显示窗口系统。该窗口系统具有很好的图形显示功能，在登录（Login）到其他的工作站或服务器上之后，便可以分享它们的处理能力、内存和硬盘存储空间，并以此来实现空间数据库的管理和处理功能。

四、系统实施、测试与优化

1. 实施规划的制定

在完成空间数据库的总体设计和详细设计之后，接着需要制定应用程序编写、数据采集、数据处理和人员培训等实施规划。

(1) 应用程序编写实施规划的要点

包括：①根据软件工程规则分解编程任务，确定编写组成员并进行具体分工和进度安排；②制定实施质量控制的规则与标准，建立时间进度和程序质量的评估体系。

(2) 数据采集实施规划的要点

包括：①选择数据采集的方式方法；②建立实施质量控制的规则与标准；③建立实施进度和预算评估的跟踪系统；④将空间数据库实施规划与整个地矿信息系统的实现结合起来。

2. 数据采集方法设计

空间数据采集所涉及的工作内容和过程主要有三种：①原始数据资料的采集：例如 GPS 数据采集、地面测量数据采集和摄影测量数据的采集；②图像和图形的数字化：图像和图形的扫描和手工数字化、属性数据的键盘输入等；③数据的转换：现存的数据格式转换成数据库所需要的格式和定义。数据采集方式设计的任务是根据数据源的特点、工作条件和实际需要，为系统选择适当的采集方式、方法和技术手段。

(1) 原始空间数据采集方法选择

地矿调查、勘探和开发所使用的原始空间数据，通常利用 GPS 方法、摄影测量方法和测量法等三种方法来采集。

1) GPS 方法。GPS 是依靠导航卫星来决定地球上某一位置坐标的技术，是一种既便宜又精确的空间数据采集技术。目前 GPS 技术已经用在大地测量、军事目标定位、区域地质填图、矿产勘查、土地详查、地壳运动和地质灾害监测、航海、汽车导航、登山等领

域。GPS 对地测量的理论精度已经达到毫米级水平，但由于人为的干扰导致实际精度大大降低，需要利用各种方法（例如差分）进行矫正。

在地壳运动和地质灾害监测中，通常使用大型的差分式高精度 GPS 来动态观测地表的位置变化。这种高精度的 GPS 需要有高级测量控制点作为参照，设备价格也相当昂贵。在矿产勘查和土地详查工作中，通常使用轻型的差分式中高精度 GPS 来测定钻孔（井）、地质点、地质体、土地边界的位置。这种中高精度的 GPS 可以不用控制点作为参照，但价格仍比较高。在 1∶ 25 万区域地质填图工作中，通常使用手持式微型低精度 GPS 来测定地质点和地质体的位置。这种低精度的手持式 GPS 小巧方便，理论误差在 15 m 左右，但由于各种因素干扰，实际误差在 30 m 左右，且误差偏移方向是随机的，仅适用于小比例尺的区域地质填图，不适宜作为大比例尺的地质填图的准确定位手段。但如果与航空照片或地形图配合，在地形地物标志不易辨认的山区、高原、沙漠、草原等地方，可以对山头、山脊、山谷、隘口、河湾、大树、古人文遗迹等进行控制性的辅助定位。

GPS 所采集的数据是 X，Y，Z 坐标点，这些点的信息可以先以文本的形式存储，然后根据数据库和 GIS 软件的要求，使用简单的程序将其转换成有拓扑关系的图形加属性信息。

2）航空照片测图方法选择。各种大比例尺的地图和数字影像地图都是使用航空照片测图方法制作而成的。航片测图方法是一种先进的数据采集方法。它精度高，但费用开销大，在确定是否使用航片测图方法时，主要应当考虑下列几个方面:

比例尺大小　航片测图的比例尺与造价关系很大。确定比例尺的依据，主要是实际工作需要和数据库的精度要求。

提取特征的选择　只提取数据库中所需要的特征，避免大量无用的信息存入。

数字正射影像数据　如果时间和费用允许的话，尽量生成数字正射影像数据。因为它是很好的背景图像，既能对整个数据库的精度检测和美观起很大作用，还能作为屏幕数字化的底图。利用软拷贝摄影测量技术，可以用来生产数字化的正射影像图，也可以支持三维的屏幕数字化。数字正射影像数据易于在 GIS 中存储和管理。

多边形数据的提取　目前应用较广泛的与摄影测量方法有关的软件大多还是 CAD 类型的软件，数据类型主要是点和线，而没有多边形。在空间数据采集和建库过程中，需要将所提取的某些线特征转换成面特征。

3）地面测量数据采集方法选择。使用各种经纬仪、水准仪、测距仪进行地面位置和高程测量，是一种传统的空间数据采集方式，各国土资源管理部门均或多或少地拥有该类数据。这类数据系统性好，而且精度很高，直到现在还被作为主要的控制数据。全国性和区域性大地测量，通常采用具有标准水准点（例如黄海海平面）和标准坐标系（例如北京坐标系）的三角网水准测量方法，由国家专业测绘部门执行。所获取的中小比例尺测量成果图件和数据，由国家测绘局及其所属的各省市测绘局保管，通过正常渠道可以购买或索取。

地矿勘查区、矿区、水利水电工程、交通工程和城乡建筑工程所使用的大比例尺图件和数据，则往往根据施工单位的实际需要，以国家大地测量三角网点为基准，分别选择小三角网或导线测量方法，交由自己的测量队伍或委托专业测量队伍来执行。这种地面测量数据都可以转换成文本格式，然后通过运行简单的应用程序转换成图形数据。值得一提的

是在一些金属矿区、煤田、油田和大型水利工程的中大比例尺测量中，常常采用局部坐标系，甚至在不同勘查阶段采用不同的坐标系，例如长江三峡坝区地质勘查工作先后使用了大坝、蛇山、北京、永久船闸、临时船闸五套坐标系统，图件的拼接和资料利用十分困难，需要将其转换为统一的坐标系统，才能进行地质分析、评价和设计。为此，在空间数据库系统中，应将不同坐标系之间的相互关系一并存入，便于坐标换算。

(2) 图像和图形数据的入库方法选择

有些现今采集的原始空间数据和早期的工作成果是以非数字化图像和图形方式表达的，必须将其转换成计算机可以识别和接受的电子数字形式才能存入数据库。这个转换过程通常被称为数字化。数字化方法有两种，其一是手扶跟踪法；其二是扫描法。

1) 图形数据的手扶跟踪数字化。手扶跟踪式数字化方式是最常用的一种方式，它是一种用鼠标的方式手扶跟踪地图上的各种地质地理特征，以获得 X，Y 坐标。这种方法适合于图形数据的数字化，可直接将图形数据转换为矢量数据，但具体操作较为麻烦，费工费时，效率较低，通常在图形较为简单的情况下使用。为了保证数据质量，在操作时应该注意如下几点:

设定控制点层　在数字化前先产生一个控制点层，每幅图至少有四个控制点，以便保证整个数据库中各图幅的精度从而保证整个数据库的总精度。

设定阈值　该阈值是指在跟踪数字化过程中任何两个节点的最小距离。任何两个节点间的距离小于该阈值的，后面输入的点位将不被记录。

2) 扫描数字化。扫描数字化是使用扫描仪将整幅图扫描成栅格图像数据，再进行屏幕跟踪矢量化的方法。该种方式既适合于图形数据的数字化，也适合于图像数据的数字化。如果仅需要栅格数据，这种数字化方法的效率是很高的；如果还需要矢量化数据，则通过屏幕跟踪来实现矢量化也是较为费工费时的。目前已经有一些对扫描数据进行自动矢量化的软件，但仅能处理简单线条图形，难以面对复杂的地质图件，特别是难以应付彩色的区域地质图。在对彩色图件进行扫描数字化时，通常要求对原始材料进行预处理，例如将地图中色彩不同的地质地理特征先分色，复制在透明薄膜上，然后再进行扫描。目前已研制成功自动的分色扫描仪，甚至提出了自动分层建库的方法 (Wang and Zhang, 1996)。

(3) 属性数据入库方法选择

属性数据的数字化有四种方法，应根据使用单位所拥有的设备条件进行选择。

键盘输入法。这是一种常规的输入方法，最为经济、简便。大多数属性数据的输入以此方式为主。

扫描字符识别法。扫描字符识别法是采用扫描方式将字符转换栅格数据，再利用专门的软件来识别并还原成文本文件的方法。目前已经有许多软件能够实现纯文字的扫描输入识别，但还未能突破图形和报表中的属性数据扫描输入后的自动识别和拾取难关。

过程赋值法。在数字化或矢量化的过程中，通过调用属性表来赋值。

人工编辑。某些实体的属性值可以利用 GIS 的编辑和分析功能进行自动赋值。例如利用基准曲线的已知值来给相间的曲线进行内插赋值。

3. 数据库的质量控制规划

空间数据库的质量控制是指对空间数据及其对应的属性数据的质量检测。

(1) 空间数据质量检测内容

包括：①空间位置的几何精度；②空间地理特征的完整性，是否所有的内容均数字化；③空间特征表达的完整性：例如面状的特征是否以面状的多边形进行表达；④空间数据的拓扑关系；⑤空间数据的地理参考系统是否正确，是否满足整个数据库使用的最低要求；⑥空间数据所使用的大地控制点是否正确；⑦边界匹配如何。

(2) 属性数据质量检测内容

主要包括：①属性表的定义是否符合数据库的设计；②主关键字的定义和唯一性怎样；③各数据项的值是否在有效范围以内；④各属性表的外部关键字是否正确；⑤关系表之间的关系表达是否正确；⑥各数据项的完整性。

(3) 常用的质量控制方法

常用的空间数据库质量控制内容及对应方法见表5-11。表中带圆圈的数字分别对应于上面的各种空间和属性数据质量检测内容。

表5-11 各种质量控制内容与方法选择

控制内容 \ 控制方法		图件	有效值	频率	包含	统计	匹配检查	程序	报告检查
空间数据	①	√							
	②	√							
	③	√							
	④							√	√
	⑤								√
	⑥	√							
	⑦	√							
属性数据	①						√	√	
	②			√					√
	③		√	√	√	√			
	④							√	√
	⑤								√
	⑥	√		√		√		√	

(据陈俊等，1999，略有修改)

其中，图件法是采用与原始图一致的比例尺和注记、符号，将空间和属性数据以制图的方式绘在透明纸上，再与原始图件进行重叠比较的一种方法。图件法使用较为广泛，是检查空间数据位置精度的有效方法，但由于最后必须通过人工方式逐一进行核查，费时又费力。包含法、有效值法、频率法、统计法、匹配法、程序法和报告法等，可以使用程序进行自动质量检查，但都有一定的应用范围，需要根据实际情况来选择。包含法用来检查连续型属性数据项的值是否在一定的范围内，例如金矿床储量的值域通常应该是在1000吨以下。有效值法通常用于离散型属性数据的有效性检查，可以是定性的也可以是定量的。例如，如果属性数据是贵金属矿种，它的有效值应当是金、银，等；如果属性数据是贵金属的密度，那么有效值通常应该在1.0 g/cm^3 ～20.0 g/cm^3 或1.0 t/m^3 ～20.0 t/m^3 之

间。频率法主要用于冗余数据和主关键字、外部关键字重复性检查。统计法主要使用常用的统计方法，例如均值、方差、最大值、最小值、中值等来检查属性数据的内容。匹配法用于检查数据表的定义是否正确，是否符合数据库设计的要求。它将标准的数据表定义与实际的数据表定义写成同一种数据格式，例如文本数据格式，然后使用程序来进行比较，以便确定数据库结构的完整性。程序法是一种很灵活的质量检测方法，可以或多或少地被应用于任何一类质量检查，但要求实施者会使用某一种或多种程序语言。报告法用于给用户提供的质量检查报告中，其内容可以根据不同项目的具体要求而定。

目前各种商品化的 GIS 软件均提供基本的质量控制功能，设计者可以根据自己的要求进行选择、裁剪或二次开发，然后对需要进行质量检测的数据进行检测定义。这种质量控制工具通常以一种标准方式制定出来，能对整个数据库进行自动检测。

4. 开发进度的监测

地矿资源空间数据库常常需要多人参与开发，如何监测整个开发速度便是一个很重要的环节。目前较为简单而常用的开发的监测方法是进度表法和区域图表法。

(1) 进度表法

进度表法（表 5-12）是将各进度的步骤详细列出，制定出每步骤的起始和终止日期。这种方法可以从表中一览整个数据库开发的进度安排。

表 5-12　进度表法示意

日期 / 内容 / 阶段	预处理		数字化		属性数字化		制图		拼接		QC		…	
	起	终	起	终	起	终	起	终	起	终	起	终		
1	6.1	6.6	6.6	6.16										
2														
3														

（据陈俊等，1999，略有修改）

(2) 区域图表法

区域图表法是目前较流行的一种数据库开发进度监督的办法，它是将整个的工作区按格网分成小区后，以各小区作为工作的基本单元来安排进度（图 5-18）。对于整个过程的各个阶段使用不同的图案来表达，这样整个数据库开发实施的目前状况从图中便一目了然，而且可以通过对方格的数量占总方格的百分比来计算出整个项目的完成状况。

与进度表法相比，该法更适用于数据库实施以后进度完成状况的监督，而前者则更适用于进度的计划。两者若同时使用，便可相辅相成、相得益彰。

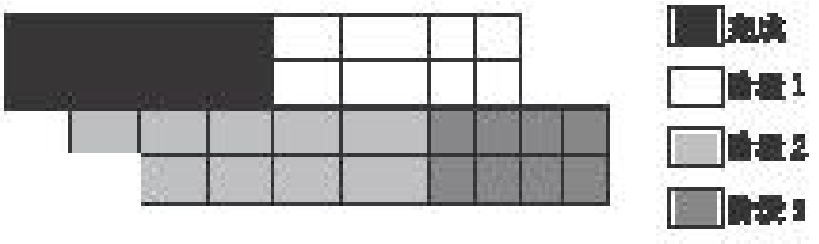


图 5-18　开发进度的区域图表法示例
（据陈俊等，1999）

除此以外，在整个自动化实施过程中，有必要对下面几个方面加以注意：①每个层的自动化方案和实施的细节过程图；②人员需求和分配；③人员的发展和评价；④硬件和软件等系统方面的准备；⑤开发的时间和经费的预算；⑥管理方面的要求等。

5. 数据库实施

数据库的概念设计和详细设计修改以后，便可以开始正式进行数据库实施了。实施的过程，应当以实施计划为指南，尽量按照计划进行实施。但是再好的计划也不可能是完全准确的，在实施过程中常常需要对实施计划做或多或少的改动。任何方面的改动都应当以书面形式备案，做到有案可查。

测试项目的顺利完成是整个数据库实施的必要保障。数据库的真正实施要求按照计划好的过程按部就班地进行。如果前面的各个过程，包括数据库的概念设计、详细设计、自动化实施计划等都能够认真地完成，这一步的实施通常是一个项目的总监督过程。

第三节　地矿资源空间数据编辑、处理与空间分析

围绕内置的空间数据库和属性数据库，空间信息子系统配备了一系列功能强劲的应用软件，其中包括空间数据编辑、提取、转换、处理与空间分析。这些应用软件使得地矿勘查与开发人员能够以一种全新的、综合的方法，来解决与地质体或矿产资源的形成、演化、预测、评价及其合理利用等方面的各种空间信息分析问题，在实践中产生了巨大的效益。

一、地矿资源空间数据的编辑

空间数据编辑是对各种来源的空间数据进行规范和纠错，或者根据特定的专题处理要求将数据重新组织和格式化的过程。无论用何种方法获取的原始数据，都可能存在着采集过程引入错误、与使用格式不相适应、比例尺和投影不统一、图幅之间不匹配、数据冗余等问题，必须通过对原始数据的提取、转换等处理，来使存储的数据净化并符合规范化标准。

1. 数据采集过程引入错误的去除

(1) 数据编辑的查错方式与过程

空间数据采集过程引入的错误包括：空间点位和线段的丢失或重复、线段过长或过短、区域中心识别码的遗漏，以及结点代码和区域属性码不符合拓扑一致性要求（即出现拓扑异常），等等。空间数据编辑的除错方式有目视检查、机器检查和图形检查三种（图 5-19）。目视检查用以检查一些显著的错误，例如线段的丢失、线段过长等；机器检查用以进行地矿空间数据拓扑一致性的逻辑检验、将弧段连接成多边形，以及进行数据容差的检查等；图形检查用于检查数字化结果与原图的吻合程度，可利用透光桌将数字化文件输出的地图覆盖在原始地图上，再进行逐一比较。必要时，可以将两种图形放大，然后再进行比较。在编辑阶段，当需要将几幅地矿图件合并为较大的数据库时，也应当按要求进行地矿资源空间数据的编辑处理，其中主要是边沿匹配（详见后述）。为保证边沿匹配的顺利进行，对并入数据库的数字文件的所有地质体都要进行认真的检查；而当进行数据合并时，有时还需要对数字文件重新进行格式化。其具体过程见图 5-20。

(2) 拓扑关系建立与校正

拓扑关系是空间实体之间的几何关系。建立拓扑关系是 GIS 与一般制图系统的主要区别之一。但在进行数据采集时，为了避免造成多边形共用边的重复输入与拼合不准，我们并不单独输入描述这个矿体的多边形，一般只将节点与弧段数字化输入，然后在图形编辑时交互式生成多边形拓扑结构或全自动生成拓扑结构，进而生成线状或面状地物。

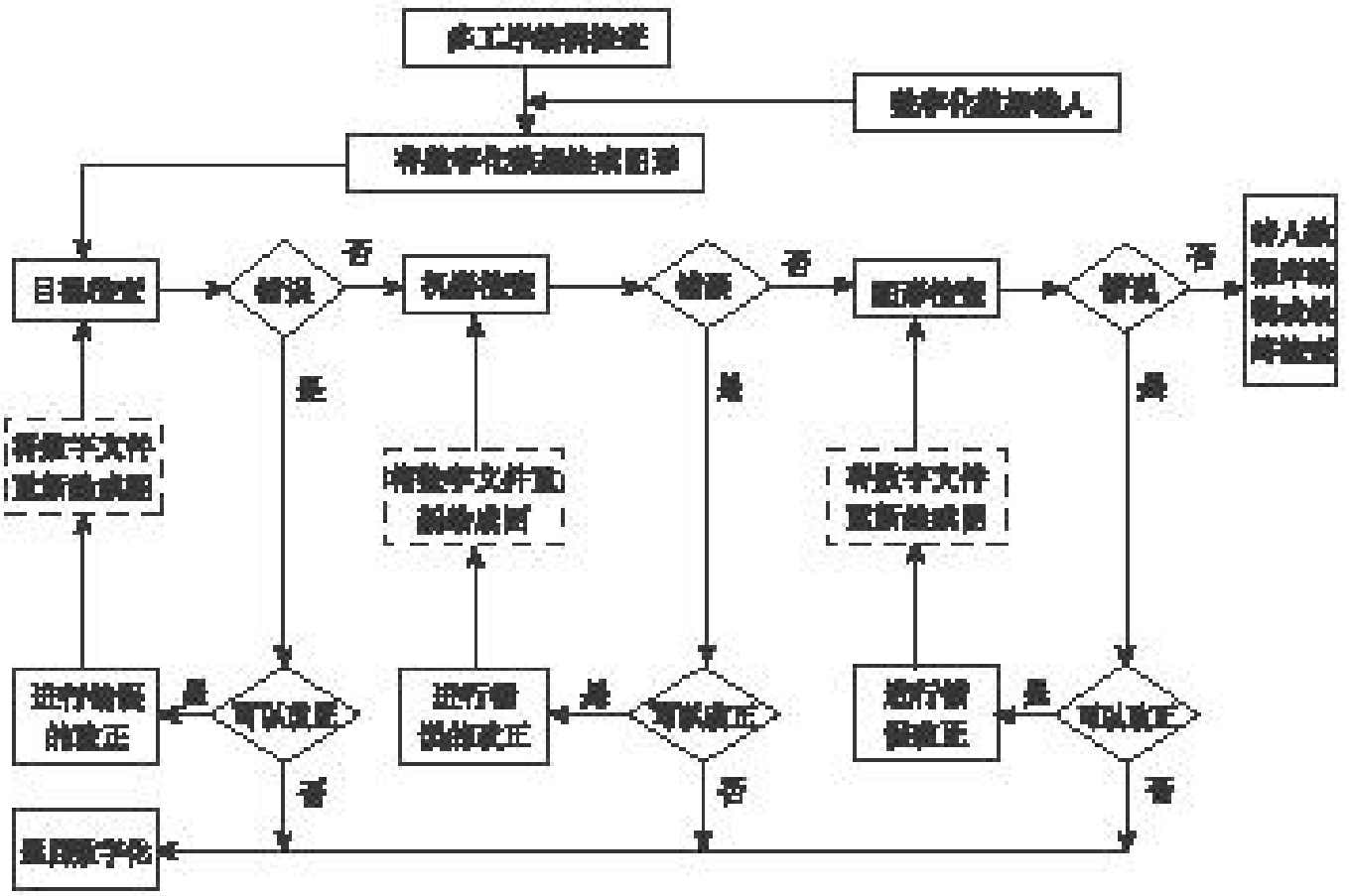


图 5-19　多工序编辑检查地矿资源空间数据的过程

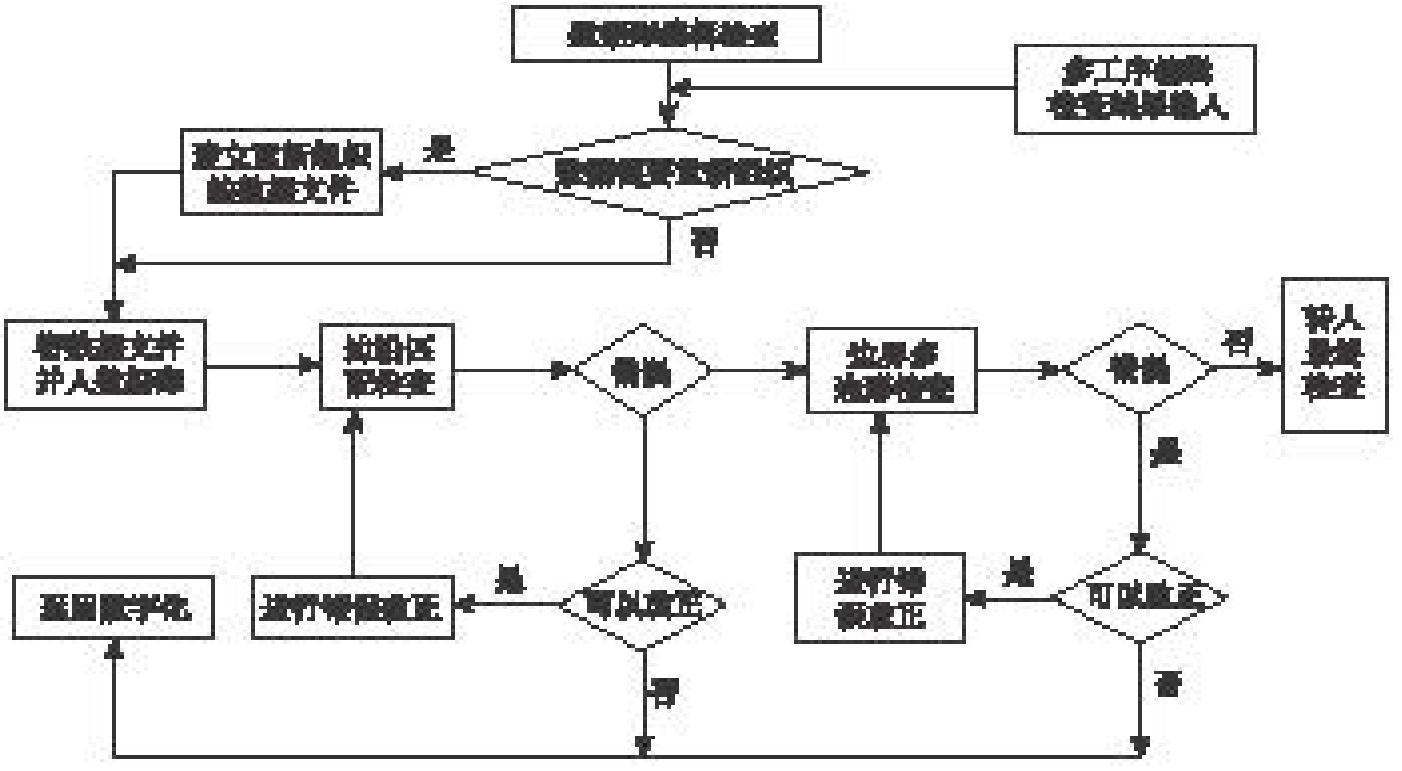


图 5-20　地矿资源空间数据库编辑检查

交互式生成多边形拓扑结构是采用鼠标人工装配方式进行的，它的操作复杂，工作量大，但是对原始数据要求不严，修改时不必重复计算。全自动生成拓扑结构是采用纯软件驱动方式进行的，生成速度快，不需要人工干预，但对原始数据要求严。

由于图形输入时的差错，经常出现多边形不封闭或边的重复输入等问题，造成拓扑结

构的混乱。编辑时必须对所输入的空间图形进行拓扑结构合理性检查。系统可以自动显示输入图形中的悬挂线、交叉边、重复边等，提示应用人员进行校正，有的系统还能自动进行节点平差，自动分割交叉边，自动将线段端点延伸到接点上，或将线段超过接点的部分自动切除，通过这些对图形进行必要的处理，减轻下一步图形编辑的负担。

除了上述多边形的拓扑关系外，还可通过图形编辑建立几何图形之间的组合关系。

(3) 图形编辑与纠错

图形编辑包括图形参数编辑与图形几何数据编辑。

图形参数用于表征图形元素。在 GIS 中，与地矿资源有关的图形元素分为点、线、面三类。其中点元素包括图形符号、数字、文字、圆、椭圆等；线元素包括河流、海岸线、交通线、行政区边界、地层界线和断层等等；面元素包括水体、岩体、盆地、构造单元、地层单元、行政区、矿区，等等。表征每种图形元素都有相应的图形参数，例如一个用于标注地质体的图形符号，它的图形参数包括符号代码、符号高、符号宽、旋转角、颜色、层号、属性代码等；又如一条代表断层的曲线，它的参数包括描述断层性质的线型和线色、描述断层级别的线宽、描述断层赋存状态的层号、属性代码；再如一个表示地层的多边形，它的参数有表征地层岩性、时代等的填充方式、填充图案类型、填充颜色、层号、属性代码等。

地矿资源空间信息子系统为图形参数的编辑和几何数据的修改提供了丰富的功能。在图形编辑中，参数控制功能最为常见，主要用于：①安置数据采集参数，例如节点捕捉容差等；②设置地形图定向参数；③存储一般的数据获取与编辑参数；④设置一些功能开关。

对于所有这些图形参数和几何数据，用户可以利用图形编辑模块进行可视化修改，非常直观方便。修改既可以是单独进行，也可按层、按区域成组地进行。系统允许按层显示地图要素，以便控制图面负担，突出研究的地矿专题信息。每个描述地质体的图形元素都具有层号和属性代码，用于连接该图形元素与属性数据库中相应的属性数据。

(4) 图幅拼接

由于在对地矿图件进行数字化时存在的人工操作误差、机器精度误差、图幅转换误差以及其他因素影响，两个相邻图幅的地图接合处时常出现逻辑裂隙与几何裂隙。所谓的逻辑裂隙是指一个图幅的图形元素的属性或图形参数与另一图幅相应的图形元素不一致，也就是说同一地质体在两个数据库中具有不同的附加信息；几何裂隙是指两个图幅相应的图形元素——数据库边界被分开的一个物体两个部分不能精确地衔接。为此，通常在图形编辑系统中可以定义相邻图幅的接边范围，开窗检索搭接范围内的图形及属性信息，将跨接图幅的目标，根据接边的方向，横或纵向坐标进行排序，对逻辑相同的地物在几何上自然连接，对于属性一致的图形，可自动清除接边误差。对于属性不一致的图形，可使用交互方式将它们拼接在一起，达到逻辑上的一致。

(5) 属性数据编辑

对地矿属性数据的编辑一般是在属性数据处理模块中进行的，但是在图形编辑系统和数据库管理系统中的地矿属性数据应保证在同一数据库中，只是在不同数据处理模块和界面中进行的。将一个地质体的属性数据连接到相应的几何目标上，为了建立属性描述数据与几何图形的联系，通常需要在图形编辑系统中设计对地矿属性数据的编辑功能。在数字化和建立图形拓扑关系的同时或之后，对照一个几何目标，直接输入相应属性数据。一般

来说，属性数据库管理系统比在图形编辑系统中功能更强，编辑系统除了提供插入、删除、修改、拷贝等功能外，还具有移动数据的行和列，合并数据等功能。

当文件或数据库的全部编辑任务完成以后，首先要将这种净化后的数据存入新的存储介质，然后再提供后续对这些数据的提取、转换、分析应用，同时关闭空间数据编辑站。

2．图幅数据的变换和匹配

当使用数字化仪和扫描仪对地矿图件进行图数转换时，由于数字化仪桌面坐标系与地矿图件直角坐标系不一致，或者图纸变形等原因，所获得的数字化坐标（x，y）不可能与地图的真实坐标（x，y）一致。有时，不同来源的地矿图件的地图投影及地理坐标还可能有差异，地图比例尺之间、地图比例尺与数字化仪的长度单位之间也可能不一致。另外，野外现场观测数据的获取，也是根据具体的任务要求，结合现场操作的方便和习惯进行的，在地矿资源空间数据的类型、结构、规范标准及其坐标系、比例尺方面，都会有一定的局限性。为了获得与地图直角坐标一致的数据，也为了使不同地图的投影坐标取得统一，必须对图幅内的图形坐标数据进行变换，包括数字化坐标数据比例尺的变换、变形误差的消除、投影类型的转换，以及坐标的旋转和平移（图 5-21）。

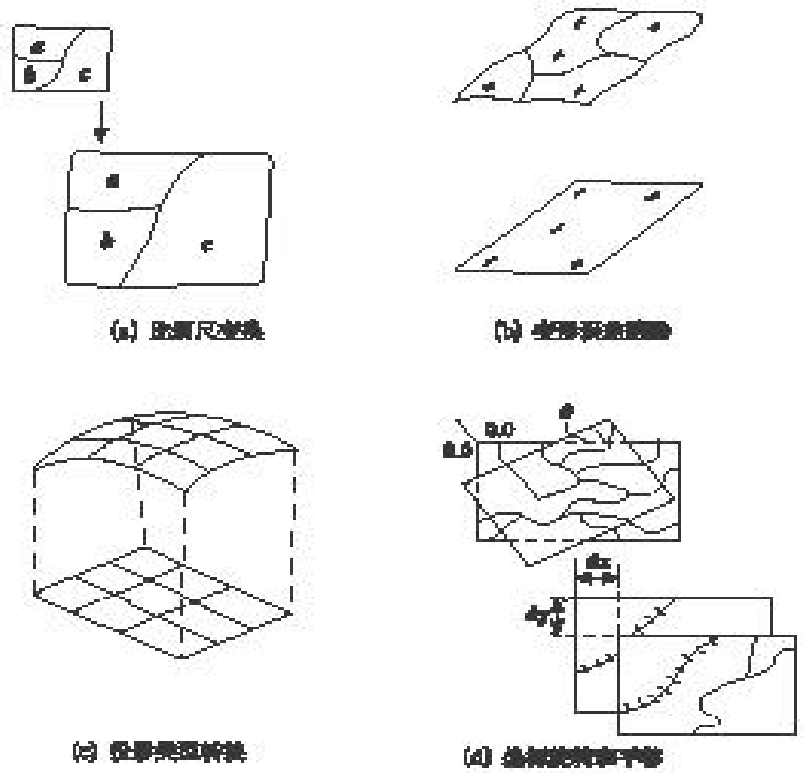


图 5-21　图幅数据的变换和匹配处理

（据 Jack Dangermond，1984）

（1）数字化坐标的变换

如图 5-22，设 XOY 为地质图件直角坐标系，XO′Y′为数字化仪坐标系，两坐标系之间的坐标轴夹角为 α，O′相对于 XOY 坐标系原点的平移距离为 A_0、B_0，地质图实际比例尺为 m，则根据图形学原理，可写出变换公式：

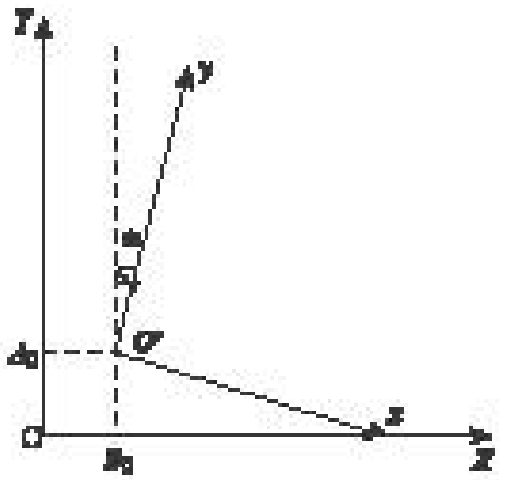


图 5-22　坐标变换原理

$$\begin{aligned} X &= m(x\cos\alpha - y\sin\alpha) + A_0 \\ Y &= m(x\sin\alpha + y\cos\alpha) + B_0 \end{aligned} \tag{1}$$

令 $A_1 = m\cos\alpha$，$B_1 = m\cdot\sin\alpha$

则（1）式可以简化为

$$\begin{aligned} X &= A_0 + A_1 x - B_1 y \\ Y &= B_0 + B_1 x + A_1 y \end{aligned} \tag{2}$$

当已知点数 $n=2$ 时，各个系数有最简单的计算形式：

$$A_1 = \frac{(x_2 - x_1)(X_2 - X_1) + (y_2 - y_1)(Y_2 - Y_1)}{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

$$B_1 = \frac{(x_2 - x_1)(Y_2 - Y_1) - (y_2 - y_1)(X_2 - X_1)}{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

显然，每个数字化坐标（x，y），对应的实际坐标（X，Y），均可由式（2）求得。注意：式（2）是假设地图的实际比例尺在 X 和 Y 方向一致，称为相似变换。实际上，由于地图图纸所标注的方向不同，其 X 和 Y 方向因变形而引起的实际比例尺常常是不相同的。这时，令 m_1 和 m_2 分别表示 X 和 Y 方向的比例尺，则可变换公式（1）为：

$$\begin{aligned} X &= (m_1\cdot\cos\alpha)x - (m_1\cdot\sin\alpha)y + A_0 \\ Y &= (m_2\cdot\sin\alpha)x + (m_2\cdot\cos\alpha)y + B_0 \end{aligned} \tag{3}$$

令 $A_1 = m_1\cos\alpha$，$A_2 = -m_1\sin\alpha$；$B_1 = m_2\sin\alpha$，$B_2 = m_2\cos\alpha$

则式（3）简化为：

$$\begin{aligned} X &= A_0 + A_1 x + A_2 y \\ Y &= B_0 + B_1 x + B_2 y \end{aligned} \tag{4}$$

由此，建立误差方程：

$$Q_x = X - (A_0 + A_1 x + A_2 y)$$

$$Q_y = Y - (B_0 + B_1 x + B_2 y)$$

同理，由 $[Q_x^2]$ 为最小和 $[Q_y^2]$ 为最小的条件，得到两组法方程：

$$\begin{aligned} A_0 n + A_1\Sigma x + A_2\Sigma y &= \Sigma X \\ A_0\Sigma x + A_1\Sigma x^2 + A_2\Sigma xy &= \Sigma xX \\ A_0\Sigma y + A_1\Sigma xy + A_2\Sigma y^2 &= \Sigma yX \end{aligned} \quad \text{和} \quad \begin{aligned} B_0 n + B_1\Sigma x + B_2\Sigma y &= \Sigma Y \\ B_0\Sigma x + B_1\Sigma x^2 + B_2\Sigma xy &= \Sigma xY \\ B_0\Sigma y + B_1\Sigma xy + B_2\Sigma y^2 &= \Sigma yY \end{aligned}$$

求解法方程，得系数值 A_0、A_1、A_2 和 B_0、B_1、B_2。

当已知点数 $n=3$ 时，可以根据以下线性方程组

$$\begin{bmatrix} X_1 \\ Y_1 \\ X_2 \\ Y_2 \\ X_3 \\ Y_3 \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 0 & 0 & 1 & 0 \\ 0 & 0 & x_1 & y_1 & 0 & 1 \\ x_2 & y_2 & 0 & 0 & 1 & 0 \\ 0 & 0 & x_2 & y_2 & 0 & 1 \\ x_3 & y_3 & 0 & 0 & 1 & 0 \\ 0 & 0 & x_3 & y_3 & 0 & 1 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ B_1 \\ B_2 \\ A_0 \\ B_0 \end{bmatrix}$$

求解各个系数。于是，各个数字化坐标（x，y），由式（4）可以求得对应的实际坐标

(X, Y)。该变换公式考虑到在 X 和 Y 方向上因变形而引起实际比例尺的不同，因此称为仿射变换。

(2) 地图投影的变换

当系统使用的地矿资源空间数据取自不同地图投影的图幅时，需要将一种投影的数字化数据变换为所需要投影的坐标数据，以便实现不同比例尺地图的合成，以及跨分度带的地图的拼接。系统中有三种投影变换的方法模块，可供选择使用。

正解变换　通过建立一种投影变换为另一种投影的严密或近似的解析关系式，直接由一种投影的数字化坐标 (x, y) 变换到另一种投影的直角坐标 (X, Y)。

反解变换　由某种投影坐标反解出地理坐标 (x, y→B, L)，再将结果代入另一种投影的坐标公式中 (B, L→X, Y)，实现由一种投影坐标到另一种投影坐标的变换 (x, y→X, Y)。

数值变换　根据两种投影在变换区内的若干同名数字化点，采用插值法，或有限差分法，或有限元法，或待定系数法等，从而实现由一种投影坐标到另一种投影坐标的变换。

(3) 图形的缩放与旋转变换

在进行图形分析和编辑时，为了观察和修改的方便，人们常常需要将图形平移、放大、缩小，甚至旋转到某一个新的角度。在 GIS 中，图形的平移、放大、缩小和旋转，是通过空间数据的拓扑运算来实现的。

(4) 图幅数据边缘匹配

为了将图幅的数据连接在一起，组成连续统一的图幅数据，以便加入大型数据库或输出拼接后的地矿图形 (图 5-23)，必须消除相邻图幅的边缘部分由于种种原因造成的误差，包括同一地质体的线段或弧的坐标数据不衔接、坐标系统不一致和编码方法不一致，等等。边缘匹配处理的任务包括以下几个方面:

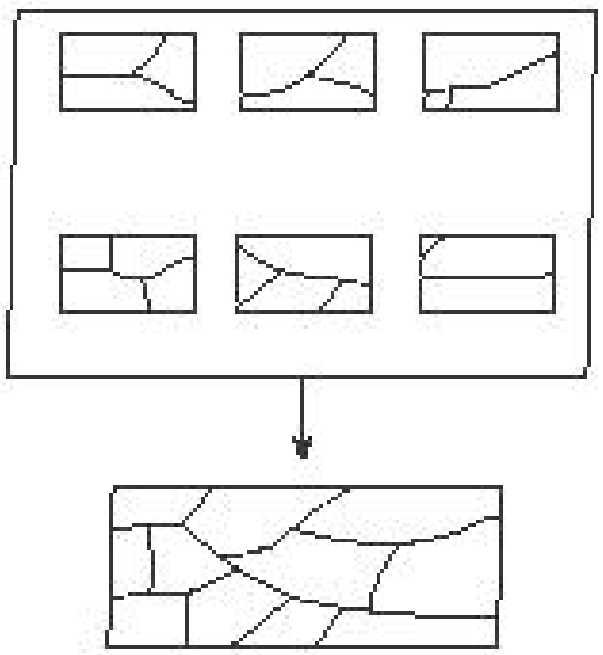


图 5-23　图幅数据边缘匹配示意图

识别和检索相邻地矿图幅的数据　相邻两幅图的数据识别，可以通过图幅编号与分幅数字化的数据联系起来的方法来解决。如图 5-24 所示，各图幅中的十位数指示图幅编号的横向顺序，个位数指示纵向顺序。当进行横向图幅拼接时，主要目标是十位数编号相同的图幅；当进行纵向图幅拼接时，主要目标则是个位数编号相同的图幅。图幅数据的边缘匹配处理主要是针对跨越相邻图幅的线段或弧而言的，为了减少数据容量，提高处理速度，一般只检索并提取地矿图幅边界 2 厘米范围内 (如图 5-24 中编号为 22 的斜线部分) 的数据来建立匹配处理文件。

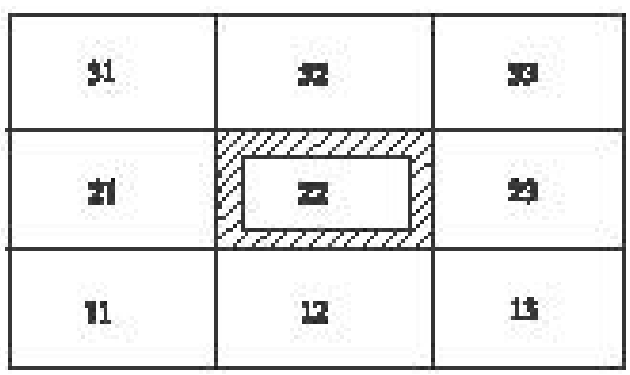


图 5-24　图幅编号及图幅边缘数据的提取

相邻图幅边界点坐标数据的匹配　相邻图幅边界点坐标数据的匹配采用追踪拼接法。

追踪拼接时可采用向右追踪、向左追踪、向上追踪和向下追踪等四种方式（图 5-25）。一般地说，只要符合下述两个条件，两条线段或弧即可匹配衔接：①相邻图幅边界两条线段或弧的左右码各自相同或相反；②相邻图幅同名边界点坐标在某一许可值的范围内（例如≤ ±0.5 mm）。匹配衔接时以一条弧或线段链作为处理单元，当边界点位于两个结点之间时，须分别取出相关的两个结点，然后按照结点间线段链方向一致性原则进行记录和存储。例如，图 5-26（b）就是根据图 5-26（a）的原始数据，按照跟踪拼接法处理后输出的图形。

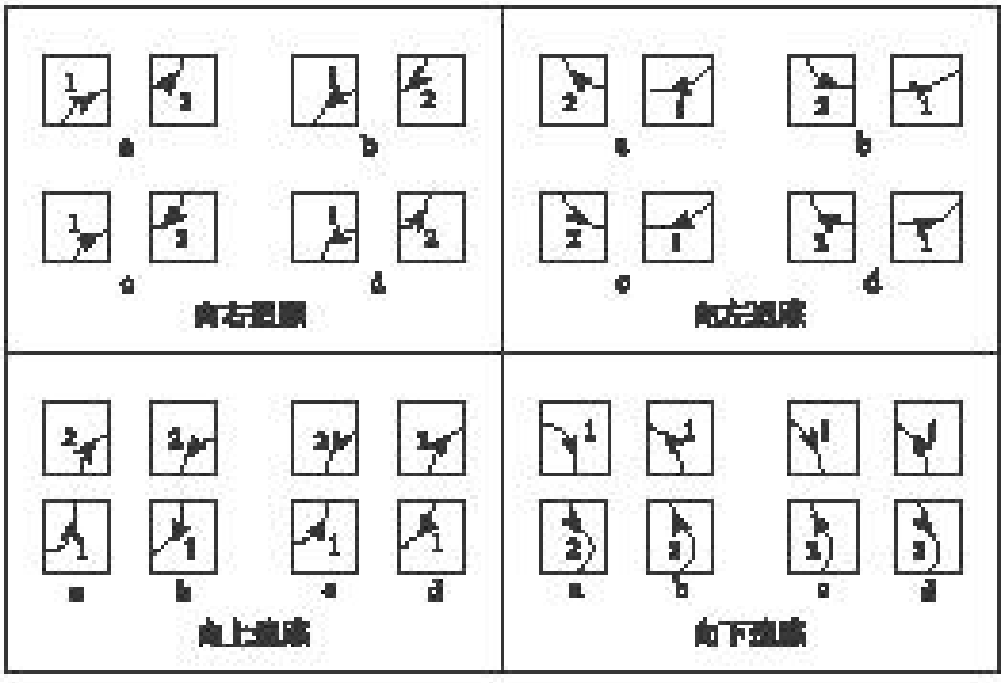


图 5-25　追踪拼接法

（据胡友元等，1988）

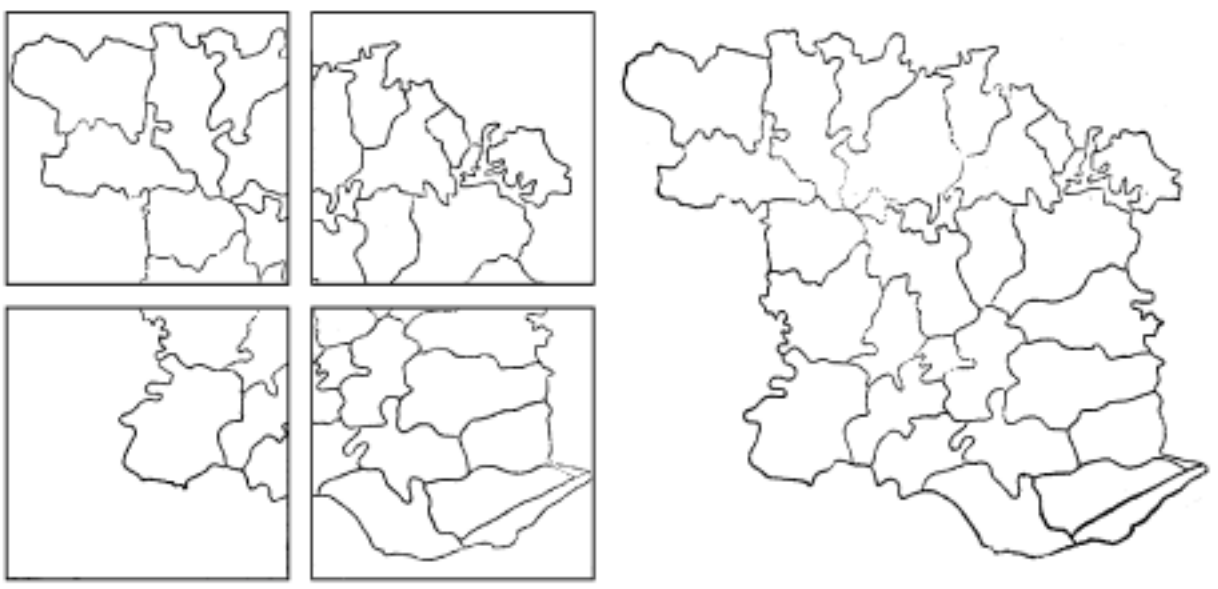


图 5-26　图幅的边沿匹配示意

（据胡友元等，1988）

相同属性多边形公共界线的删除　当图幅数据通过边缘匹配处理，组成较大区域的连续图幅的数据以后，常常伴随着比例尺的缩小和属性类型的合并。相同属性的多边形合并之后，原先各自的边界封闭线就成为多余的了，需要将其统统删除掉。地矿资源空间信息子系统提供了这种删除功能，用户可以很方便地加以利用。

3．图形整饰

为了制作正规的成果图，必须对图面进行整饰。系统提供丰富的图形工具，可以自动生成标准图廓、经纬线或方里网，有的还能让用户方便地制作拼图表、比例尺、指北针以及图例、图注等，甚至有的软件还具有简单的文字排版功能。

地矿资源图件的线型、符号种类繁多，需要加以规范化、标准化。但为了补充实际应用的需要，地矿资源空间信息子系统通常都提供自定义功能，让用户根据需要制作一些特殊的图形符号、特殊线型、填充花纹和调色板。用户可以用鼠标直接描绘，也可以通过组合原有的符号来形成新符号，然后存入符号库中，供图件整饰时使用。

二、地矿资源空间数据类型的转换

在空间信息子系统中，地矿资源空间数据的类型有矢量数据和栅格数据两种。矢量数据和栅格数据的混合处理和应用，已成为当前系统发展的主要趋势。而对矢量、栅格和地矿专业数据的共同管理和处理，也已成为应用型 GIS 设计的基本标准。这种共同管理和处理的主要内容，涉及矢量数据与栅格数据之间的互相转换、表示和运算。

1．矢量数据的栅格化

矢量数据向栅格数据转换处理的根本任务是通过一个有限的工作存储区，使得矢量和栅格数据之间不可避免的读写操作限制在最短的时间范围内。根据转换处理时，基于地质体弧段数据文件和多边形数据文件的不同，分别采用不同的算法。

(1) 弧段数据的栅格化方法

弧段数据栅格化的算法可以分解为数据分割和转换计算两步。

数据分割　数据分割的任务是按照可用的工作存储区，将矢量数据体划分成数据段或栅格带（图 5-27），并建立其相应的矢量数据文件。具体方法是：对整个原始矢量数据文件进行一次性扫描，计算原点（IX_0，IY_0）、弧段总数 NA、栅格行数 NR、列数 NC、分带数 NW，再将各个弧段按照它们在纵轴上的位置归入相应的子数据体。用 max（I）表示

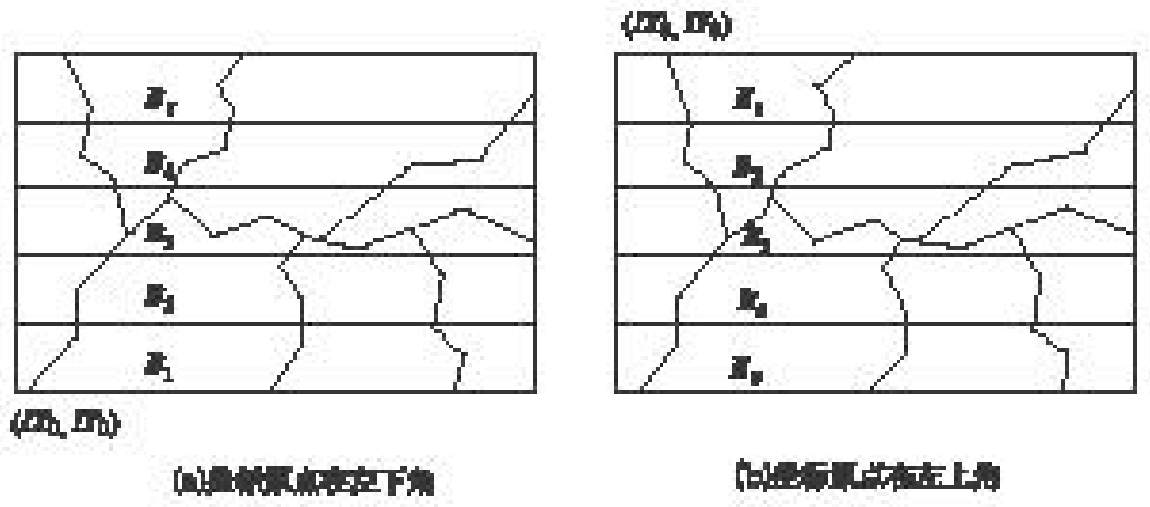


图 5-27　栅格矩阵按不同坐标系分带示意图

该弧段归入的栅格带最大编号，min（I）表示该弧段归入的栅格带最小编号；I为弧段的识别码。在实现了矢量数据的分割之后，便可以按照规划的栅格带，进行矢量数据向栅格数据的转换计算。

转换计算　转换计算的任务是将任意的X、Y坐标转换为由行号（I）和列号（J）表示的栅格数据。具体转换的方法采用扫描线与有关的弧段求交，然后将交点的X坐标、该弧段对应的左右区码记入该扫描线所在行的数组，最后通过对一行所有X值按照由小至大排序、进行左右区码的配对和施行奇点的处理等技术，在相邻X值之间，逐段生成栅格数据，直到全部扫描行都完成从矢量向栅格的转换为止。

（2）多边形数据的栅格化方法

这种方法的要领是以多边形作为栅格化的处理单元，首先将一个多边形的x坐标数据，分别按照顺时针和逆时针的方向排序，即

$$x_1, x_2, \cdots, x_{n-1}, x_n \text{ 和 } x_n, x_{n-1}, \cdots, x_2, x_1$$

对任一格网点P，设其坐标为X_p、Y_p，则根据该点的x值在上述两组坐标序列中相应区间出现的次数，可以求出相应个数的y值。设y_a对应于顺时针排列的x区间的内插值，y_b对应于逆时针排列的x区间的内插值，则

$$y_a = y_i + \frac{y_i - y_{i-1}}{x_i - x_{i-1}} \times (x_p - x_{i-1})$$

$$y_b = y_j + \frac{y_j - y_{j-1}}{x_j - x_{j+1}} \times (x_p - x_{j+1})$$

设A_1，A_2和A_3为中间变量，而且A_3为A_1和A_2之和，A_1和A_2的取值由下式决定

$$A_1 = \begin{cases} 1 & \text{当 } y_a \geq y_p \text{ 时} \\ -1 & \text{当 } y_a < y_p \text{ 时} \end{cases} \qquad A_2 = \begin{cases} 1 & \text{当 } y_b < y_p \text{ 时} \\ -1 & \text{当 } y_b \geq y_p \text{ 时} \end{cases}$$

如果$|A_3| \geq 2$，该格网点P位于该多边形之内，将该多边形的属性值赋予该格网。

按照这种方法逐点处理，直到全部多边形生成栅格数据或栅格游程编码数据，提供系统分析和使用。

2．栅格数据的矢量化

栅格数据矢量化的目的，首先是为了能将自动扫描仪获取的栅格数据加入矢量数据库，其次是为了将栅格数据分析的结果，通过矢量绘图装置输出，此外，为了数据压缩的需要，也常将大量的面状栅格数据转换为由少量数据表示的多边形边界。在进行转换处理时，须根据图像数据文件和再生栅格数据文件的差异，分别采用不同的算法。

（1）图像数据的矢量化

图像数据是指通过自动扫描仪，按一定的分辨率对不同灰阶和色调的影像或线划进行扫描采样，而得到的以不同灰度值（0～255）表示的数据（图5-28）。目前扫描仪的分辨率可达0.0125 mm，因此对一般粗细（例如0.1 mm）的线条，其横断面扫描后平均也有8个像元，而矢量化的要求只能允许横断面保持一个栅格的宽度，因此在将栅格数据转换成矢量数据时，需要进行一些处理（图5-29），例如二值化处理、细化处理和跟踪处理。

二值化处理　线划图形经扫描后所产生栅格数据，是按从0～255的不同灰度值量度的，设以G（i，j）表示。为了将这种256或128级不同的灰阶压缩到2个灰阶，即0和1两级，首先要在最大和最小灰阶之间定义一个阀值，设为T，则根据下式可得到二值图:

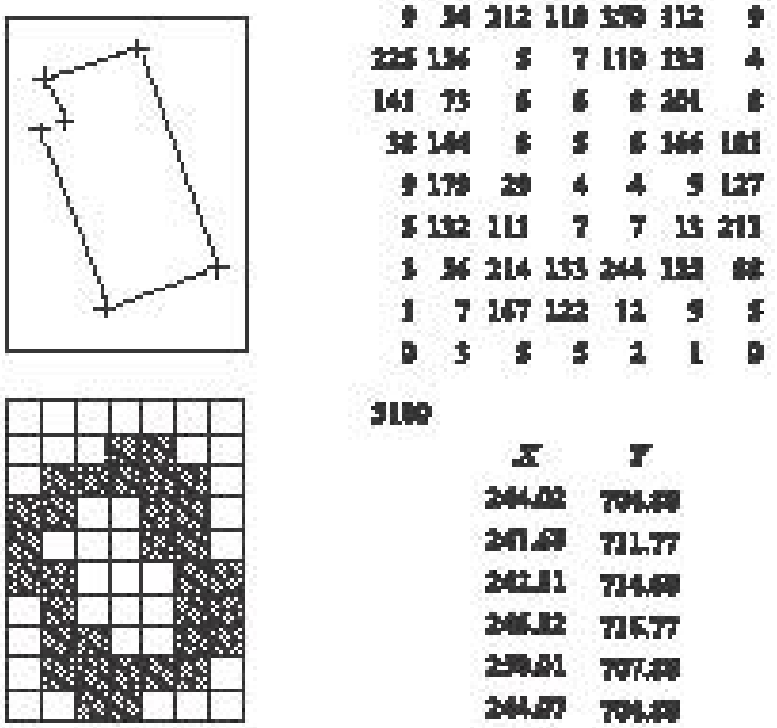


图 5-28　栅格与矢量数据表示法

(据 W. Lichiner 等，1987)

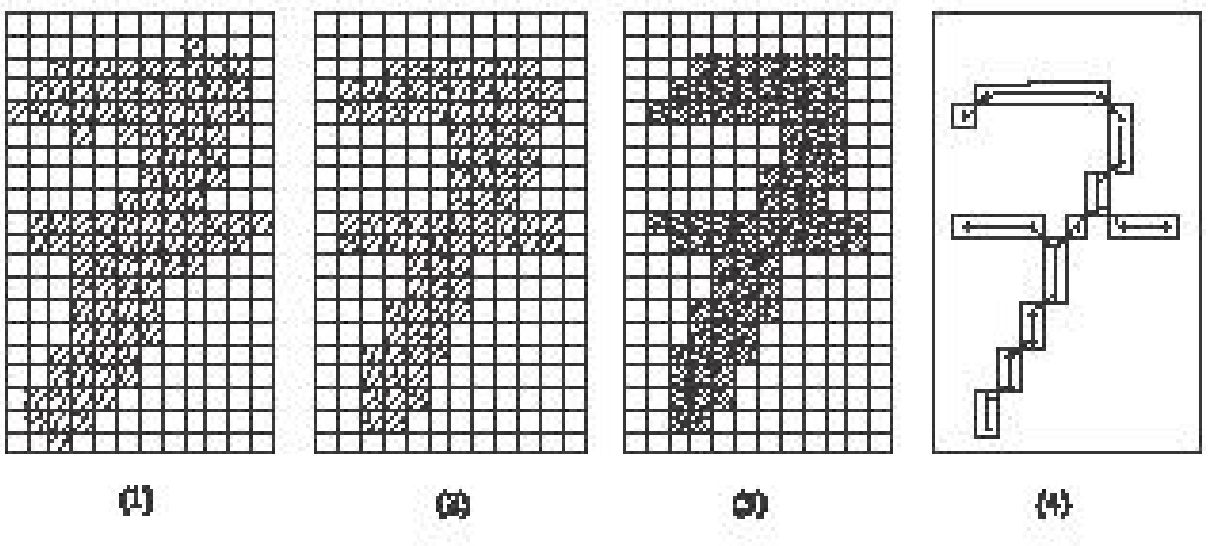


图 5-29　栅格-矢量转换过程

(据 W. Lichiner 等，1987)

$$B(i,j) = \begin{cases} 1 & \text{如果 } G(i,j) \geq T \\ 0 & \text{如果 } G(i,j) < T \end{cases}$$

细化处理　细化处理用于消除线划横断面栅格数的差异，使得每一条线只保留代表其轴线或周围轮廓线（对面状符号而言）位置的单个栅格的宽度。栅格线划的“细化”方法，可分为“剥皮法”和“骨架化”两大类。剥皮法的实质是从曲线的边缘开始，每次剥掉等于一个栅格的一层，直到最后留下彼此连通的由单个栅格点组成的图形。由于同一条线在不同位置可能有不同的宽度，在剥皮过程中不允许剥去会导致曲线不连通的栅格。其解决办法是，借助一个在计算机中存储着的、由待剥栅格为中心的 3× 3 栅格组合图来

决定。如图 5-30 所示，一个 3× 3 的栅格窗口，其中心栅格有 8 个邻域，因此组合图共有 2^8 种不同的排列格式。若将相对位置关系的差异只是转置 90、180、270 或互为镜像反射的方法进行归并，则共有 51 种排列格式。显然，其中只有格式 2、3、4、5、10、11、12、16、21、24、28、33、34、35、38、42、43、46 和 50 可以将中心点剥去。这样，最多核查 256× 8 个栅格，便可确定中间栅格点应该保留或应该删除。将应保留的栅格系列写入数据文件之后，细化处理过程便告结束。

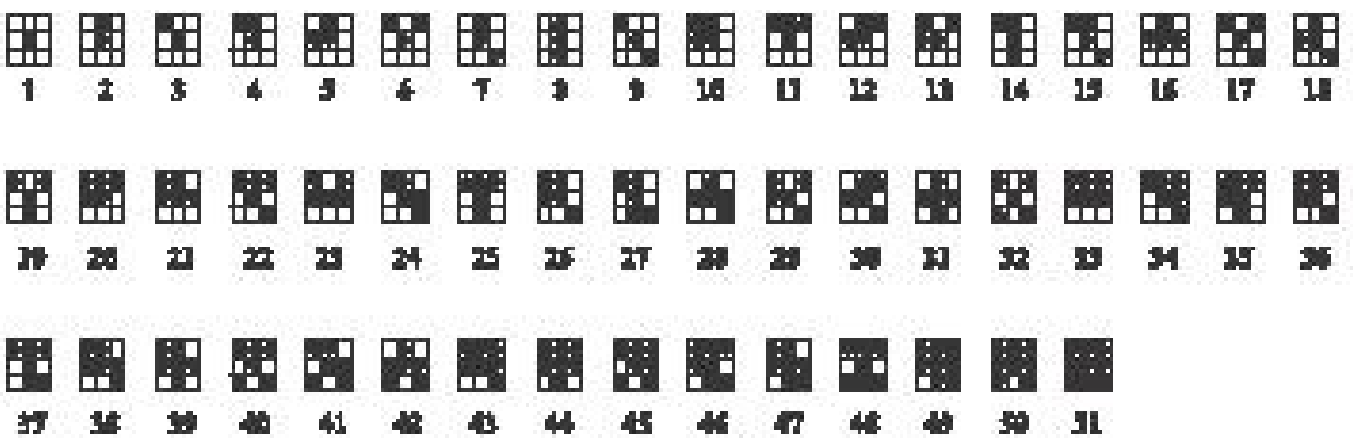


图 5-30　3× 3 栅格组合图

(据 W. Lichiner 等，1987)

跟踪处理　跟踪处理的目的是将经细化处理后写入数据文件的栅格数据，整理为从结点出发的线段或闭合的线条，并以矢量形式存储于特征栅格点中心的坐标。跟踪应当从图幅西北角开始，按顺时针或逆时针方向，对八个邻域进行搜索，依次跟踪相邻点，并记录节点坐标，然后搜索封闭曲线，直到完成全部栅格数据的矢量化，写入矢量数据库。

在本系统内，栅格数据的二值化、细化和跟踪处理，可以采用人机交互方式完成，也可以采用全自动方式完成。

(2) 再生栅格数据的矢量化

再生栅格数据是指根据弧段数据或多边形数据生成的栅格数据。再生栅格数据的矢量化目的，只是为了通过矢量绘图装置输出高质量的图件。这种数据除了与图像数据相匹配者需要进入数据库之外，一般只用于空间分析，不必作为永久文件保存。

三、空间数据的提取和空间分析

空间数据提取是指借助各种检索手段将存储在空间数据库中的数据搜索出来，并以屏幕显示、转存或打印形式加以表达的一种操作。空间分析则是指对所提取的空间数据进行拓扑运算、属性分析以及拓扑与属性联合分析，具体地说，包括地形分析（DTM 和 DEM)、空间特征的几何分析、网络分析、叠置分析、缓冲区分析等。

1. 空间数据的提取

空间数据的提取操作贯穿于数据分析、数据处理和应用的全过程，是地矿资源空间数据库管理系统的一种经常而重要的工作形式。

(1) 空间数据的提取方式

空间数据提取包括基本提取、空间检索和布尔处理等三种方式。

空间检索专指按照范围限定条件找出所需空间数据的操作，而基本提取是指按照一般

属性或常规窗口找出所需空间数据的操作。空间检索的主要内容包括：①相邻分析检索，指采用相邻分析的技术方法，检索出与某一已知空间实体有邻接关系的空间实体；②相关分析检索，指凡是把与查询窗有关的点、线、面状空间实体都检索出来，而不论它们是整个地或局部地落在这个窗内；③叠置分析检索，指利用多边形叠置分析技术，来检索落在窗口界线范围内的空间实体，而对落在窗口外的部分则剪裁掉；④边沿匹配检索，指应用边沿匹配处理技术，在多幅地图的数据文件之间建立跨越图幅边界的多边形，并提取与查询窗关联的图幅数据，再将这些数据自动地组织进连续的窗口内。例如，在查询诸如“窗口 A（一个用户自行确定的多边形）范围内所有地矿的分布时”需要使用到上述的查询检索技术。具体检索要求如图 5-31 所示。

布尔处理是指根据用户规定的布尔标准或条件表达式所进行的限定性的多重属性的查询和数量统计（图 5-32），例如“检索图斑面积大于 0.25 平方公里的所有金矿分布区”，这是属于双重属性的布尔数据提取处理。多数 GIS 都提供如 SQL 语言的信息查询提取方法。但要在常用的数据库查询语言的基础上支持地矿空间信息查询提取，例如提取穿过一个矿床的断层、某断层两侧 2 公里内的所有矿点等，还需进行必要的扩充或重新设计。

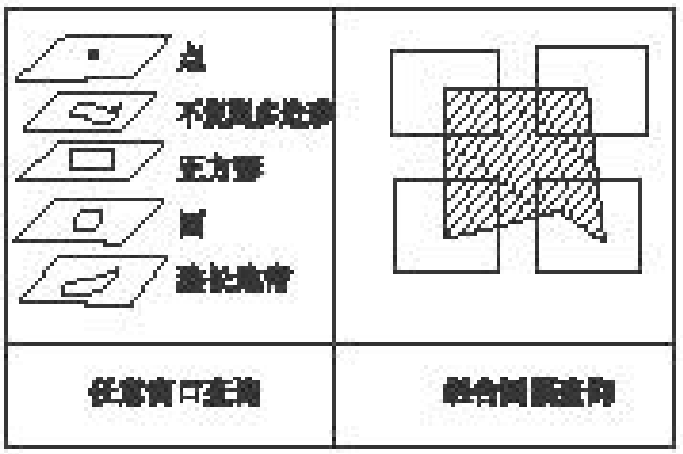


图 5-31　数据的查询检索
（据 Jack Dangermond，1984）

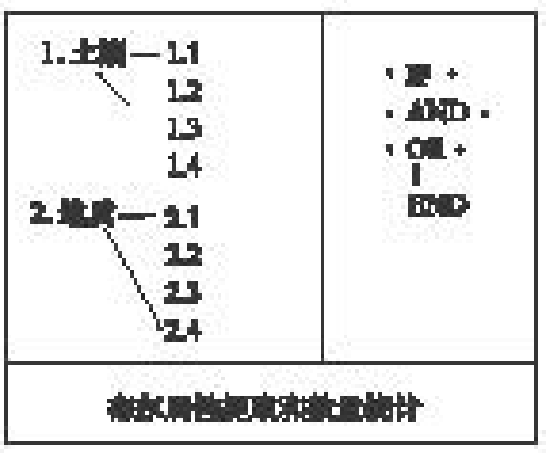


图 5-32　数据的布尔处理
（据 Jack Dangermond，1984）

（2）拓扑空间的提取

空间目标之间的拓扑关系有两种，一种是几何元素的节点、弧段和面块之间的关联关系，用来描述几何元素之间的拓扑数据结构；另一种是 GIS 中地质体之间的空间拓扑关系，一般通过关联关系和位置关系来隐含表达，并通过专门途径进行提取：①面与面关系，例如检索与某个面状地质体相邻的所有地质体（图 5-33（a））；②线与线关系，例如检索所有与某一主断层关联的次级断层（图 5-33（b））；③点与点关系，例如检索到某矿点的距离小于 2 公里的所有点状地质体（图 5-33（c））；④线与面关系，例如检索某断层所经过的所有矿床（图 5-33（d）），或检索某矿区内所有断层（图 5-33（e））；⑤点与线关系，例如检索某河流上

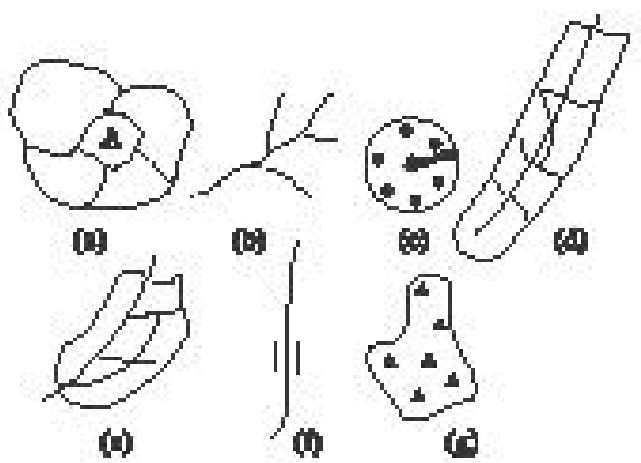


图 5-33　拓扑关系查询

的桥梁（图 5-33（f））；⑥点与面关系，例如检索某含矿地层所有的矿床分布点（图 5-33（g））。

（3）分层检索提取

如前所述，GIS 中的地矿属性数据与空间数据都是分层存放的。例如，矿产地质图的信息分为成矿有利地层、含矿岩体、断层构造线、矿化点等多个层次。如果我们要提取与成矿有关的断层构造线，只需要检索断层构造线这一层数据即可。

（4）定位检索提取

系统允许进行地质体的图形和属性的双向交叉提取。用户可以用光标指点屏幕上的图形，提取它的相关地矿属性数据；也可用光标指点数据表格中的数据项，在屏幕上显示该数据项所关联的地质体图形。

（5）区域检索提取图

可在屏幕上指定一个矩形或任意多边形，提取落在该区域内的所有地质体以及与它们相关的属性。也可提取距离某个点、线、面一定范围内的所有地质体及其相关属性。

（6）条件检索提取

可用数据项与运算符组成的条件表达式检索地质体与属性。这些运算符号可以是 +、-、×、≥、≤、>、=、< 等。例如，给定条件（面积 > 100）&&（地层代号 = = “Pt_2”），我们可以检索出满足面积大于 100，而且地层代号为 Pt_2 的所有相应地质体。

2. 空间分析

空间信息子系统拥有强大的空间分析功能，为地矿勘查和开发人员提供了多种分析问题和解决问题的有效手段。在实际操作中，空间运算、查询和分析往往是同时进行的，相互之间没有严格的区别。

（1）缓冲区分析

缓冲区分析是根据应用目的，在空间数据库的基础上自动建立有关点、线、面实体周围一定范围的多边形缓冲区，用于表达某种因素的影响范围（图 5-36）。例如，在断层或构造线两侧划出缓冲区，分析该缓冲区范围内所存在的矿点类型及其分布密度，可以判断在离断层或构造线多远的距离内成矿可能性最大；在边坡有滑塌危险矿坑，需要根据观测结果划出一定的缓冲区，作为警戒的指示。缓冲区分析涉及不同地理特征的空间接近度或临近性问题，也涉及所形成的新多边形数据层的存储和管理问题。

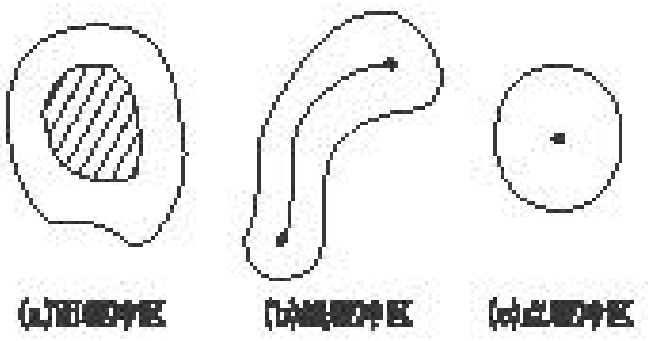


图 5-34　缓冲区分

根据研究目标的空间特点，缓冲区有单点、单线和单面缓冲区（图 5-34）和多点、多线和多面缓冲区之分。GIS 通常都可以自动处理两个目标的缓冲区重叠的情况，并能够自动取消由于重叠而落在缓冲区内的弧段。另外，在进行地矿勘查与开发资料分析时，往往需要对同一空间实体的不同段落划定不同的缓冲区，例如断裂两侧影响带或者矿脉两侧矿化蚀变带的宽窄受到许多地质因素的影响，要以露头、钻探、物探和化探资料为依据来圈定。空间信息子系统具有求取可变缓冲区的能力，可以支持这项工作。

（2）叠置分析

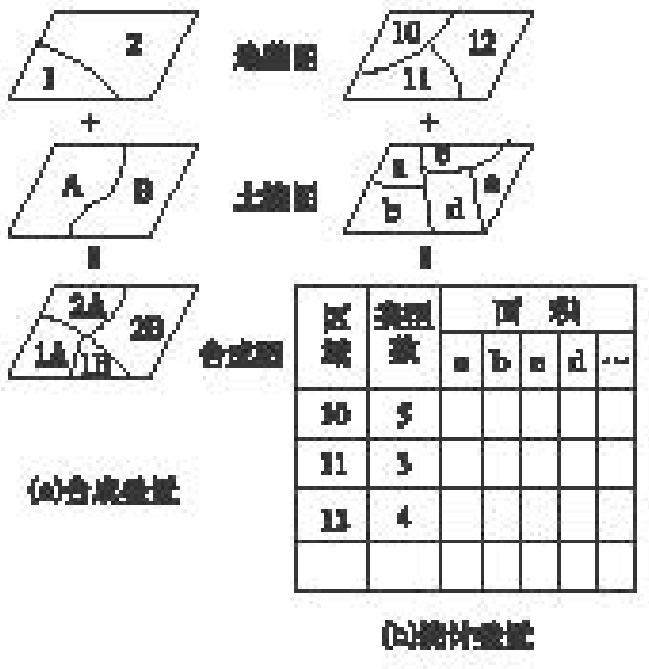
在进行资源预测时，常常需要综合多种空间信息和属性信息，要求空间信息子系统具有把不同数据层的相关信息叠合起来，形成新数据层的能力。叠置分析以叠加运算为基础，将同地区、同比例尺的两组或多组地质体多边形要素的数据文件进行叠加，然后根据两组多边形边界的交点建立具有多重属性的新多边形，或者对新多边形范围内的属性特征进行统计分析（图 5-35）。前者称为合成叠加，后者称为统计叠加。数据模式不同叠加运算的方式也不同，叠置分析据此可分为栅格数据叠置和矢量数据叠置两种。

1）栅格数据的叠置分析。对于栅格数据，层间叠置可通过像元之间的各种运算来实现。设 A，B，C 等表示第一、第二、第三等层同一坐标处的属性值，函数 f 表示各层上该处属性与用户需要之间的关系，U 为叠置后属性输出层的属性值，则

$$U = f(A, B, C, \cdots)$$

叠置操作的输出的结果可能是：①各层属性数据的平均值（简单算术平均或加权平均等）；②各层属性数据的最大或最小值；③算术运算结果；④逻辑条件组合。

许多地矿勘查的原始资料，例如化探资料、矿化露头资料等都是离散数据，容易转换成栅格数据，因而便于栅格方式的叠置分析。



(a)合成叠置　　(b)统计叠置

图 5-35　叠置分析

2）矢量数据的叠置分析（拓扑叠置）。矢量数据的叠置分析实际上是一种空间关系的拓扑叠置分析。叠置可以是多边形对多边形的叠置，也可以是线对多边形的叠置、点对多边形的叠置、多边形对点的叠置，点对线的叠置。不同数据层的矢量数据叠置之后，必然产生一些新的实体和新的拓扑关系。这些新的实体和新的拓扑关系都可以通过叠加运算来求得，因此，具有完整拓扑关系的矢量数据层叠加后，所得的结果也必然具有完整的拓扑关系。拓扑叠置还能够把不同数据层的同一实体的属性合并到一起，实现特征属性在空间上的连接。矢量数据的叠置有主动层与被动层之分，地位不同，叠置的结果也不同。例如，点对多边形的叠置与多边形对点的叠置，结果有显著的差别，不能混淆。

多边形与多边形叠置　两个多边形叠置时其边界互相切截将形成新的多边形，新数据层的多边形数目通常多于两个数据层叠置前的多边形总和。此外，由于各数据层本身存在误差，在多边形叠置操作时也会产生许多多余的小多边形——碎多边形或伪多边形。数据层本身误差来自同一根线被分别输入两个数据层时的两次数字化操作。人们在实际工作中经常遇到同一根线有两种或两种以上含义的情况，例如，一段河流可能是地壳大断裂的一部分，也可能是两套地层分界线的一部分，还可能是某种行政区域界线的一部分。这四种数据通常被分别存贮在四个数据层（图层）中，即使数字化的精度再高，也不可能使四次输入的数据完全相同。为了消除伪多边形，有些系统允许用户设置一容差值，以便使这类线段重合。但这一容差值较难把握，因为有些时候不同数据层的对应线段确实有些差别，若容差过大，有些真实的多边形将被删除；若容差太小，又不能完全剔除错误的多边形。

多边形与多边形的叠置可以有合并（Union）、相交（Intersect）、相减（Substraction）、判别（Identity）等方式。它们的区别在于输出数据层中的要素不同：合并保留两个输入数据层中所有多边形；相交则保留公共区域；相减从一个数据层中剔除另一个数据层中的全部区域；判别是将一个层作为模板，而将另一个输入层叠置在它上面，落在模板层边界范围内的要素被保留，落在模板层边界范围以外的要素将被剪切掉。

线对多边形叠置　线对多边形叠置的结果将生成一个新的线数据层。新线数据层中的弧段将具有它们所在的多边形的属性。例如，断层以线的形式存放在一个数据层中，将它与地层多边形数据层叠置后，其结果能够用来决定每条断层落在不同地层中的长度。线对多边形叠置的运算有相交、相减、判别等方式，叠加运算结果分别是穿过多边形的要素部分、多边形以外的线要素、被多边形切断的所有线要素（包括两个交点）。

点对多边形叠置　点对多边形叠置实质是计算多边形对点的包含关系，叠置结果将生成一个新的点数据层。该新数据层内有一串带有附加属性的点要素，即点所在的多边形的属性被连接到点的属性中。例如，露头地质点的位置以点要素的形式存放在一个数据层中，地层分区以多边形要素的形式存放在另一个数据层中，进行点对多边形叠置的结果便可以确定露头地质点在各地层分区内的分布。点对多边形的叠加运算也可以有相交、相减、判别等方式，运算结果分别是落在多边形内的点要素、多边形以外的点要素、所有点要素。

多边形对点叠置　多边形对点叠加的结果是生成新的多边形数据层，但只保留那些有点落在上面的多边形，这种叠置不作属性连接，结果多边形的属性和原始多边形相同。

点对线叠置　点对线叠置生成新的点数据层，保留点的所有要素。叠置时系统将自动找到距离某点最近的线并计算出点线间距离，然后将线号和点线距离记录到该点的属性中。

点线距离定义如下：

对任意点 D 和曲线 L，假设 L 由 n 个离散点 d［0］、d［1］、d［2］…d［n］构成，则 D 到 d［0］、d［1］…d［n］的距离分别为 S0、S1…Sn，D 到直线段(d［0］，d［1］)、(d［1］，d［2］)…(N［n-1］，d［n］)的法线距离分别为：

$$\begin{cases} D\text{到}(d[i-1],d[i])\text{的法线距离} & \text{若法线距离存在} \\ \infty & \text{若法线距离不存在} \end{cases}$$

那么点 D 到曲线 L 的距离　S = min（S0，S1，S2，…Sn，l1，l2，…ln）。

点对线的叠置具有求解点与最近距离的线的作用，其用途是显而易见的。例如，在进行矿床经济评价时，要求计算未来矿山的交通运输成本。如果将矿床以点要素的形式存放在一个数据层中，各种水陆交通线以线要素形式存放在另一个数据层中，则进行点对线的叠置后将求出离每个矿床最近的运输线及相应的距离。

在合成叠置的基础上还可以进行统计叠置分析，以便统计一种地矿要素在另一种地矿要素中的分布特征，例如统计某省各类土地的面积，统计重力、磁法、化探、遥感、地质信息叠置后的空间特征，获得更为深刻丰富的矿产资源分布信息，进而确定找矿靶区、圈定地质构造等。叠置结果可以以统计报表的形式给出。

(3) 地形分析功能

地形起伏的空间连续变化可以用数字高程模型（DEM）来表达。地矿资源空间信息

子系统提供建立数字地形模型（DTM）的功能，可用于生成三维景观、分析地形特征、制作剖面图、计算工程方量、计算地表汇流面积，以及其他地形的定量化分析，等等。

1）数字高程模型的建立。数字高程模型主要有格网 DEM、不规则三角网 TIN 以及由两者混合组成的 DEM 三种形式。格网 DEM 数据简单，容易管理，但高程是原始采样点的派生值，内插过程将要损失高程的精度，仅适合于建立中小比例尺的 DEM。TIN 直接利用原始高程取样点重建表面，能充分利用地貌特征点、线，较好地表达复杂的地形，但存储量大，不便于大规模规范化管理，也难以与 GIS 中的图形矢量数据或栅格数据以及遥感影像数据进行联合分析和应用。

2）地形分析。常用的地形分析包括如下几种:

等高线分析　地矿专业人员通常利用等高线图来表示和识别地形特征，包括地形的起伏程度，区域内各部分的高程等等。等高线图可从格网 DEM 中获取，也可在 TIN 中生成。

透视图分析　GIS 具有绘制地质体透视图的功能，还可以在三维格网上涂色，地矿人员可以直观地观察地形的概貌，也可以直观地剖析地质体表面的起伏变化。

坡度坡向分析　建立了数字高程模型后，很容易在格网内或三角形内计算地质体坡度和坡向，派生出坡度和坡向图供地形分析用。

断面图分析　断面主要用于地矿生产中的工程设计和工程量算，地矿人员可以在断面图上观察某地质体剖面地形的起伏和计算剖面面积。

地形表面面积和挖填方体积计算 以每个格网或每个三角形为基本单元，根据地质体地形的起伏计算每个单元的斜面面积，进而求得整个地质体数字高程模型的表面面积。用户还可以根据两个不同时期的数字高程模型计算挖填方的体积和范围。

（4）网络分析

网络分析是根据网络拓扑关系（线性实体之间，线性实体与结点之间，结点与结点之间的联接、连通关系），利用网络元素的空间数据和属性数据，对网络的性能、特征进行各种分析和计算。网络分析模型通常用来研究工程施工中物资和能量的流动、分配等，以选择最佳路径、规划物资的合理调运，确定工程中各项设施布局和服务范围，分析金属成矿物质和油气的运聚路径、成矿流体活化的集散度及其在构造网络中的流量分配，还可以用来研究不同成矿模式所形成的矿田、矿床和矿体的布局等。

1）网络数据模型。在空间信息子系统中，网络专指由若干线性实体通过交汇点联结而成的一类数据实体。网络数据模型是真实世界中网络系统的抽象表示，其中包括人工网络（例如交通网、通讯网、地下管网等）和天然网络（例如地表水系、地壳断裂网络等）。

构成网络系统的最基本元素有两种，即线性实体（网线或链，Link）及其联接交汇点（结点，Node）。网线构成网络的骨架，是系统中物质、能量和信息传输的通道，可以代表公路、铁路、航线、水管、煤气管、河流、断层、矿液或油气运移通道等等；结点既是网线的端点，又是网线的交汇点，可以表示交叉路口、中转站、河流汇合点、断层交汇点等等。此外，根据需要，网络还可以设定若干附属元素，例如在路径分析中用来表示途经矿石或矿产品装卸站点（Stop）、在资源分配中用来表示矿石或矿产品发散或汇聚中心（Center），以及对矿石或矿产品传输或通讯联络起阻断作用的障碍（Barrier）等等。网络附属元素的分类和设定的不同，体现了空间信息系统的网络数据模型的差异。

为了满足网络分析的需要，网络系统的网线和结点除自身的常规属性外，通常还要附带一些特殊的属性数据。例如，用于路径分析和资源分配的网线数据，应包含正反两个方向上的阻碍强度（如流动时间、耗费等）和资源需求量（如学生人数、水流量、成矿流体和油气运移量等）；与此相应，该网络的结点数据也应当包含资源需求量等。某些 GIS 平台（如 ARC/INFO、MAPGIS）还允许每个结点带有一个转角数据——转向表（Turntable），用于说明资源从某一网线经该结点到另一网线时所受的阻碍强度。与资源分配中心（Center）相联系的数据，包括该中心的资源容量、阻碍限度（资源流出或流向该中心所能克服的最大累积阻碍）。有些 GIS 系统还允许赋予中心一定的延迟量，以表达该中心相对于其他中心进行资源分配的优先程度。此外，与站点（Stop）相关的数据一般附带有传输量（即资源装卸量）、阻碍强度等。障碍（Barrier）一般无须附带任何相关数据。

网络分析系统研制者的任务是建立一个抽象的、具有相当适应面的，并且也是便于实现的网络模型；而使用者应当根据自身管理分析的需要，深入理解现实网络系统中各组成部分的特点及其相互关系，用网络模型中的不同元素合理地表示这些组成成分。

2）常用的网络分析。在地矿勘查与开发过程中，经常使用的网络分析功能包括路径分析、连通分析、资源分配、流分析和选址等几种。

路径分析　路径分析的核心是对最佳路径和最短路径的求解。从网络模型的角度看，最佳路径求解就是在指定网络中的两结点间寻找一条阻碍强度最小的路径。最佳路径的确定取决于网线和结点转角的阻碍强度（如果模型中的结点具有转角数据的话）。如果要寻找的是最快的路径，阻碍强度应当预先设定为通过网线或在结点处转弯所花费的时间；如果要寻找的是费用最小的路径，阻碍强度应当预先设定为费用；如果要寻找的是最短的路径，则网线的阻碍强度都应当预先设定为网线的长度。当网线在顺逆两个方向上的阻碍强度都是该网线的长度，而结点无转角数据或转角数据都是零时，最佳路径就成为最短的路径。

有时候需要利用路径分析来求解最佳游历方案，其中包括网线游历和结点游历。网线游历方案求解，就是给定一个网线集合和一个结点，分析由指定结点出发至少经过每条网线一次而回到起始结点的最佳路径。结点游历方案求解，则是给定一个起始结点、一个终止结点和若干中间结点，确定由起点出发遍历全部中间结点而达终点的最佳路径。

连通分析　第一类连通分析问题是连通分量求解问题，即要求弄清从某一结点或网线出发能否到达的全部结点或网线，以及其通达程度如何。这是在日常工作和生活中经常遇到的问题。例如，在分析油气二次运移时，需要知道通过由砂体、不整合面和成藏期活动的断层、裂隙带等组成的输导网络，烃类流体从源岩附近出发能否到达每个圈闭。某些圈闭与排烃的源岩如果不在一个连通分量之内，那里就不可能有油气的聚集。

另一类连通分析问题是最少费用连通方案的求解问题，例如，公路部门拟修建足够数量的公路，使某个县的几个乡镇直接或间接地相互联结，如何使费用最少呢？如果把每一条可能修建的公路作为网线，把相应的预算费用作为网线的耗费，上述问题就转化为求一个网线集合，使全部结点连通且总耗费最少。

资源分配　资源分配是指为网络中的网线和结点寻找最近（按阻碍强度的大小确定）的中心（资源发散或汇集地）的分析方法。例如，为城市中和农村的每一个社区、每一个村庄确定供水（灌溉）的水厂（水库）和供电的变电站，并模拟资源是如何在中心

(水厂、水库、变电站等) 和它周围的网线 (水管、水渠、线路等)、结点 (社区、村庄等) 间流动的。这种方法也可用于分析成矿流体和含油气流体的来源，只是阻碍强度类型不同而已。

资源分配是沿最佳路径进行的。其要领是：根据中心容量以及网线和结点的需求，将网线和结点按最佳路径分配给中心；当网络元素被依次分配给某个中心后，该中心拥有的资源量就依网络元素的需求而缩减，直至中心的资源耗尽。

流分析　所谓流，是指将资源由一个地点运送到另一个地点。流分析的问题主要是按照某种最优化标准 (时间最少、费用最低、路程最短或运送量最大等) 设计运送方案。在进行流分析时，要根据最优化准则下的不同扩充网络模型，把中心分为代表资源运送目标点的“收货中心”和代表资源运送起始点的“发货中心”。其中，“发货中心”的容量代表待运送资源量，“收货中心”的容量代表所需要的资源量。对于不同研究领域的不同最优化准则，网线的属性数据会有所不同。如果最优化准则是运送量最大，就要定义网线的传输能力；如果目标是使费用最低，则要定义网线的传输单价；如果目标是使损耗最低，则要定义网线的传输损耗率。这个方法适用于商业流通领域，也适用于成矿成藏研究领域。

选址　选址分析也是基于最佳路径的一种求解方法。在地矿资源开发领域，除了确定服务性设施位置之外，还需确定某一矿区内优先开发的矿体，以及设计工业广场和矿井的最佳位置。在网络分析中的选址问题一般限定设施必须位于某个结点或位于某条网线上，或者限定在若干候选地点中。由于对“最佳位置”的解释标准不同以及要定位的设施数量不同，选址问题是多种多样的，其实现方法和技巧也是多种多样的。

(5) 邻域分析

邻域分析是追索同类目标在邻近区域的分布特征的一种空间分析方法。利用这种方法，可随时检索已知的某种实体 (地矿目标) 在邻近区域内的分布信息。例如，在矿产资源预测中的就矿找矿、根据成矿序列法寻找新类型矿床、老矿区深部和外围的成矿预测，以及研究某种岩石类型和矿床类型的空间分布特征，等等。

四、空间数据格式的转换标准

一般的 GIS 数据格式是不公开的，且不易为人工阅读。为了方便而高效地实现空间数据间的传输和交换，必须有一种统一的标准格式。当前的空间数据转换标准可分为两大类，即矢量标准和栅格标准。这里仅介绍一些目前较为流行和公认的转换标准，它形成基于这样的原则：“充分考虑实际的或潜在的问题，为通用的及重复性使用建立保障，在约定的背景中达到最佳的次序”(Arnold 和 Dnce，1990)。

1. 矢量数据标准转换格式

(1) NTF (National Translation Format) 格式

NTF 即英国国家转换格式。该格式是英国测量局为其所有数字产品选择的传送格式，目前已被 AGT (英国地理学会) 挑选作为用于空间数据交换的优选格式。

NTF 由一系列以数值标号 1 至 5 的层所定义。层 1 和层 2 允许转换简单的地质体矢量数据，包括线、点等；层 3 允许用数据来表达不同地矿要素间的相互依赖关系，也允许不同数据模型间的转换；层 4 是专门为转换数据以适应 DIGEST 所定义的全拓扑数据模型而设计；层 5 可以包含不包括在基本定义格式中的新数据域和数据记录，用于转换那些未被

其他层所涵盖的数据模型。从目前的情况看，NTF 仅仅转换数据而不包括任何有关数据的描述信息，除要素编码信息量不足外，还不包括点要素符号、线要素类型及色彩、面填充式样、数据绘图、图廓资料等的转换。NTF 将这些视为软件控制输入及输出的内容。

(2) DXF 格式

DXF (图形交换格式) 由 AutoDesk 公司开发，用于其计算机辅助设计 (CAD) 软件包 AutoCAD。由于它在 CAD 领域内无所不及的影响，已被其他领域广泛采用。

DXF 用于转换完全的图形信息和几何元素。DXF 包含所有的用于描述数据的线型及颜色的定义、不同的文本类型、通过数据分层描述的要素编码信息，以及许多用于 AutoCAD 的系统变量的定义 (AutoDesk, 1992)。DXF 的缺陷是没有描述拓扑关系。

(3) DIGEST 格式

DIGEST (数字地理信息交换标准) 由北约军方制图组织代表组成的数字地理信息工作小组 (DGIWG) 开发出来，主要用于地图产品机构间基本数据和产品的交换。

DIGEST 定义了大量不同的详细格式，分别用于不同数据模型的转换。它们包括一种依赖于关系表格的矢量关系格式 (VRF)。VRF 过去常被称为 VDF (矢量产品格式)，最初是由 ESRI (资源环境研究所) 连同“世界数字航图”工程一并开发出来的，后来才加入到 DIGEST 之中。此外，DIGEST 还包含有一个可被用于数据转换的要素描述总目录，即《要素与属性编码目录》(FACC) (DGWG, 1992)。FACC 的目的是确保要素能从可能包含于被转换数据中的要素编码属性里正确地解译出来。采用 DIGEST 标准的优点是解译的可靠性较高，特别适合于语言差异掺杂于误解译之中的国际性转换场合。

(4) SDTS 格式

SDTS 即“Standard of Spatial Data Transition——空间数据转换标准”的缩写。这种矢量文件格式是由美国地质调查局开发的，在 1992 年被采用作为“美国联邦信息处理标准 (US FIPS)”(Fegeas 等，1992) 而强制使用，迫使许多软件厂商着手为他们的软件增加 SDTS 的输入和输出接口。由于美国在计算机活动领域内的影响，它在世界各地可能会变得越来越重要。该标准定义包括有矢量数据和栅格数据。通过定义为 SDTS 子集的方式对特定产品加以局限化的标准，称为子标准，例如《拓扑矢量子标准》(TVP) (Bar, 1992)。TVP 已经被用于具有类似于 DIGEST 全拓扑数据模型的矢量数据转换。如同 DIGEST 采用《要素与属性编码目录》一样，SDTS 也包含有一特征编码表。

2. 栅格数据标准转换格式

在栅格数据中，仅存在一种可能的数据模型，即由仅通过其几何位置表达其相互关联的元素阵列组成。每一个元素可以附带一值，该值可被用于解译和推断元素间的关联。已经出现的常规标准有《特征图像文件格式》(TIFF)、Postscript、《计算机图形元文件》(CGM) 和《等弧秒栅格产品标准》(ASRP)。

(1) TIFF 格式

TIFF (特征图像文件格式) 是用于常规目标空间数据的最为重要的栅格数据转换标准，由 Aldus 公司、Microsoft 公司和其他商业公司于 1986 年所共同推出。TIFF 文件是一种二进制文件而非文本文件，其标识是文件第三和第四字节为“42”。TIFF 有两个重要特色，第一是其标记的格式，其二是其支持不同形式数据压缩的能力。在这种标记的格式中，每一节数据前有一相应短码标识 (或称“标记”)，依靠这些“标记”可以对数据进

行正确解码。压缩数据是栅格数据模型的重要用途，TIFF 中包含有多种压缩方法，如运行长度编码、二维运行编码等等，具有广泛的适应性和实用性。TIFF 的技术规范要求使所有的软件均能对基本 TIFF 文件读写，用户应当确保所接受的 TIFF 形式是他们可以理解的。

(2) CGM 格式

计算机图形元文件（CGM）已被完善地制定为一种国际标准，并且在计算机辅助设计领域中被广泛采用。CGM 有 3 种编码形式，即二进制编码、字符编码和清洁文本编码，是人工可读的，并且是一种相对简单的格式。其二进制编码十分紧凑，压缩强度随转换图形图像应用以及机器结构而不同。

(3) Postscript 格式

Postscript 是作为输出图像至栅格设备（如激光打字机）的一种输出格式。该格式与设备无关，在实际应用中，栅格格网的大小在文件中不指定，而是由独立输出设备解释。因此该转换标准可以适用于各种各样的输出设备。

(4) ASRP 格式

ASRP（等弧秒栅格产品标准）是专为空间数据而设计的栅格标准（Gennery，1992）。它指定像素密度为每厘米 100 点。该标准设计用于转换较大地理范围的数据，包含有坐标系统（WGS 84）的定义及用于转换经纬坐标至笛卡儿坐标的投影定义，可以在东—西方向指出图像的实际变形。ASRP 文件以行程编码方式压缩。

第四节　地矿资源空间信息子系统应用举例

在国土地矿调查与勘探中，应用空间信息子系统可以有效地避免野外手工记录、室内数据矢量化的冗余操作，实现有关数据的一次性入库；同时为在野外对地质草图及各种前期勘查图件进行直接修编提供了便利条件。地质工作者可以根据野外的实际情况，在现场对数字化的地质草图和各种前期勘查图件进行修改，在室内稍作整饰就可以打印、输出。本节以野外地质填图、工程勘察和资源预测为例，介绍空间信息子系统的应用。

一、基础地质调查中的空间数据采集与整理

这里介绍一种国产的应用软件——GeoSurvey。该系统由中国地质大学（武汉）国土资源信息系统研究所开发，目前装载于超轻型便携机上，也可以部分移植到掌上机上。该系统利用关系数据库（RDBS）来采集、组织和管理属性数据，而利用 GPS 和 GIS 来采集和管理空间数据，再通过内部编码实现两种数据的关联，能够实现从野外数据采集到室内综合整理，再到地质图编绘的全过程计算机辅助化。这里着重介绍其空间数据采集功能。

1. GeoSurvey 数据采集流程及图层划分

(1) 所采集的空间数据类型

GeoSurvey 所收集、采集和处理的空间数据，包括地形图、地质剖面图和地质（平面）图等图形资料，遥感图像、航空照片、数字照片、地质素描等图像资料，以及地质点露头或钻孔岩心的结构、构造、岩石、矿物等文字资料。这些数据所描述的地质体空间特征可归结为点状、线状和面状等三大类特征（图 5-36）。其中，点状特征主要是描述地层界线点、采样点、化石点、产状测量点、钻探点等；线状特征主要是描述地质观察路

线、地质界线、断层、褶皱轴迹等；面状特征主要是描述地层、岩体、水体等分布区。

(2) 空间数据采集工作流程

GeoSurvey 的数据采集可分为数据准备、数据采集、数据整理三个过程（图 5-37）。在数据准备阶段，主要进行室内辅助资料如地形图、航空照片、遥感图片的扫描、矢量化及坐标转换等处理；在数据采集阶段，遵照所布置的地质观察路线，首先确定地质观察点的点位，然后按实地观察所见勾绘地质界线的走势，并进行其他地质特征的属性描述和记录；数据整理阶段也是地质草图修编阶段，主要工作是进行图面地质界线的连接、地质体边界位置和形状的修改，以及面状地质体的生成。数据整理阶段的工作主要是为地质图编辑和输出作准备，其中相当一部分工作也可以在野外现场开展。

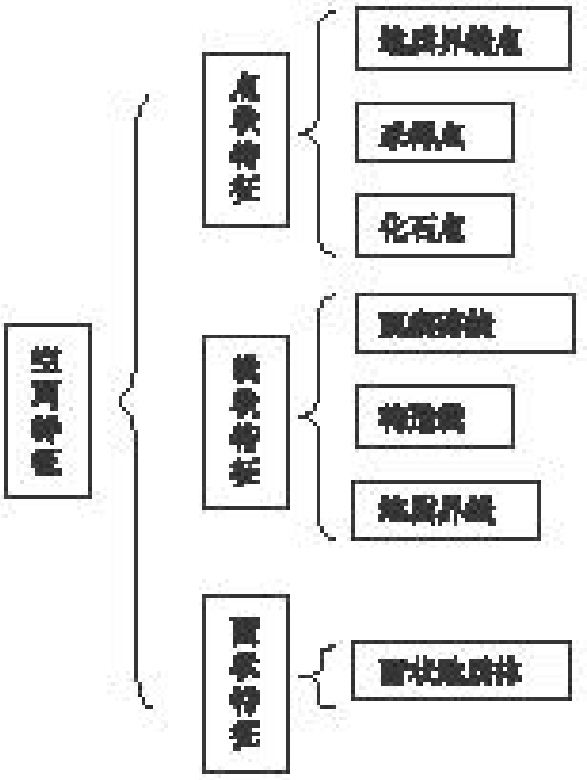


图 5-36 地质体空间特征分类

(3) 系统中的图层设置

在 GeoSurvey 系统中，空间数据图层设置为辅助图层和基本图层两大类。在完成空间数据采集和整理后，系统能自动地将用户当前勾绘、圆滑的地质界线和封闭的面状区域，保存到相应的图层中。

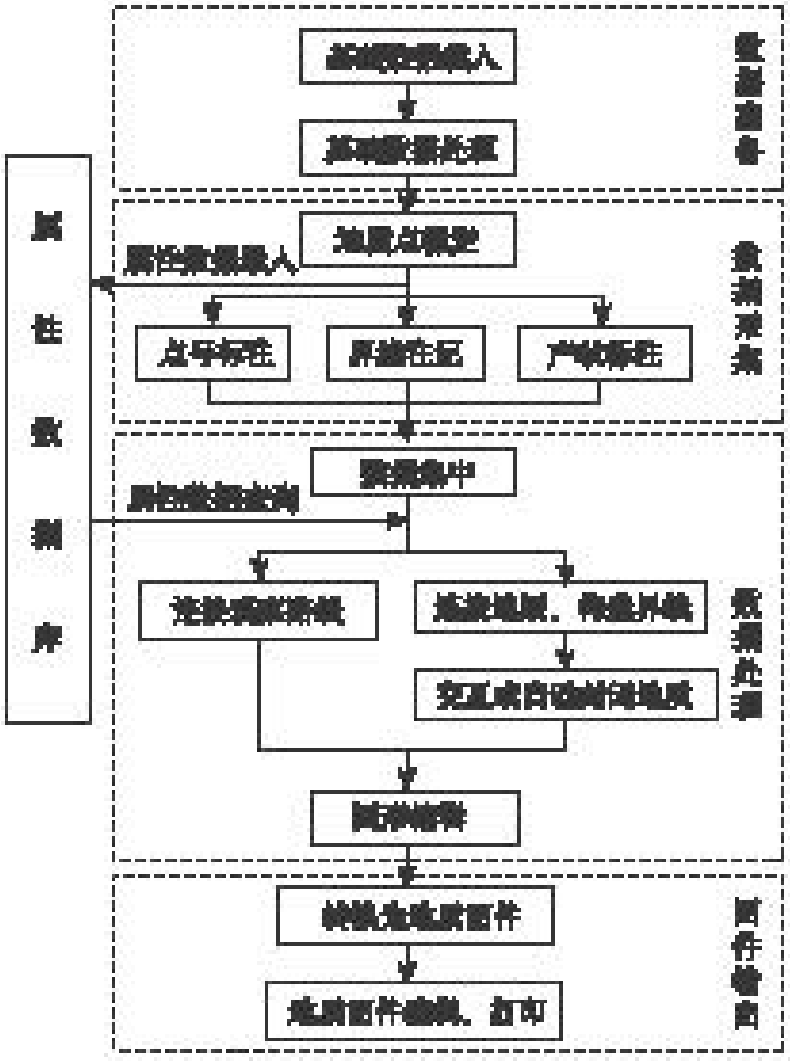


图 5-37 GeoSurvey 的空间数据采集与地质图修编流程

1) 辅助图层。地形等高线图层 表示填图区域的地形，通过地形图矢量化获取。

方里网格线及图框图层 图幅上的方里网格线以及该图幅的边界线，通过地形图矢量化获取。

水体图层 表示池塘、湖泊、海域等以面状显示的对象。设此图层是为了对地质体进行掏挖，使得在上述区域的图面上不出现地质体。

道路水系图层 此层可以放置各种道路和水系，用于辅助确认地质点的位置及其拓扑关系。

其他图层 航空摄影图片、遥感影像图片等单独成层，便于单独调用。

2) 基本图层。地质观察点图层 保存野外填图中所定的地质观察点；

地质路线图层 保存填图过程中的地质观察路线及其特征点；

界线注记图层 保存地质观察点

上观察到的地质界线的走势;

地质线图层　保存根据地质点连接而成的地质界线、断层线及其他构造线等;

地质体图层　保存封闭地质线生成的地质体;

GPS 点图层　保存根据手持 GPS 接收机数据确定的地质观察点;

图框图层　表示填图区域的边界，利用本系统绘制或地图边界的矢量化结果。

除此以外，系统还设置了一个标注文件，用于保存游离在图层以外的线条、符号，例如标记在图上的地质产状、地质点号、地层单位代号等对象。

2. GeoSurvey 的基本功能

该系统所具备的基本功能是对多源、多维空间数据进行有效地采集、管理、综合整理，并提供数据查询、检索，以及地质草图修编和地质图编辑输出。

(1) 野外空间数据采集

基础地质调查中的空间数据采集主要包括各类地质点的标定、各类地质产状的测量与标注。采集空间数据是 GeoSurvey 系统最基本的功能，其中贯穿着由点到线、由线到面的空间数据采集、分析和标注（绘）工作——即由地质点观测和标定推演出地质界线，而由地质界线连接和封闭成地质体（面)。野外空间数据采集子系统的结构 - 功能如图 5-38 所示。

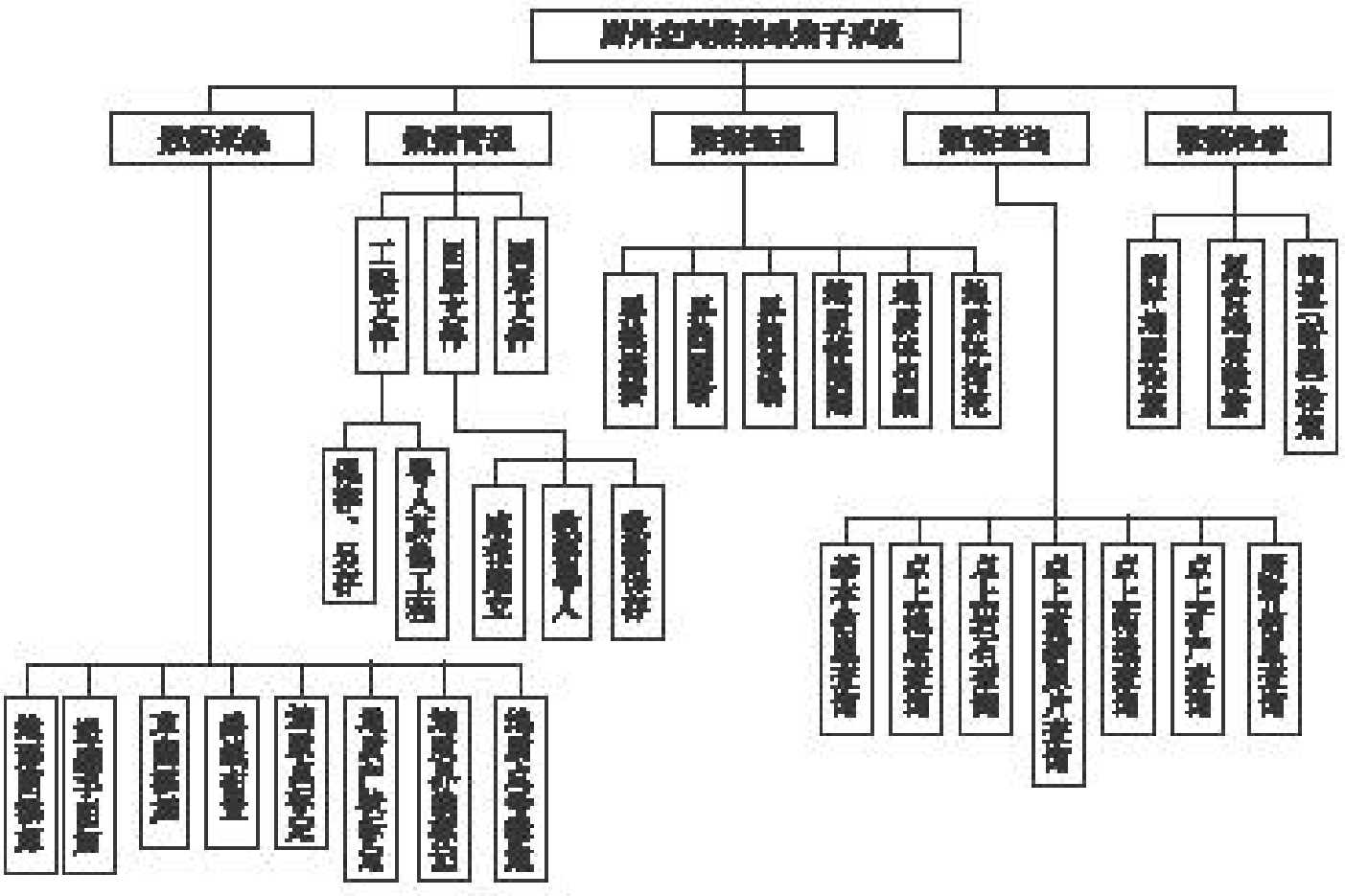


图 5-38　野外空间数据采集子系统的结构-功能图解

GeoSurvey 允许采用机载地形底图（手图)、航卫照片和 DEM 数字影像地图，直接进行点位、地层界线和地层产状标注（图 5-39，图 5-40)，也可以用手持 GPS 来标定地质点位置。GeoSurvey 还设置了专门的地质剖面测制模块。该模块采用常规的实测剖面数据采集表格，可以在野外直接采集导线方位、坡角、地层产状、地质界线位置、地质构造和地质体特征等参数，并且由计算机自动完成全部参数的投影计算，同时完成综合柱状图、路线地质

图、地质剖面图及整个图面的设计和编绘（图 5-41）。

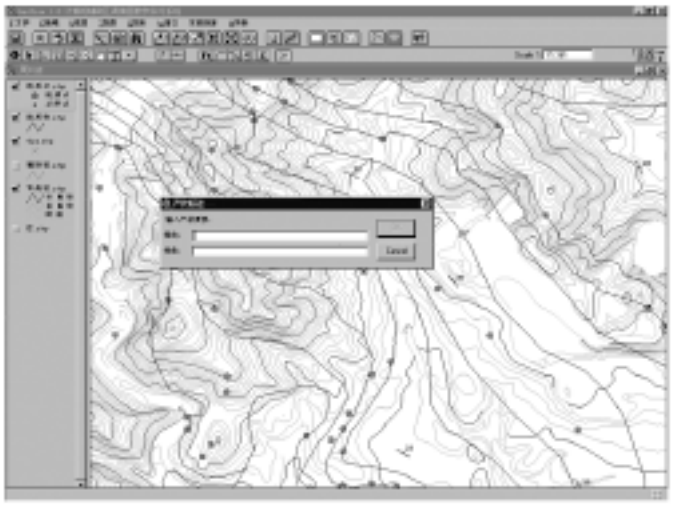


图 5-39　在便携机装载的地形图上直接进行点位、地层界线和地层产状标注



图 5-40　在便携机装载的航片和地形图的叠合图上进行野外辅助定点与导航

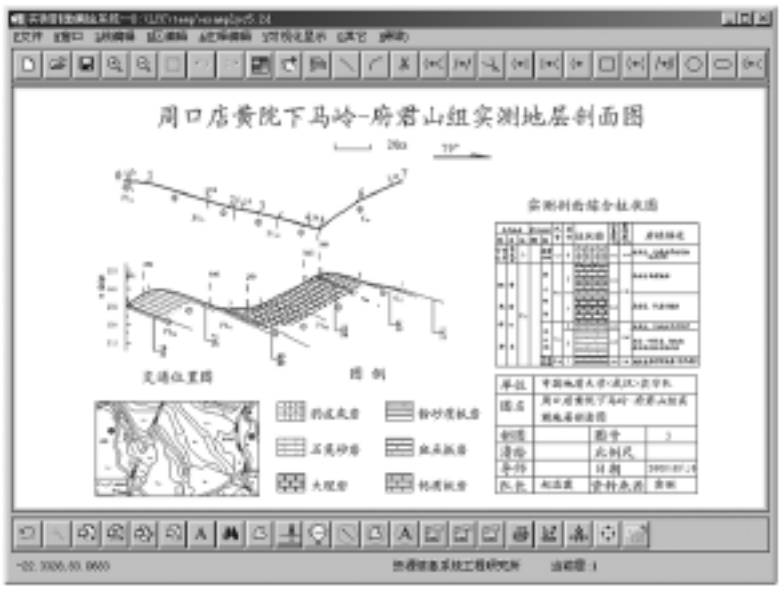


图 5-41　利用 GeoSurvey 自动生成包括综合柱状图、路线地质图在内的完整实测剖面图

(2) 空间数据管理

GeoSurvey 提供了工程文件管理、图层文件管理、标注文件管理和接口文件管理等功能。其中，工程文件用于记录与各图层文件有关的联络信息，不包含各图层中的空间实体对象。其作用是对该幅图的所有图层、图形进行组织、管理，便于用户调用图层文件。图层文件用于保存相应图层中空间实体的位置、拓扑关系和属性信息。用户只需事先对填图图幅进行属性设置，同时指定图层文件存放的专用目录，就可以保证数据的独立性，不需要进行其他干预。

标注文件专用于管理用户标注在图上的点号、产状、线号等对象。用户只要指出目标数据在计算机中的位置，就可以准确提取并存放到当前的图层文件中去。

接口文件提供了导入数据的接口，可以在维持各图幅和各野外小组的数据相互独立的同时，方便地在工程文件之间导引、传输和集中数据。这样既可保证一个地质图幅由多组人员进行填制，又可保证一组人员同时填制多个图幅，而不发生数据混淆。

此外，系统也可方便地读取 AutoCAD 的 DWG 文件及 DXF 文件，还可与 MapGIS、shapefile 进行明码格式文件转换，读取明码格式文件中的点、线、面文件。

(3) 空间数据整理

GeoSurvey 的空间数据整理主要是连接地质观察路线和地质界线，生成面状地质体以及进行地质图的图面编辑。

1) 地质观察路线的连接

系统可以依据地质观察点的编号进行分类，将属于同一路线的点依编号进行连接，生成的地质路线保存在地质路线图层中。

2）地质界线的连接

GeoSurvey 允许直接在机载地质草图上修订地质界线（图 5-42），或直接在机载地形图上勾绘地质界线，也可按“V”字形法则自动或人机交互生成地质界线。地质界线的自动生成是通过设置固定的缓冲区半径拾取地质观察点的有关信息，再根据“V”字形法则计算地层与地形的关系判断地质界线的走势来实现的。

3）面状地质体的处理

面状地质体指的是通过封闭地质界线得到的多边形，图上表现为一个二维的面，因此以面图层进行组织、管理。本系统可通过确定填图区范围、剔除湖泊、塘地等面状空间特征和利用地质界线对地质体进行分割等步骤自动生成面状地质体。

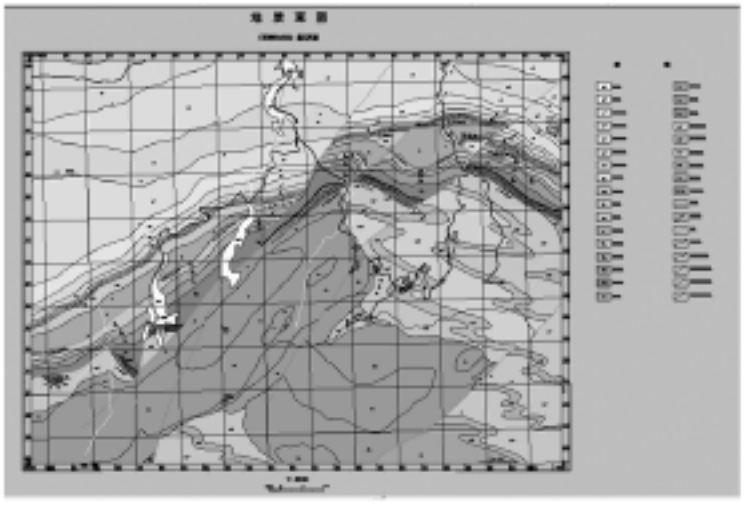


图 5-42　在机载地质草图上修订地质界线

4）图形图像的叠合显示

地质填图工作涉及的原始图件资料比较多，由于不同时期制作的图件所采用的比例尺和地图投影方式往往不同，在对相关矢量化图形进行叠合时需要事先做空间位置配准。另外，诸如遥感解译图像、航片、重力异常图、磁异常图等栅格型数据文件，虽与空间位置有关却常常没有坐标。要使这些图像与有确定坐标的图形正确地叠加，也需要进行空间位置的配准。GeoSurvey 提供了完善的空间位置配准功能。

5）图面编辑。GeoSurvey 系统提供如下 9 种基本的图面编辑功能：

圆滑线条　采用三次样条函数插值法对线条或图形进行圆滑处理；

抽稀线条　对特定线条的节点做抽稀处理，减少其数量以便于进行其他编辑；

裁剪图层　利用特定的封闭图框对选定的图层进行裁剪，位于封闭图框内的图层特征将被保留到新的图层中；

鼠标剪断线条　将鼠标点中的线条在鼠标点击处剪断；

线条剪断线条　可以用两相交线条中的一根作为“剪刀线条”，剪断另一根线条并且将其较短的一截弃用；

连接线条　可以通过距离最近的两个端点，使两根线条连接成一根线条；

扯动线条　可以用光标“扯动线条”，使地质界线经过某地质观察点；

封闭线条　将一根线条的首尾端点连接起来，使之成为一个封闭的弧；

各类符号标注　可以通过鼠标点击目标对象调用数据库信息，自动标注地质点的点号、产状符号和地层单元代号等，并且可以调整符号的大小、位置和方向。

6）地质构造分析。为了便于在野外开展地质构造分析，GeoSurvey 提供了自动编绘节理走向和倾向玫瑰花图、节理和断层产状极密投影图、地层和层理产状赤平投影图和构造应力 - 应变分析图等项功能。

7）地质素描、数字照片的实时显示。在进行野外地质调查时，常常需要绘制、拍摄大

量的素描和照片。GeoSurvey 系统通过数据库对素描、照片进行有效地组织、管理，然后采用“热连接”方式实现地质素描、数字照片的实时显示。GeoSurvey 的属性数据库不但可以记录大量的描述信息，还可以有效地改变各种素描、照片的混乱无序状况。

(4) 空间数据的查询、检索

1) 空间数据查询。用户通过鼠标选择要查询的地质观察点，一旦选中，系统弹出图 5-43 所示的对话框，对话框的标题显示的是该地质观察点点号。点击其中的任意按钮，便可得到该地质点的相关数据（图 5-44）。如果按钮呈现灰色，表明数据库中未存放该项数据。若结果显示屏的对话框上有“详细信息”按钮，则点击后可进一步显示有关描述内容。



图 5-43 地质点查询

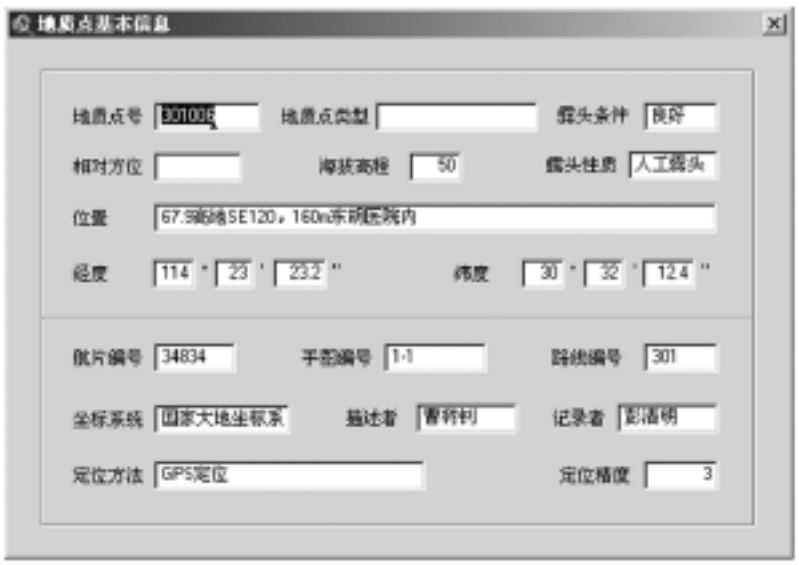


图 5-44 地质观察点基本信息查询

2) 空间数据检索。利用 GeoSurvey 的空间数据检索功能，可以对所有地质观察点进行单一地层单元和构造（断层）现象的检索，也可以对两套地层单元进行复合检索。

3. GeoSurvey 的应用

利用 GeoSurvey 提供的上述各种功能，可以在野外方便地采集路线和地质点的全部属性数据、空间数据和影像数据，实现从野外数据采集到室内资料综合整理，再到地质图编绘的全程计算机辅助化。

GeoSurvey 还可以方便地将野外调查资料、物探资料、化探资料和钻探资料集成起来，达到“多 S”信息交互检索、综合处理和综合应用的目的。

由于该系统采用行业统一的数据模型、标准的代码体系、规范的图式图例、约定的处理方式和通用的软件接口，可以支持组建区域性的“基础地质信息系统”。

二、地质异常的空间分析与资源预测

空间信息系统技术在矿产资源预测评价中的应用正在日益加强，并且已经取得了一系列引人注目的成果。这里着重介绍它在地质异常分析和金属矿产资源预测中的应用。

1. 地质异常的基本概念

赵鹏大等（1991）指出：地质异常是在物质成分、结构构造或成因序次上与周围环境有着明显差异的地质体或地质体组合。其表现形式不仅在物质成分、结构构造和成因序

次上与周围环境不同，而且还经常表现在地球物理场、地球化学场及遥感影像异常等的不同上。如果用一个数值区间（或阈值）来表示背景场，凡是超过或低于该阈值的场就构成地质异常，它具有一定的空间范围和时间界限。大型、超大型矿床都分布在大大小小不同级别的地质异常中。因此，地质异常理论的研究对寻找大型、超大型矿床有独特的作用。

地质异常的尺度水平不同，其表现的形式也不一样。大的如板块的俯冲带和地缝合线(板缘构造)，小到岩石、矿物、矿物内的包裹体甚至矿物的各种超微结构。有的是有形的，如地质体的不连续界面或不同地质体的分界面、地质体内部及外部特征突然变化或突出变化的部位，以及不同成因地质体的嵌入；有的是无形的，例如单位面积或体积内各种地质体或同一地质体不同属性组合熵异常，地质体的构造复杂程度、相互间的相似或关联程度等。有形的地质异常可以直观地判断，而无形的地质异常要通过分析、综合才能识别。

地质异常是在不同地质历史时期演化发展的产物。地质异常形成的地质时代、构造背景、地质环境和岩石类型，决定了异常的性质及其赋存的矿产资源种类和规模。因此，地质异常具有空间和时间上的演化序列。

2. 地质异常的提取方法

不同尺度水平的地质异常，其研究和提取方法所依据的资料和数据各异，需要根据实际情况选择使用。下面介绍几种基于求异理论的地质异常圈定方法。

(1) 地质复杂系数（C 值）法

根据计算公式可分为复杂系数（C_1）和复杂度（C_2）两种。

复杂系数（C_1）　复杂系数是指某一个单元相对于全国全部单元平均复杂程度的度量。首先按地质图上的地层构造、岩浆岩等凡能观测的变量进行统计，取其平均值作为全国全部单元平均复杂程度的度量（大背景或理想的正常场），用向量

$$X = (x_1, x_2, \cdots, x_p)'$$

来表示，然后计算每个单元相对于平均复杂程度的复杂系数，计算公式如下:

$$C_{1i} = \sum_{j=1}^{p} (x_{ij} - x_j)^2 \quad (i = 1, 2, \cdots, n)$$

其中 x_j 为向量 X 的第 j 个分量，共计 p 个地质变量（分量），可近似将它看成背景场，x_{ij} 为第 i 个单元第 j 个变量的取值，n 为单元总数。复杂系数有正负之分，其符号确定如下:

$$C_{1i} = \begin{cases} +, & \sum_{j=1}^{p} (x_{ij} - x_j) \geq 0 \\ -, & \sum_{j=1}^{p} (x_{ij} - x_j) < 0 \end{cases} \quad (i = 1, 2, \cdots, n)$$

当 C_{1i} 值为 0 时，表示该单元的复杂程度与全国全部单元平均复杂程度（正常场）相当；当 C_{1i} 值为正时，表示该单元的地质条件比正常场复杂，数值越大，复杂程度越高；当 C_{1i} 值为负时，表示该单元的地质条件比正常场简单，数值越小，地质条件越简单。

因此，根据每个单元的 C_1 值，我们可以圈定地质条件复杂或简单的地质异常区。

复杂度（C_2）　复杂度是另一种确定复杂程度更简单的方法。如果说复杂系数是相对的度量方法，那么复杂度是绝对的度量方法。它将每个单元所有变量的取值进行累加，

就可以得到每个方格的复杂度（C_2）。计算公式:

$$C_{2i} = \sum_{j=1}^{p} x_{ij} \qquad (i = 1, 2, \cdots, n)$$

经实践检验，其效果与复杂系数相同，只是变化幅度小些，而且在确定地质简单地区没有复杂系数直观（没有负值）。

应用地质复杂系数（C 值）圈定我国地质异常时应用效果非常显著。在所圈定的简单地质异常中分布着我国已知或正在勘探的全部油田，还预测了我国主要的浅海油田；在复杂地质异常中分布着我国大部分已知的大型内生金属矿产。

(2) 熵（H 值）法

熵是信息论中度量信息量的一种方法，它也反映事物发生的不确定度。一般来说，复杂系统的不确定度高。因此，地质构造特征越复杂，其不确定度越大，即熵值越高。这就是我们用熵来表示地质构造特征的复杂程度的原因。

对定和数据，如地层、岩浆岩等出露面积（其和为 1），熵（H 值）可用公式

$$H_i = -\sum_{j=1}^{p} x_{ij} \lg x_{ij} \qquad (i = 1, 2, \cdots, n)$$

来计算。对非定和数据，如断层条数等，熵（H 值）可以用公式

$$H_i = -\sum_{j=1}^{p} \frac{\frac{x_{ij}}{n}}{\sum_{i=1} x_{ij}} \lg \frac{\frac{x_{ij}}{n}}{\sum_{i=1} x_{ij}} \qquad (i = 1, 2, \cdots, n)$$

来计算。

其中 p 为变量数，n 为单元总数，x_{ij}为第 i 个单元第 j 个变量的取值（原始数据），对数可取自然对数或者以 10 为底的普通对数等。

(3) 地质相似系数（S 值）法

地质相似系数是某一个地质单元与周围单元之间相似程度的度量。一旦某个地质单元与周围的地质情况有变异时，相似程度就降低了，于是可以根据相似程度差别的阈值圈出地质异常的分布范围。具体步骤如下：①先将地质图划分成一系列单元 i（图 5-45），再统计从地质图中获取的地层、构造、岩浆岩等变量数值；②计算每个单元 i_0 与周围 8 个单元（i_1，i_2，…，i_8）的相似系数，然后取其平均值作为该单元与周围单元的相似系数。

i_1	i_2	i_3
i_4	i_0	i_5
i_6	i_7	i_8

图 5-45　研究单元与周围单元的空间关

地质相似系数的计算公式如下:

$$S_{ij} = \frac{\sum_{k=1}^{p} x_{ik} x_{jk}}{\sqrt{\sum_{k=1}^{p} x_{ik}^2 \sum_{k=1}^{p} x_{jk}^2}} \qquad (i = 1, 2, \cdots, n; j = 1, 2, \cdots, 8)$$

其中 n 为单元总数，p 为变量数，其他同上。

计算了 S_{i1}，S_{i2}，…，S_{i8} 以后，则有

$$S_i = \sum_{j=1}^{8} S_{ij}/8$$

作为第 i 单元与周围单元的地质相似系数（S 值）。

为了便于圈定地质异常，可以将相似系数转换成不相似系数，即

不相似系数 = 1 - 相似系数

除了计算每个方格与周围方格的不相似程度（不相似系数）外，还可以计算每个方格与平均值之间的不相似程度。最后可通过各种等值线来圈定地质异常分布范围。

(4) 地质关联度（R 值）法

用关联度圈定地质异常的基本原理与地质相似系数相同。关联度的含义也是单元之间的相似程度，只是计算方法不同而已。

3. 应用地质异常法进行矿体定位预测

矿体定位预测的地质异常法，是在地质异常致矿思想指导下，综合多学科信息，以非线性科学和信息技术为手段，以圈定不同尺度和不同类型的地质异常为途径，逐渐逼近工业矿体的一种新的定量成矿预测方法。

(1) 矿体定位预测的工作流程

由于不同尺度地质异常对不同等级矿产资源域（体）有对应控制关系，利用地质异常进行矿体定位预测，可归纳为对“5P”找矿地段的圈定（赵鹏大等，1998）。所谓“5P”地段即：成矿可能地段（probable ore-forming area）、找矿可行性地段（permissive ore-finding area）、找矿有利地段（preferable ore-finding area）、矿产资源体潜在地段（potential mineral resources area）和矿体远景地段（perspective ore body area）。

通过各种方法圈定的具有成矿基本条件的地质异常地段都是“成矿可能地段”。在成矿可能地段中，有可能找到预期矿床的异常（或异常带）地段即“找矿可行地段”；而结合更多找矿信息（物化探异常、重砂异常等）所确定的最有希望找到预期矿床的地段即“找矿有利地段”。在找矿有利地段内，运用中大比例尺（1∶5 万～1∶1 万）地质物化探综合信息所圈定的矿化显示地段，称为矿产资源体潜在地段，其中通过地质工程（槽探、坑探、钻探）、基岩化探、电磁测井等手段圈定的矿化地段，称为矿体远景地段。

矿体定位预测的地质异常法工作流程如图 5-46 所示。随着地质异常研究的深入，找矿异常信息由少到多，找矿靶区范围由大到小，靶区级别由低到高，找矿成功概率逐渐增大，靶区的经济价值也逐步增加。这是一种有序性很强的定量成矿预测方法，易于实现成矿预测的自动化。因此，利用数学和空间信息分析处理技术来不断完善其理论和方法，提高成矿预测的精度和自动化、智能化和可视化程度，是十分必要的。

(2) 地质异常的空间分析和矿床预测实例

在 GIS 支持下所建立的地矿资源空间分析子系统，是地质异常分析和“5P”地段圈定的有力工具；而各类异常找矿有利度分析和成矿预测空间模型的建立，是应用这一工具的基础。基于地质异常的矿床预测步骤如下：建立成矿预测空间模型；圈定找矿可行地段；确定用于预测找矿有利地段的异常组合及数值区间，并决定其权重；所选定的各异常叠合，形成叠合图；确定各级别找矿有利地段的归并标准，归并后圈出找矿有利地段。下面以滇中元江地区大红山群铜矿床找矿预测为例，对基于地质异常的空间分析和“5P”地段预测方法（池顺都等，1998）做一些简要介绍。

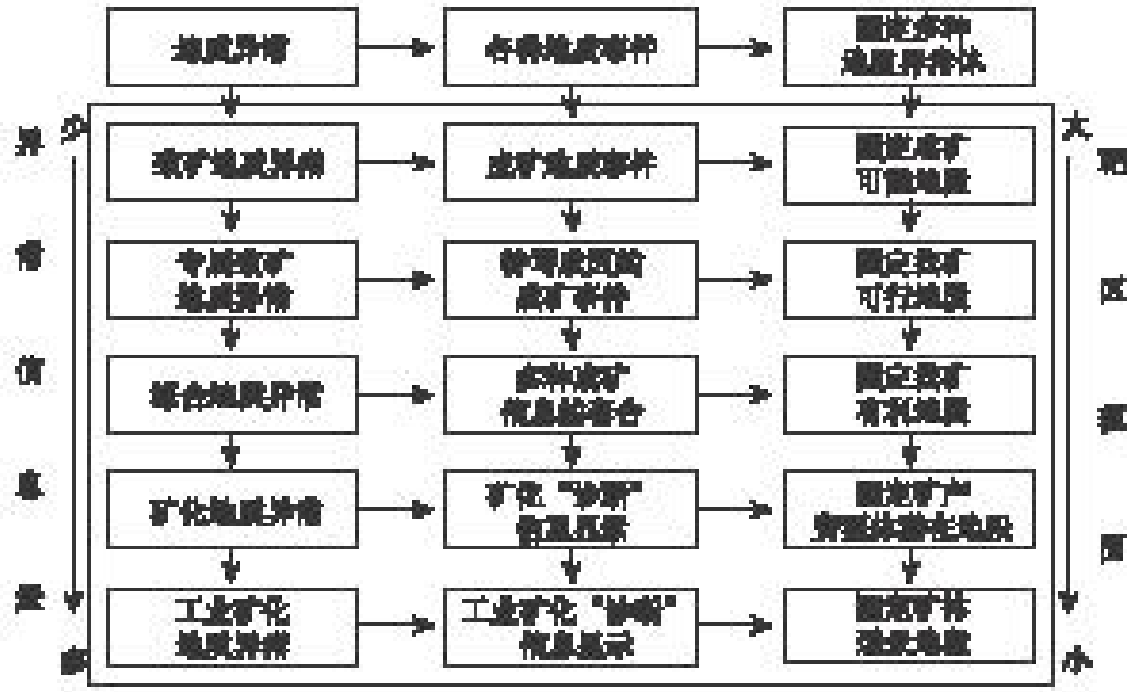


图 5-46　矿体定位预测的地质异常方法流程

1）大红山群铜矿成矿预测空间模型的建立。经过长期实践，滇中地区大红山群铜矿床预测找矿的基本规律可以总结如下:

主要地质条件　包括: ①产出的主要层位。曼岗河组中上部是大红山群铜矿床产出的层位; 老厂河组下部是中型矿床岔河式铜矿的产出层位; 底巴都组只有铜矿化未见铜矿床; ②控制大红山群铜矿床产出的构造。主要是北西向断裂及其更低级别的断裂及层间剥离构造; ③铜矿与火山一沉积建造依存关系明显。在早期中、基性熔岩喷溢时，只有小的透镜状铁矿，而无铜矿产出; 在火山喷发间歇期的碳酸盐岩沉积时期，形成了稳定的巨大的铜矿床; 在后期火山喷发阶段，出现条带状铁铜矿。

主要找矿标志　包括: ①化探 Cu 异常。Cu 异常的强度和范围与矿床的剥露程度关系密切; ②化探 Co、Ni 异常。Co 和 Ni 元素与 Cu 有较强的相关关系; ③较强的磁异常。

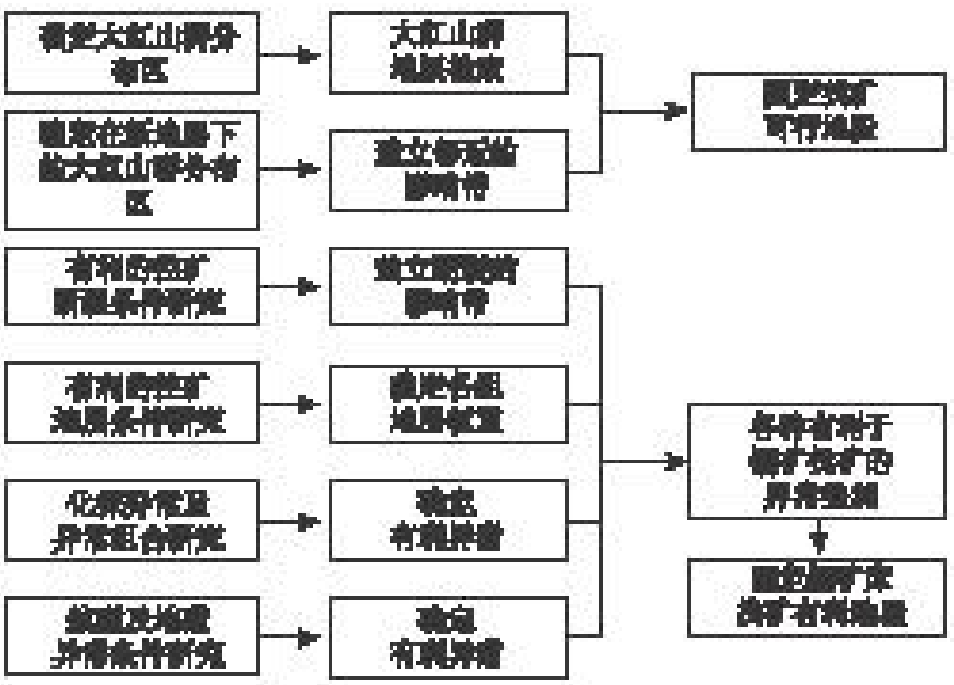


图 5-47　GIS 成矿预测的空间模型

综合上述成矿地质条件与找矿标志，可以得出如图 5-47 所示的成矿预测空间模型。由此而论，见阳裂谷地质异常区是在滇中寻找前寒武纪铜矿床的成矿可能地段；大红山群红山组、老厂河组和曼岗河组分布区，是寻找大红山群铜矿床的可行地段；而邻近大红山群地层分布区以外 1 km 的地段，是寻找同类型铜矿的潜在找矿可行地段。

2) 找矿可行地段和找矿有利地段的确定。根据成矿预测空间模型，可以利用装载了研究区各种空间信息的 GIS，通过缓冲分析和叠加分析等方法来圈定找矿可行地段和找矿有利地段。

利用缓冲分析圈定找矿可行地段　应用 GIS 作如下操作，即可在图上得到找矿可行地段：①检索出大红山群各组地层分布区；②沿大红山群边界线，以 1 km 为半径作缓冲区；③如果在上述圈定的地区内有大岩体，则扣除岩体的分布范围。

利用叠加分析确定找矿有利地段　根据找矿有利度分析得出最重要信息层的空间数据叠加，产生新的数据层。根据找矿信息量的大小，可以划分出找矿有利地段。具体工作步骤如下：①在所有的找矿标志中，找出单位矿产当量最大的那些标志，再按照一定的原则着色；②将所有选出的证据层，进行区对区合并的空间分析。分析的结果得到许多不同颜色的区，分别反映各种成矿有利因素之和，也即反映各个区有利成矿的程度。色标的数值越大，就越有利于找矿；③对所有颜色的色标作统计分析，根据找矿有利地段既要包含所有的大、中型矿床，又要使面积尽可能小的原则，确定找矿有利地段与非找矿有利地段、Ⅰ级找矿有利地段与Ⅱ级找矿有利地段的界限；④去掉属于同一级找矿有利地段的区域界线，使之成为统一的找矿有利地段；⑤统计各个找矿可行地段和找矿有利地段的面积。

在各类异常找矿有利度研究的基础上，筛选出 4 种有利的控矿因素：①红山组、曼岗河组（Pt_2h/Pt_2mg）及老厂河组（Pt_2lc）；②以 1.13 km 为半径的断层影响带（F - 1.13）；③航磁异常；④有利的化探异常（Cu-Ni 和 Cu-Co 的组合异常）。这 4 种控矿因素可分为 5 个数据层（表 5-13），其预测权重用主观赋权的方法给定。但同一控矿因素中不同预测数据层及同一数据层中不同对象的权重，则根据单位矿产当量的大小按比例给定。表中的单位矿产当量是按大红山群出露的范围计算出的，其值与在整个预测区内计算的略有不同。

表 5-13　大红山群铜矿床预测的数据层及其色标

因素	因素权重	数据层	对象	单位矿产当量 个·km^{-2}	权重（色标）
地层	30	有利地层组	Pt_2h/Pt_2mg	3.717	26
			Pt_2lc	0.542	4
断层	20	F - 1.13		0.619	20
航磁	20	航磁异常区		0.180	20
化探	30	Cu - Co		4.692	17
		Cu - Ni		2.932	13

按照表中的权值确定各个数据层的色标，在 GIS 上进行区对区合并分析后，得到 5 个数据层的叠合图。将叠合图作些归并，其标准为：找矿有利度 > 51 为Ⅰ级找矿有利地段；35 ~51 为Ⅱ级找矿有利地段。分析结果表明，Ⅰ级找矿有利地段是有利的地层因素和航磁异常，再加上具有化探组合异常或断层影响带等 3 种有利的因素，直至包括全部其余 4

种有利因素的组合；而Ⅱ级找矿有利地段是在有利于铜矿找矿的地层分布区内，有航磁异常的地段或有 Cu、Ni、Co 异常的地段，或其中的一种加上断层影响带。

3）预测结果分析。基于地质异常空间分析和“5P”地段预测，总共圈定了找矿可行地段 7 处，找矿有利地段 21 处（图 5-48）。其中，Ⅰ级找矿有利地段 6 处，Ⅱ级找矿有利地段 15 处。在大红山地区铜矿找矿可行地段内，有两个Ⅰ级找矿有利地段（面积 9 km^2）及 3 个Ⅱ级找矿有利地段（面积 20 km^2），并且已经有超大型铁铜矿床大红山铁铜矿床产出，是继续寻找大红山群铜矿的最有利地区。在撮科铜矿找矿可行地段内，也有两个Ⅰ级找矿有利地段（面积 7 km^2）和 3 个Ⅱ级找矿有利地段（面积 4.5 km^2），并有中型岔河铜矿产出，是第二个找铜矿的有利地区。阿不都、腰街、希拉河等找矿可行地段的找矿条件也比较好，同样是继续寻找大红山群铜矿较为有利的地区。

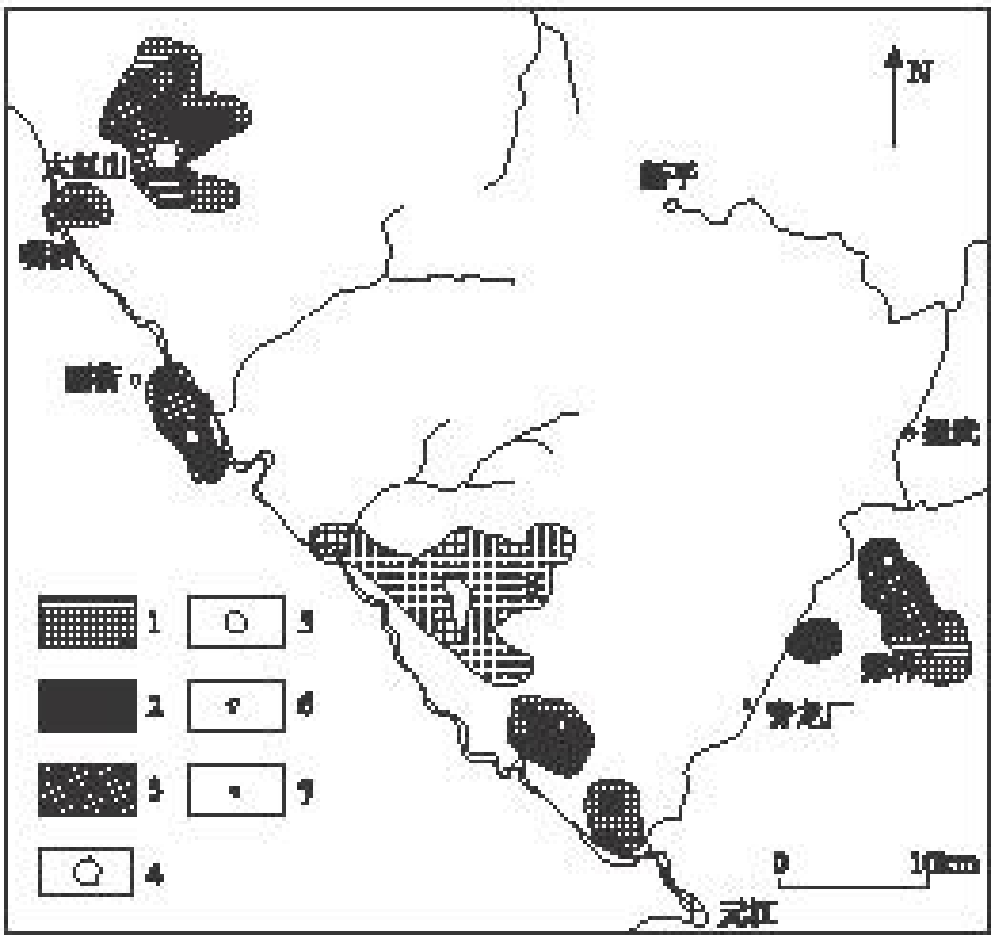


图 5-48　云南元江地区铜矿 GIS 预测图

（引自池顺都，1998）

1—找矿可行地段；2—Ⅱ级找矿有利地段；3—Ⅰ级找矿有利地段；4—大型矿床；5—中型矿床；6—小型矿床；7—矿点

实际预测的初步成果（池顺都等，1998）表明，这种方法的优越性是：①能充分地吸收定性研究的经验和定量研究的统计数据，并体现在成矿预测空间模型中；②所圈定的“5P”地段的范围，是由所选定的各种因素叠加后得出的，比较客观且形态自然；③所得到的“5P”地段面积比较准确。显然，GIS 是地质异常提取和矿产资源预测的有力工具。

三、空间信息系统在工程地质中的应用

在各类工程地质勘察与施工过程中，常常会遇到大量边坡、马道、平硐等编录图以及

录像、照片等空间图形数据。将这些空间数据管理起来，并利用这些空间信息进行工程地质条件分析和评价，是工程地质勘察与设计的一个重要内容。本节以三峡水库坝区边坡工程地质信息系统为例，说明这方面的应用。

1. 系统的结构与组成

该系统由数据采集模块、文件管理模块、图形编辑查询模块、应用专题模块、图表输出模块、系统维护模块组成，而每个模块都含有 2 个以上的子功能模块（图 5-49）。

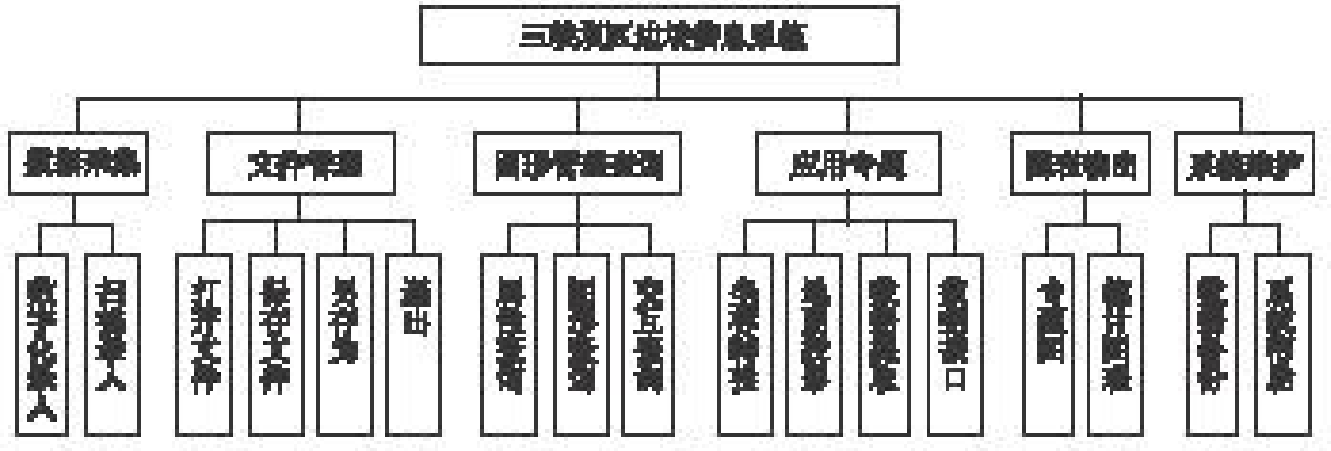


图 5-49　三峡坝区边坡空间信息系统结构 - 功能图

2. 边坡地质编录内容与系统工作流程

(1) 边坡地质编录内容

归纳起来，长江三峡水利枢纽的边坡地质编录的内容大致有如下几个大类：①边坡基本情况，起始桩号，终止桩号，坡面，坡角，控制点坐标；②岩层、岩体、岩脉和风化特征；③断裂构造、节理和裂隙；④关键块体和岩体质量；⑤水文地质。

边坡编录的记录和成果表达方式是：断裂、岩脉、裂隙卡片和边坡地质编录图。

本系统采用了面向对象的数据模型，并且将图件资料分为 12 个图层来管理：图形概述；岩体结构；风化界线；裂隙；断裂；岩性；边坡坡段说明；控制点；特定面及面边界；岩脉；块体；地下水露头。

本系统数据库设计参照的标准和技术规范包括：水利部工程地质勘察内业整理规范（SL - 78，SL - 95），三峡库坝区地质勘察图式图例及图件编制、描绘一般规定，1∶ 5 万、1∶ 1000、1∶ 2000地形图要素分类与代码（GB - 14804 - 93）。

(2) 文件命名规则与术语代码标准

1) 文件及图层的命名规则。三峡水库的边坡编录文件和图件数量庞大，为着统一管理和信息交换，系统设计者根据相关的国家标准和行业标准，制定了统一的命名规则（图 5-50）。

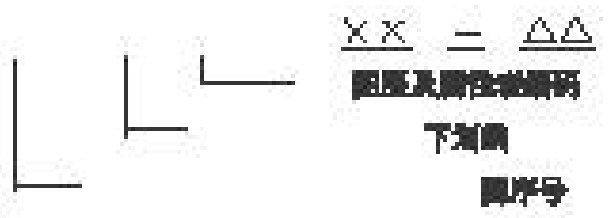


图 5-50　三峡坝区边坡信息系统的文件和图件命名规则

2) 部分属性描述标准术语。参照 SL - 78、SL - 95 等水利水电工程地质勘察内业整理规程，制定了如下描述标准：

岩性　　覆盖层：粘土、粉质粘土、粉质壤土、粉细沙、细沙、沙卵石、沙夹块球体、块球体夹沙；

基岩：闪云斜长花岗岩、细粒闪云斜长花岗岩、片岩捕虏体；

脉岩：辉绿岩脉、闪斜煌斑岩脉、花岗岩脉、伟晶岩脉；

风化分带　全（Ⅳ）、强（Ⅲ）、弱上（$Ⅱ_1$）、弱下（$Ⅱ_2$）、微（Ⅰ）；

断裂　碎裂**岩、碎裂岩、碎斑岩、角砾岩、糜棱岩；

水文地质　出水类型：线状水、滴水、有水流痕迹、无水流痕迹；

透水性分级：W>0.1　极强透水；W=0.1～0.05　较强透水；W=0.05-0.01　中等透水；W<0.01　微透水；

裂隙　倾角分级：陡倾角>60，中等倾角35°～59°，缓倾角<35°；

裂面形态：平直光滑、平直稍粗、起伏稍粗、极粗糙；

裂隙规模控制参数：结构面中点坐标、端点坐标、产状；

节理编号：以编录图编号为准。

岩体结构　整体结构（Ⅰ）、块状结构（$Ⅱ_1$）、似层状结构（$Ⅱ_2$）、次块状结构（Ⅲ）、镶嵌结构（Ⅳ）、碎裂结构（Ⅴ）、散体状结构（Ⅵ）；

岩体质量　优质岩体（A）、良质岩体（B_1、B_2）、中等岩体（C）、差岩体（D）；

3）编录图数据代码。根据编录图内容及应用功能设计的需要设计。

3. 各模块的基本功能

（1）数据采集模块

边坡编录图空间数据的采集可以采用跟踪数字化方式，也可以采用扫描加矢量化方式。图形采集模块可以完成设备初始化、数据采集和图形数据转换等项作业。其中初始化数字化仪是进行数字化的准备工作，包括设置数字化范围，数字化图幅边界点，设置数字化原点坐标，设置数字化的比例尺等。图形数据转换是指将由扫描或跟踪数字化得来的图形坐标，转换到三峡坝区目前所使用的坐标系，或是将由其他系统（如 AUTOCAD）采集的图形数据转换为本系统所用的图形数据格式。

（2）图形接边模块

由于各种编录图的数字化输入是分幅进行的，数字化后相邻图幅边界的几何位置与属性难免会出现不吻合，需要进行图形拼接处理。

模块中设置了自动化与人机交互式两种接边处理方式。人工交互式接边是由用户选定准备连接的图形，再通过屏幕对话方式逐一对交接处的空间实体进行连接处理。自动化接边则是由计算机对所需连接的所有图形的相对空间位置及属性参数进行加权计算，再选择权值最大的图形与之进行连接。裂隙接边参数有位置、倾向、倾角等。

（3）文件管理模块

主要用于执行建立、打开、保存编录图工程文件及点线面图形文件。它可以直接操作各建筑部位的工程文件（*.apr）。

（4）图形编辑查询模块

图形编辑查询模块提供多种图形属性查询方式，其中包括条件查询、SQL 语言查询、围限查询、图形查询、单一查询等。

（5）专题应用模块

专题应用模块是本系统的核心模块，用于编录资料管理、分析。它包含了坐标转换（图 5-51）、投影变换（图 5-52）、数据接口、裂隙连接等次一级模块。

（6）图表输出模块

该模块的功能是将分析、查询的最终结果以统计图、统计表格、专题图等方式输出，也可以生成直方图、饼状图、色块图和玫瑰花图等。

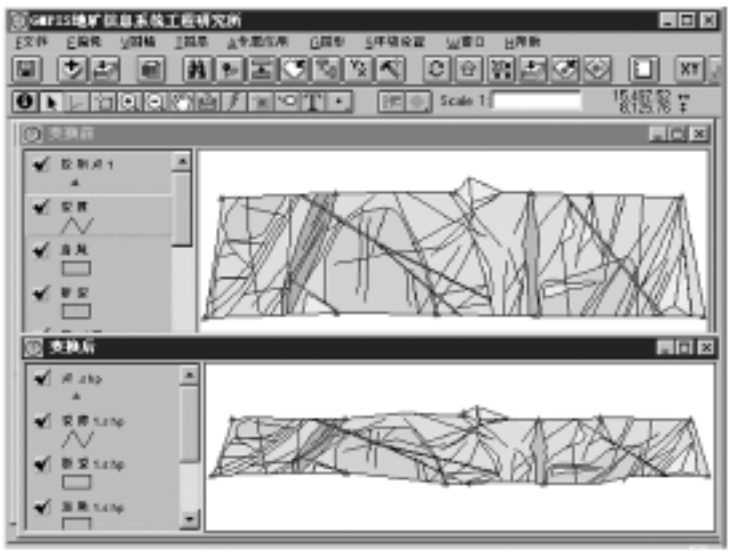


图 5-51　边坡编录图坐标变换
(下图为新坐标)

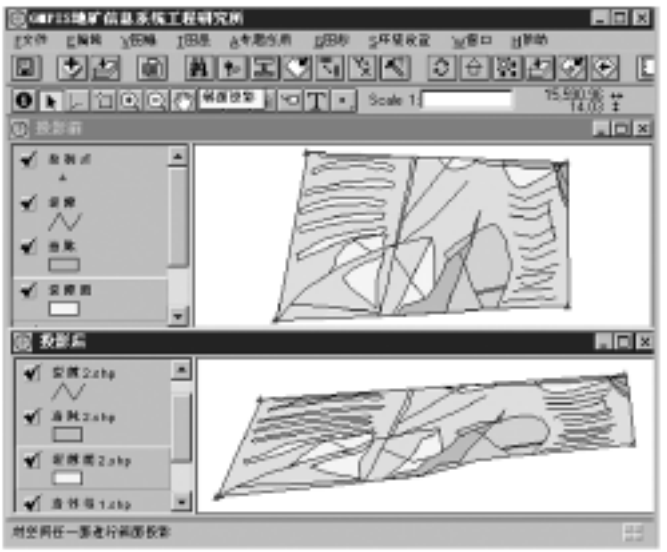


图 5-52　边坡编录图投影变换
(下图为投影结果)

(7) 系统维护模块

该模块用于保证系统正常运行，并为系统的使用者提供数据备份、系统安全性设计、系统帮助等服务。其中，数据备份能实现全部或部分数据转移和恢复；系统安全设计包括操作级别、访问权限、口令等的规定；系统还为用户提供操作模块的在线帮助信息。

思　考　题

1. 阐述资源空间信息子系统在资源信息系统中的地位、作用及其结构组成。
2. 理解空间数据库在资源空间信息系统中的作用及意义。
3. 试比较空间数据库中几种数据结构模式的异同。
4. 空间数据中可能会存在哪些错误（误差），针对这些错误需进行怎样的校正?
5. 为什么要进行坐标变换及投影变换？它们有何异同?
6. 试举例说明各种空间分析方法在金属成矿和油气成藏分析中的应用。
7. 矢量数据的标准转换格式主要有哪些?
8. 栅格数据的标准转换格式主要有哪些?

第六章　地矿资源图件机助编绘子系统的建立和应用

大量的地矿资源勘查与开发工作成果，要用图件清晰、正确地表达出来，地质工程师在分析复杂的地质事件和过程时，地矿资源图件又是他们得力的助手和时空向导，因此，可以说“图纸是工程师的语言”。计算机辅助制图和计算机辅助设计技术的应用，对于提高基础地质调查、矿产资源勘查和工程地质勘察工作的质量和效率具有重大的意义。

第一节　地矿资源图件机助编绘子系统概述

地矿资源图件计算机辅助编绘子系统是地矿资源信息系统的重要组成部分，它体现了计算机图形处理技术、资源数据库子系统、地质制图方式的有机结合。这也正是地矿资源图件计算机辅助编绘子系统与通用计算机图形系统的不同之处。

一、图件编绘与计算机辅助设计技术

1. 计算机辅助设计

图形是描述客观事物的空间位置、几何形状和形态的一种方式。与语言、文字或数据相比，图形具有直观、形象、易于理解的特点。从信息论的角度看，图形是承载和传送信息的一种媒体，它与语言、文字、数据、声音等一样，是人们交流信息的一种不可缺少的工具。人们获取信息最自然最直接的方式是采用眼 - 脑识别机构，观察逼真的图形，如同置身于现实的空间世界中，可以快速、高效而准确地理解和把握各种客观事物。使用计算机来实现信息符号化和图形化的学科，是计算机图形学。

ISO 数据处理字典将计算机图形学定义为：通过计算机完成数据与图形之间相应转换的方法和技术。来源于计算机图形学的各种图形处理技术，是关于图形的输入、存贮、加工、变换、显示和输出的理论、方法和工具。

从技术方法角度看，图件编绘系统所采用的核心技术为计算机辅助设计技术。计算机辅助设计（Computer Aided Design，简称 CAD）是指将计算机作为基本工具，为设计工作提供广泛的支持，如数据处理、逻辑判断、数值计算、图形显示、图形编辑和自动绘图等。之所以称为“计算机辅助设计”，是因为它不能完全代替人的工作，只有把人的直观感知能力、综合判断能力、经验、创造性和计算机高速度、大容量、高精度的处理能力结合起来，才能产生很好的效果。将计算机图件编绘系统与科学计算及分析系统等紧密结合起来，可构成计算机辅助设计系统。在发展初期，CAD 技术多用于机械、电子、建筑设计等行业，现正迅速地向各相关领域渗透、扩展，其一般工作流程如图 6-1 所示。

CAD 技术的优越性在于：①减少绘图劳动量；②缩短工作阶段的周转时间；③降低直接设计费用；④提高精度和可靠性；⑤易于修改和补充；⑥易于实现标准化。地矿资源图件计算机辅助编绘子系统的开发，可以在通用的 CAD 系统或 GIS 平台上进行。

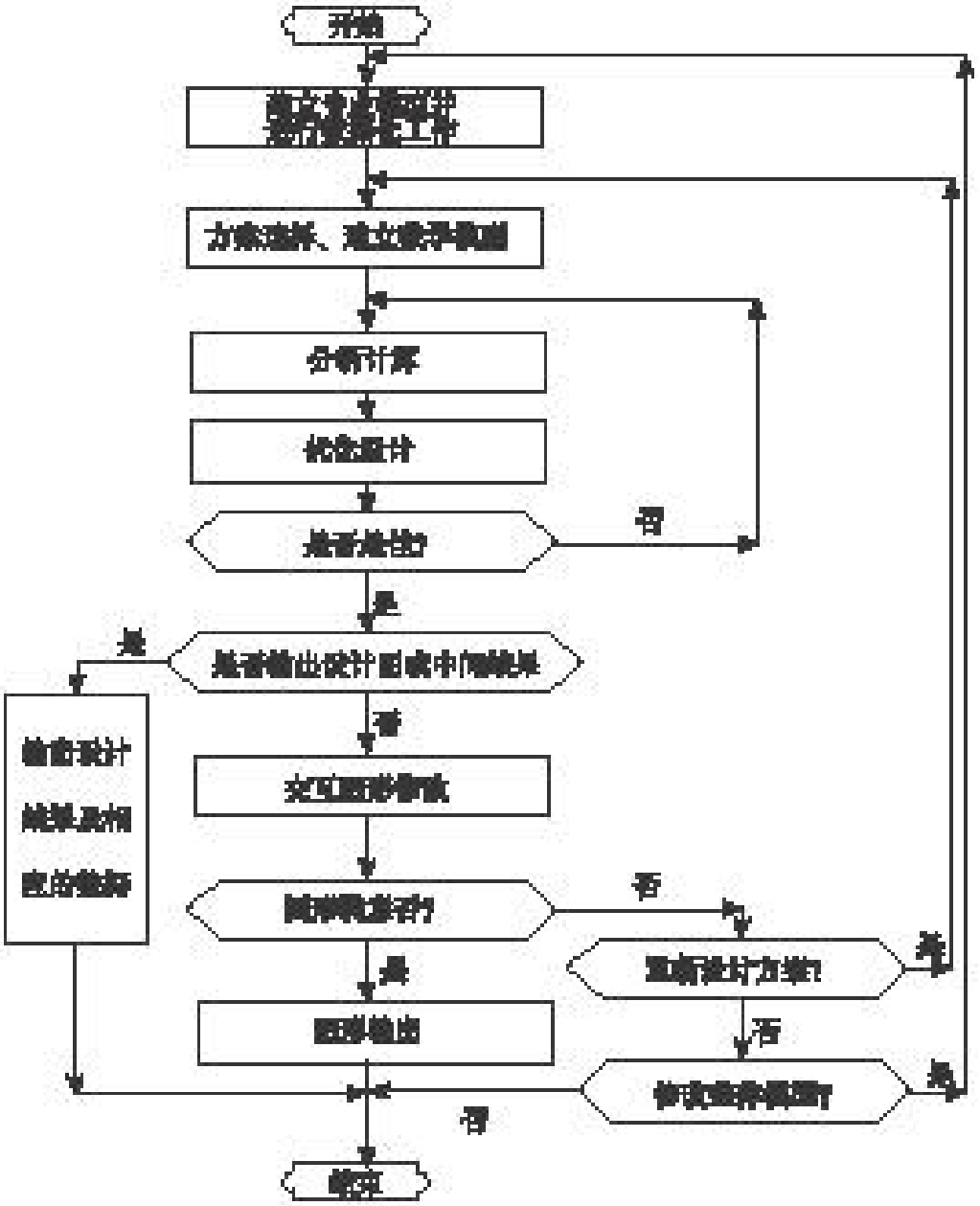


图 6-1　地矿资源图件计算机编绘子系统工作流程图

2. 地矿资源图件编绘子系统的开发前提

地矿资源图件的特点是: 图形复杂、线条随机变化、花纹繁多、推测数据多、精度要求高等。在系统开发时，地质工作者和计算机工作者要密切配合，并注意如下几点:

1) 数据来源应当是资源勘探与开发的基础原始资料，如钻孔、测井等，才能保证图形的精度和正确性，实现灵活的比例变换等功能。因此，需要地矿点源主题数据库的支持。

2) 应能将图形系统同地质分析及评价决策系统连接起来，实现地质基础资料综合应用和信息的有效提取，以可视化的直观形式协助地质专家分析、判断各种地质现象。

3) 图形系统开发，是在已有的通用绘图系统上进行二次开发，还是从头开始进行基础核心系统开发，应根据工作需要实时制定。所谓二次开发是指在功能较为齐全并且具有开放性和可扩充性的基础图形系统上，进行再开发，扩充系统功能和完善系统操作。

4) 计算机制图与编图之间存在着明显的差别。地矿资源勘查开发的计算机制图与常规的地矿资源制图相应，是指各类地矿资源图件（平、剖面图等）的清绘、编辑、制板、印刷和出版的全套工艺流程。整个资源制图过程由制图专业技术人员控制。为制图提供的

底图则是由地质人员根据实际地质资料设计完成的，称为编图。实现计算机编制或辅助编制地矿资源图件，是资源图件计算机辅助编绘系统所追求的目标。

3. 资源图件编绘子系统的开发环境

图件编绘系统的建立，往往对计算机软、硬件环境有比较高的要求。

(1) 系统的硬件环境

在当前条件下，可采用以微型计算机为基础的硬件系统。常见的硬件配置如下:

主机: IBM PC 系列及兼容机，32 位 CPU (Pentuim 或更高级型号)，内存容量 32M 字节。硬盘容量不小于 2G 字节，带数字协处理器。

显示器: 彩色高分辨率屏幕配 SVGA 型显示卡。

定点设备: 鼠标或数字化仪 (A1 以上幅面)。

打印机: 激光打印机。

绘图仪: 喷墨绘图仪 (A0 幅面)。

以上为构成绘图系统必要的设备。建议有条件的单位配置高档一些的绘图设备，以加速图形生成和提高图件生产质量。

(2) 系统的软件环境

从软件角度看，开发环境应当包括微机操作系统、基础图形系统和数据库系统。在一般情况下，操作系统可采用 Windows 98、Windows NT 以及它们的升级产品。基础图形系统可采用交互型自动绘图系统 AutoCAD 或各种成熟的 GIS，如 ArcGIS、MapGIS 等；数据库系统则可使用 Visual Foxpro、SQL 和 Oracle 等。本书主要介绍基于 AutoCAD 的开发与应用。

(3) 系统的语言工具

CAD 系统所采用的语言工具，大致分为汇编语言、高级语言和专用图形语言等三类。

1) 汇编语言。汇编语言是编制图形软件最基本的工具，常用于编写软件系统中的核心模块和外部设备的驱动程序。利用汇编语言编制的软件高效、灵活，但这种语言不易掌握。

2) 高级语言和图形函数库。高级语言 (例如 C++ 语言) 加上图形函数库是理想的图形编程工具。为了适应图形软件开发需要，各类高级语言 (如 Visual BASIC、Visual C++ 等等) 均增强了图形处理能力。但就对 CAD 系统进行二次开发而言，以 C++ 语言为佳。其优势在于: ① 既具有高级语言易写易懂的特点，又具有速度快、通用性强的长处，因而开发周期短、灵活性好; ② 带有功能强大的图形函数，易于实现复杂应用系统开发; ③ 可充分发挥已有软件的成果。

3) 专用的图形语言。利用专用的图形语言系统可以方便快速地建立用户化的应用图形系统。许多通用性绘图系统，如著名的 AutoCAD 和 MicroStation 等，都设置有专用的图形语言 (AutoLisp 等)。这类图形语言功能强，可与 C++ 语言相通，用户可以方便地用其对该系统进行二次开发。

二、资源图件编绘子系统的功能与结构

1. 资源图件编绘子系统的设计目标

在某种意义上，计算机辅助编绘系统可以看作是地矿资源信息系统的一个功能处理子系统。编图系统应当具有结构合理、经济实用、符号统一、图例标准、操作简单、快速高

效等特点，还应当能与点源主题数据库对接。概括起来，机助编图系统功能设计的目标是：图件种类齐全、系统操作方便，数据准备简单，输入、输出多样，图例花纹标准和图件内容精确。

(1) 图件编绘种类

地矿调查、勘探与评价面临的图件种类繁多，归纳起来有柱状图、剖面图、平面图、曲线图和立体模式图等五大类。根据具体用途，可以细分出数十种基础图件和分析图件。一个完善的资源图件计算机辅助编绘系统，应当能够全面实现这些图件编绘的计算机辅助化。

(2) 系统操作功能

为了使地质人员能迅速掌握并利用本系统编图，系统操作采用汉字菜单选择方式和人机对话方式，将高级语言自动编图程序与 CAD 软件系统联结起来，使用户能够自由地对计算机自动生成的图件进行补充设计、修改和编辑。

(3) 数据准备功能

本系统的数据管理采用随机数据点的提供方式。数据文件既可以由数据库检索后自动生成，也可以直接由文本编辑系统（例如 NOTEPAD 等）编辑。

利用数据库对数据实行集中管理的优点在于：①减轻应用程序员对所有标准数据的管理任务；②易于实施标准化；③应用程序员可不考虑系统性能方面的问题；④减少数据的冗余；⑤降低数据的不一致性；⑥共享所存贮的数据；⑦保持数据的完整性；⑧易于执行安全性措施；⑨为系统的随机查询和其他应用打下基础。

(4) 输入、输出功能

输入、输出的多样化历来是计算机绘图系统的用户所期待的目标之一。输入的多样化是指可以用各种方式输入数据，包括文字、代码、数值、图像和声音等。本系统设计了接收文字、代码、数值（通过键盘输入）和曲线图形（通过数字化仪或数字化板输入）等几种输入方式。

输出的多样化一般指可用各种设备显示、打印或绘制出文字、表格、图像和图件等结果。但在本系统中，输出特指图形方式输出。本系统既可直接在各种绘图仪上绘制各类地矿资源图件，也可直接在高分辨显示器上显示各类地矿资源图形，还可以生成图形交换文件，并且实现了图形和汉字的同时输出，以及图形比例尺的自由选择。

(5) 图例、花纹的规格化

经验表明，各种图件的图例、花纹必须统一、规范。这不仅有利于信息的广泛交流，而且有利于应用者和开发者提高工作效率。近年来，各部门在图件标准化方面做了许多工作，针对不同部门所开发的专用地矿资源图件机助编绘系统，通常都采用各自的部颁标准来制定可供绘图程序调用的图例花纹库。为了提高共享性，今后将会逐步采用行业标准。

(6) 图件内容的精确性

资源勘探成果图件具有指导勘探和生产的效能，图件内容必须准确和可靠。

机制图件的精确性主要表现在已知点的定位、未知点的推测以及比例精度和闭合精度等几个方面。已知点的定位内容，包括已知地质点点位上的坐标、高程、矿层厚度等参数值的外推，其可靠性主要依赖于数学模型的合理性。某些图件，例如矿层厚度和矿层底板高程的等高线，一旦数学模型出现偏差，将导致钻探工作、储量计算和开采掘进的失误。

为了满足精度要求，在进行等值线程序设计时，通常采用模拟人工方式的三角网机械插值法。

已知点的定位精度、图形比例精度及闭合精度，一方面取决于绘图机本身的性能指标，另一方面取决于图形输出的顺序调节。为了减少累积误差，对于笔式绘图机的输出设计，通常采用格架性图形信息优先处理、优先输出的策略。

2. 资源图件编绘子系统的结构组成

在逻辑结构上，一个完整的地矿资源图件计算机辅助编绘系统应当包括数据采集、数据存贮和图件编绘三个大的模块或子系统（图 6-2)。其中，计算机图件编绘模块可以定义为利用计算机实现图形和符号生成、变换、显示的人机交互系统。其基本组成如图 6-3 所示。该系统的硬件部分包括计算机和人机交互设备（信息输入、输出设备)；软件部分包括图形数据库、应用图形系统和基础图形系统。

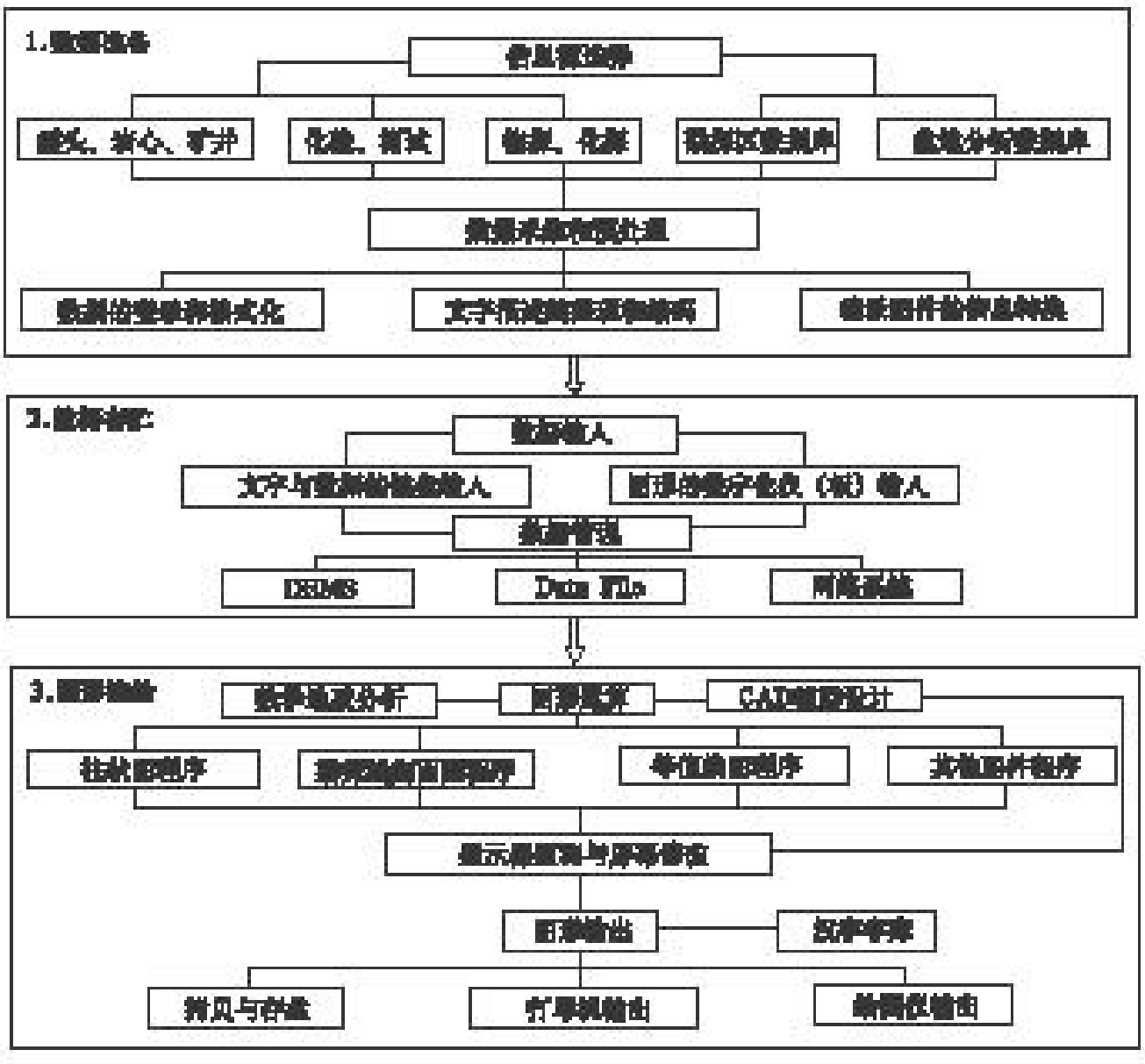


图 6-2　资源信息系统计算机编绘子系统逻辑结构框图

基础图形系统是由计算机厂或专业软件公司提供的一种通用绘图系统，具有多种基本绘图功能。其中最基本是绘制直线，因为任何曲线都可以由充分小的直线段近似构成。

应用图形系统则是图形数据库和基础图形系统之间的桥梁，完成由数据到图形的转化任务。应用图形系统包括图形专业建模模块、图形几何建模模块和图形计算模块等几个部

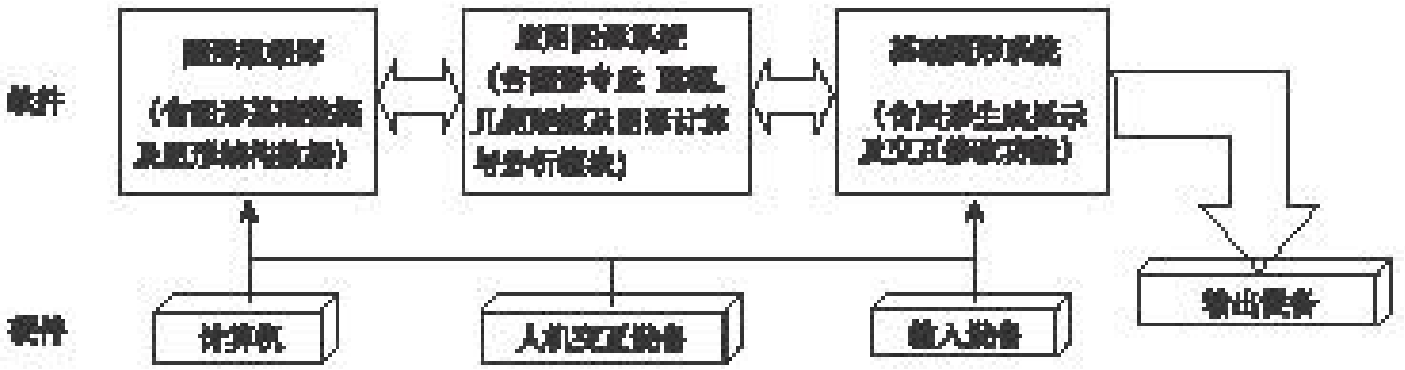


图 6-3　计算机图件编绘系统的基本组成

分，一般由高级语言（PASCAL、FORTRAN、BASIC、C 或 C^{++}等）写成。图形专业建模模块的任务是以专业知识、专业规范和专业发展为依据，以专业成果表达为目的，建立专业图形模型，指导其他相关模块的建立。图形几何建模模块的任务是，正确地描述对象的几何形体并建立其相应的数学模型。图形计算与分析模块的基本任务是：①与数据库建立接口；②将数据库中与几何形体有关的可视化数据进行计算，转化成二维或三维图形，实现几何变换，包括投影变换、坐标变换等；③利用基础图形系统将上述转化后的图形绘制出来；④对几何模型进行分析和综合。

第二节　地矿资源图件机助编绘子系统的开发基础

通用化的 CAD 系统在图形花纹较为规则的机械、电子、建筑和纺织等领域，可以直接应用；在图形、花纹较为复杂且不规则的地矿资源勘查领域，直接应用较为困难，需要做大量的二次开发工作。这种二次开发的系统设计，应遵循由地质模型（图件型式）到数学模型（图形计算），再到计算机模型（程序设计和调试）这样一个过程。本节将在介绍图形表示与生成的基本原理的基础上，重点介绍图件编绘系统的几个关键子程序的内容和开发方法。

一、图形的表示与生成原理

完成从数据到图形的转换，涉及一系列的图形运算。

与机械、建筑等行业的图件相比，地矿行业的图件机助编绘难度更大。前者以规则的几何图形为主，且为实物的具体描写；地矿资源图件主要由不规则任意曲线和复杂的花纹系统组成，且地质工作包含着许多判断、推测的成分。因此图形算法较为复杂。

1．点的表示

所有的图形基本上都可以看作是点和线的集合，而任何线又都是点的集合。在二维空间里，任一点都可用坐标来表示，形成二维向量（X，Y）；在三维空间里，任一点都可用一组三维向量（X，Y，Z）来表示。图形中的每一个点，相对于局部坐标系来说，都是一个位置向量，都可以作为一个数字矩阵储存于计算机内。确定一个点的位置，可以用绝对坐标或相对坐标。绝对坐标指相对于坐标原点的坐标，相对坐标是指相对于另一个点的位移量。

用计算机软件和硬件，可以模拟和跟踪点的运动轨迹，将产生的直线、曲线或图形输出。

2. 曲线的描述与生成

(1) 曲线的描述

曲线的数学描述，可以有如下三种形式:

1) 显函数描述。显函数一般表示为 $y = f(x)$ 的形式，例如单位圆可表示为 $y = \pm\sqrt{1 - x^2}$。这种描述形式清晰明确，在计算机中容易实现，但难以表示封闭的或多值的曲线。在多数情况下，要想得到曲线的显函数描述形式是困难的。

2) 隐函数描述。隐函数一般表示为 $f(x, y) = 0$ 的形式，例如单位圆可以表示为 $x^2 + y^2 = 1$。这种描述形式不能直观地表达因变量与自变量的关系，而且每描述一个点都必须解方程 $f(x, y) = 0$，这就使得计算和作图都非常困难。

3) 参数方程描述。一般表示为

$$x = f(t), y = g(t) \quad (a \leq t \leq b)$$

例如，单位圆可表示为

$$x = \cos(t), y = \sin(t) \quad (0 \leq t \leq 2\pi)$$

参数方程描述形式具有显著的优点，在计算机辅助几何设计中得到广泛应用。

(2) 曲线的生成

曲线的生成，是指求解曲线的数学表达式进而模拟和追踪点的运动轨迹，然后显示出图形的过程。这一过程通常按如下三种问题的处理方式实现:

1) 给定模型问题。对于曲线，特别是规则曲线的生成，通常是根据总体设计意图和具体设计要求来给定数学模型，然后对其求解，这就是给定模型问题。

2) 插值问题。在曲线设计中常常会遇到有些曲线函数是未知的情况，而已知的仅仅是一组有序的型值点。通过这些型值点构造光滑曲线，这就是插值问题。通常可以采用拉格朗日插值、埃尔米特插值或样条插值等方法生成这类曲线。

3) 逼近问题。有时给定的数据点本身就含有不确定的成分，这时并不要求构造的曲线一定通过给定的型值点，而是要求曲线在某种定义下最好地逼近给定的全部型值点，这就是逼近问题。通常可以采用最小二乘法、Bezier 法或 B 样条函数等来生成这类曲线。

3. 圆弧的表示与生成

下面以简单的二次曲线——圆弧为例，说明一般曲线的表示与生成。

圆弧的参数与非参数形式的比较如图 6-4 所示。

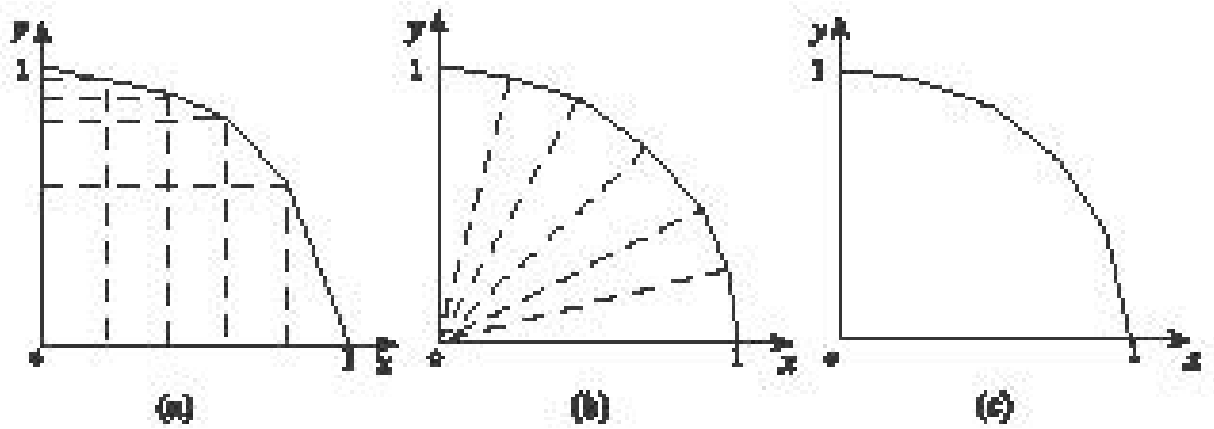


图 6-4　圆弧的表示与生成

(1) 圆弧的非参数描述形式

圆弧的非参数描述形式是: $y = \pm \sqrt{1 - x^2}$ $(0 \le x \le 1)$, 如图 6-4 (a) 所示。图中, 圆弧上的点是用相等的 x 增量 Δx 得到的。对于相等的 Δx, 相应的 y 增量 Δy 不同, 因而产生的每个弧长是不相等的。显然, 在整个象限的大约 2/3 部分, 产生的图形是粗糙的折线。此外, 由于每一步都要做开平方运算, 计算效率很低。

(2) 圆弧的参数描述形式

$$x = r\cos\theta,\ y = r\sin\theta,\ r = 1 \quad \left[0 \le \theta \le \frac{\pi}{2}\right]$$

如图 6-4 (b) 所示, 图中圆弧上的点是用相等的增量 $d\theta$ 得到的, 即

$$x_{n+1} = r\cos(\theta_n + d\theta),\ y_{n+1} = r\sin(\theta_n + d\theta)$$

显然, 得到的弧长是相等的, 且比较逼近圆弧。这种算法能产生良好的圆弧图形。但由于在每个点都必须计算三角函数值, 计算效率也是很低的。

(3) 圆弧的另一种参数描述形式

$$x = \frac{1 - t^2}{1 + t^2},\quad y = \frac{2t}{1 + t^2},\quad t = \mathrm{tg}\frac{\theta}{2} \quad \left[0 \le \theta \le \frac{\pi}{2}\right]$$

如图 6-4 (c) 所示。这种算法只需要很少的计算时间, 但从图中可以看到, 弧长并不完全相等, 图形质量不够理想。

(4) 圆弧的参数描述改进形式

为了在圆弧的参数描述形式中不再每步计算三角函数值, 又得到较理想的圆弧, 可将描述形式进行适当变换, 即由和角公式

$$\cos(\theta + d\theta) = \cos\theta\cos d\theta - \sin\theta\sin d\theta$$
$$\sin(\theta + d\theta) = \cos\theta\sin d\theta + \sin\theta\cos d\theta$$

可得

$$x_{n+1} = r\cos(\theta_n + d\theta) = r\cos\theta_n\cos d\theta - r\sin\theta_n\sin d\theta$$
$$y_{n+1} = r\sin(\theta_n + d\theta) = r\cos d\theta_n\sin d\theta + r\sin\theta_n\cos d\theta$$

所以

$$x_{n+1} = x_n\cos d\theta - y_n\sin d\theta$$
$$y_{n+1} = x_n\sin d\theta + y_n\cos d\theta$$

由于 $d\theta$ 为常数, 故 $\cos d\theta$ 和 $\sin d\theta$ 亦为常数。用这种高效率的算法得到的点的间隔是相等的, 描绘出来的圆弧质量很好, 与图 6-4b 相同。

下面给出按第四种算法画图的通用子程序 (FORTRAN 语言算法示例)。参数含义: R 为半径; H 和 V 分别为圆心的 X 和 Y 坐标; N 为描述点数; X 和 Y 为圆弧上各点坐标的 N + 1 维数组, P 为 θ 的增量。

```
SUBROUTINE CIRCLE( R, H, V, N, X, Y)
DIMENSION X( N + 1) , Y( N + 1)
P = 2* 3. 14159 / N
C = COS( P)
S = SIN( P)
X( 1) = H + R
```

```
      Y( 1) = V
      DO 15 I = 2, N + 1
      X( I) = H + ( X( I - 1) - H) * C - ( Y( I - 1) - V) * S
      Y( I) = V + ( X( I - 1) - H) * S + ( Y( I - 1) - V) * C
15 CONTINUE
      RETURN
      END
```

4. 二维图形变换

所谓图形变换，就是一个几何图形的每一个点都变成另一个图形的确定的点。图形变换在图形显示和自动绘制中都是十分重要的。目前几种基本地矿资源勘查图件的表现形式都属于二维范畴，所以在此仅讨论二维图形变换及与之有关的二维坐标变换问题。

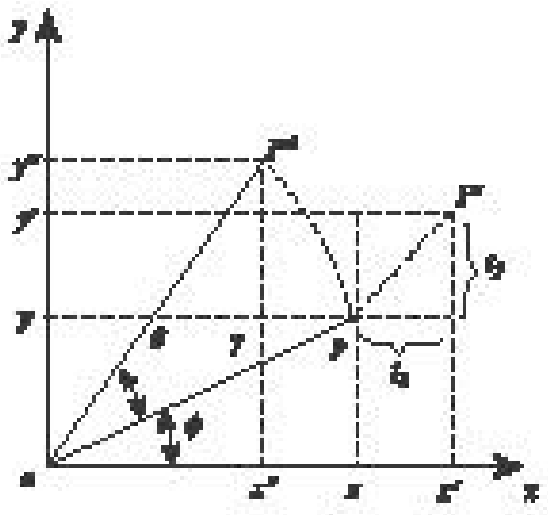


图 6-5　二维坐标平移与旋转

(1) 平移、放缩及绕原点旋转

图 6-5 表示二维坐标下图形平移和绕原点旋转的情况。

1) 平面上的某个点 P (x, y) 平移到新的点 P′ (x′, y′)，在两个坐标轴方向上的移动距离分别为 t_x 和 t_y，则平移后的新坐标为:

$$x' = x + t_x; \quad y' = y + t_y \tag{6-1}$$

2) P (x, y) 点绕原点 o 旋转 θ角到新的点 P″(x″,y″)。则旋转后的新坐标为:

$$x'' = x\cos\theta - y\sin\theta, \quad y'' = x\sin\theta + y\cos\theta \tag{6-2}$$

这里规定，绕原点反时针旋转时，θ为正。

3) 图形放缩。新点的坐标为

$$x^* = ax; \quad y^* = ey \tag{6-3}$$

式中，a 和 e 分别表示 x 和 y 坐标轴方向上的放缩比率。当 a (或 e) > 1 时，则为放大；当 0 < a (或 e) < 1 时，则为缩小。当 a = e 时，则以原点为中心，变换后的图形与原图形相似，称为等比例变换；若 a ≠ e 时，则成为非相似图形，称为不等比例变换。

(2) 二维变换的矩阵方法和齐次坐标系

一般说来，图形的变换往往需要经过上述各种变换组合而成。前面所使用的代数表达形式不利于在计算机上进行组合变换。解决这个难题的方法是采用矩阵表达方式。

1) 变换矩阵的选择。因为在二维平面变换中，点的位置是用向量 [x　y] 表示的，所以可采用二阶的基本变换矩阵 $T = \begin{bmatrix} a & d \\ b & e \end{bmatrix}$。坐标变换的表达式为

$$[x^* \quad y^*] = [x \quad y]\begin{bmatrix} a & d \\ b & e \end{bmatrix} \tag{6-4}$$

显然，绕原点的旋转 (式 6-2) 可以表示为

$$[x^* \quad y^*] = [x \quad y]\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{6-5}$$

采用这种二阶矩阵可以进行放大和旋转，但无法表示最常见的平移变换。为此往往采用齐次坐标系，即在表示坐标的 [x y] 中增加第三分量，成为 [x y 1]。这时，变换可表示成:

$$[x \quad y \quad 1]\begin{bmatrix} a & d & m \\ b & e & n \\ i & j & s \end{bmatrix} = [ax + by + i \quad dx + ey + j \quad mx + ny + s] = [x' \quad y' \quad w]$$

或者表示成:

$$[x^* \quad y^* \quad 1] = \left[\frac{x'}{w} \quad \frac{y'}{w} \quad 1\right]$$

2) 二维几何变换。常见的二维图形变换有平移变换、放缩变换和旋转变换。这些变换的数学表达式在形态上都是相同的，但其变换矩阵差异显著。

①二维平移变换

表达式:

$$[x^* \quad y^* \quad 1] = [x \quad y \quad 1]\ T_T$$

其中，平移矩阵:

$$T_T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ t_x & t_y & 1 \end{bmatrix} \tag{6-6}$$

②二维放缩变换

表达式:

$$[x^* \quad y^* \quad 1] = [x \quad y \quad 1]\ T_S$$

其中，放缩矩阵:

$$T_S = \begin{bmatrix} a & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{6-7}$$

③绕原点旋转变换

表达式:

$$[x^* \quad y^* \quad 1] = [x \quad y \quad 1]\ T_R$$

其中，旋转矩阵:

$$T_R = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{6-8}$$

上述变换都具有可逆性，即 $[x \quad y \quad 1] = [x^* \quad y^* \quad 1]\ T^{-1}$。

④绕任意点 C (x_c, y_c) 旋转变换

点 P (x, y) 绕平面上任意点 C (x_c, y_c) 旋转 (转角为 θ) 的变换，可通过下列步骤实现: 首先将原坐标原点平移变换到点 C，并作为新坐标原点，$P' = PT_T$，这时 $t_x = -x_c$，$t_y = -y_c$; 再围绕新原点 P 旋转变换，$P'' = P'T_R$; 然后通过平移，恢复原来的坐标系统，$P^* = P''T_T^{-1}$。这里，T_T 和 T_R 分别为平移和旋转矩阵。

3) 组合变换。上述绕任意点 C (x_c, y_c) 的旋转变换，事实上是一种组合变换，可采用矩阵积来代替三次分别的变换，即:

$$P^{*} = PT_T T_R T_T^{-1} = PT \tag{6-9}$$

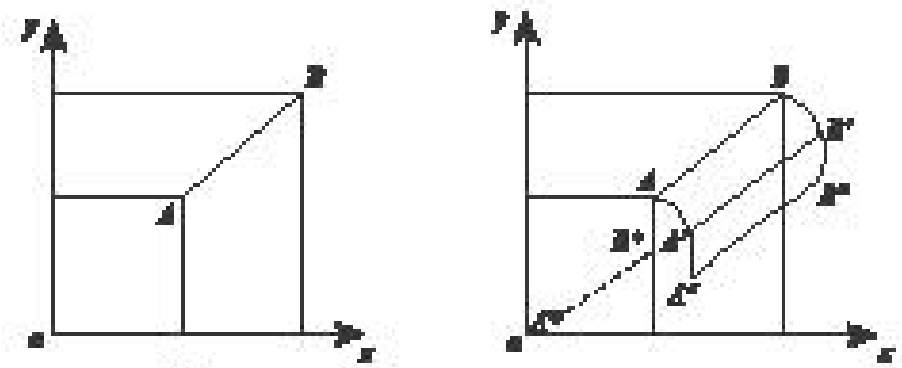
式中，变换矩阵:

$$T = T_T T_R T_T^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -x_c & -y_c & 1 \end{bmatrix} \begin{bmatrix} -\cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -x_c & -y_c & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ -x_c(\cos\theta - 1) + y_c\sin\theta & -x_c\sin\theta - y_c(\cos\theta - 1) & 1 \end{bmatrix} \tag{6-10}$$

式（6-10）**中，矩阵积** $T_T T_R T_T^{-1}$ **的一次运算可代替三次矩阵运算，可进行三种变换。矩阵方法的这个特点在对由许多点组成的图形进行组合变换时，可以大大节省计算时间，表现出明显的优越性。**

例：**分析并且表达出图** 6-6 **所示线段** AB **的旋转**（**绕原点**）**、平移和缩小**（**放缩比率** a = e = s）**组合变换。**



(a)　　(b)

图 6-6　**二维组合变换示例**

采用如下旋转矩阵:　　**平移矩阵**:　　**放缩矩阵**:

$$T_R = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad T_T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ t_x & t_y & 1 \end{bmatrix} \quad T_S = \begin{bmatrix} S & 0 & 0 \\ 0 & S & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

于是，总变换矩阵为:

$$T = T_R T_T T_S = \begin{bmatrix} S\cos\theta & S\sin\theta & 0 \\ -S\sin\theta & S\cos\theta & 0 \\ St_x & St_y & 1 \end{bmatrix}$$

利用变换矩阵 T，**可将两端点** A **和** B **分别变换成新的端点** A^{*} **和** B^{*}，**即**

$$[x_A^{*} \quad y_A^{*} \quad 1] = [x_A \quad y_A \quad 1]T; \quad [x_B^{*} \quad y_B^{*} \quad 1] = [x_B \quad y_B \quad 1]T$$

连接端点 A^{*} **和** B^{*}，**便可获得变换后的线段** $A^{*}B^{*}$，**如图** 6-6（b）**所示。**

4）**二维坐标变换。实际上是将点的齐次坐标向量**（x，y，1）**与变换矩阵** T **相乘，以得到新的坐标向量**（x^{*}，y^{*}，1）。**下面给出一个典型的坐标变换的计算机程序**（FORTRAN **程序示例**）。

```
      DIMENSION T (3, 3), AB (2, 3), C (3)
      WRITE (*, '(A)') 'AX, AY: '
      READ (*, *) AB (1, 1), AB (1, 2)
      WRITE (*, '(A)') 'BX, BY: '
      READ (*, *) AB (2, 1), AB (2, 2)
      WRITE (*, '(A)') 'TX, TY, THITA, S: '
      READ (*, *) TX, TY, THITA, S
      AB (1, 3) =1
      AB (2, 3) =1
      THITA = THITA* 3. 14159/180
      T (1, 1) = S* cos (THITA)
      T (1, 2) = S* sin (THITA)
      T (1, 3) =0
      T (2, 1) = S* sin (THITA)
      T (2, 2) = S* cos (THITA)
      T (2, 3) =0
      T (3, 1) = S* TX
      T (3, 2) = S* TY
      T (3, 3) =1
      DO 5 K = 1, 2
        DO 15 I = 1, 3
          C (I) =0
            DO 25 J = 1, 3
              C (I) = C (I) + AB (K, J) * T (J, I)
25        CONTINUE
15      CONTINUE
         DO 35 I = 1, 2
           AB (K, I) = C (I)
35        CONTINUE
5     CONTINUE
      WRITE (*, '(A)') 'New Ax, Ay: '
      WRITE (*, *) AB (1, 1), AB (1, 2)
      WRITE (*, '(A)') 'New Bx, By: '
      WRITE (*, *) AB (2, 1), AB (2, 2)
      END
```

程序中，THITA **为旋转角度**；TX **和** TY **为位移增量**；S **为放缩比率**；T **为变换矩阵**；C **为中间变量数组**；AB **为点** A **和** B **的坐标信息矩阵**，AB（1，1）**和** AB（1，2）**以及** AB（2，1）**和** AB（2，2）**开始时分别存放** A **及** B **两点的原坐标值**（x，y），**最后存放组合变换后的新坐标值**。

二、资源图件编绘子系统的基本子程序

地矿资源图件可以看作是一个特殊的符号系统。图面上以不规则自由曲线为主，由大量的中文、西文、注记、特殊的符号与花纹错综复杂地组合而成。只要设法建立任意不规则曲线和各种文字与花纹符号的数学模型，就能生成并组合成任意的地矿资源图件。图形的数学模型分析与程序的模块化设计相结合，形成了开发应用编图系统的基础。

1. 基础子程序的功能与形参

这里列出系统设置的 36 个模块化的基础功能子程序。其组成结构如图 6-7 所示。

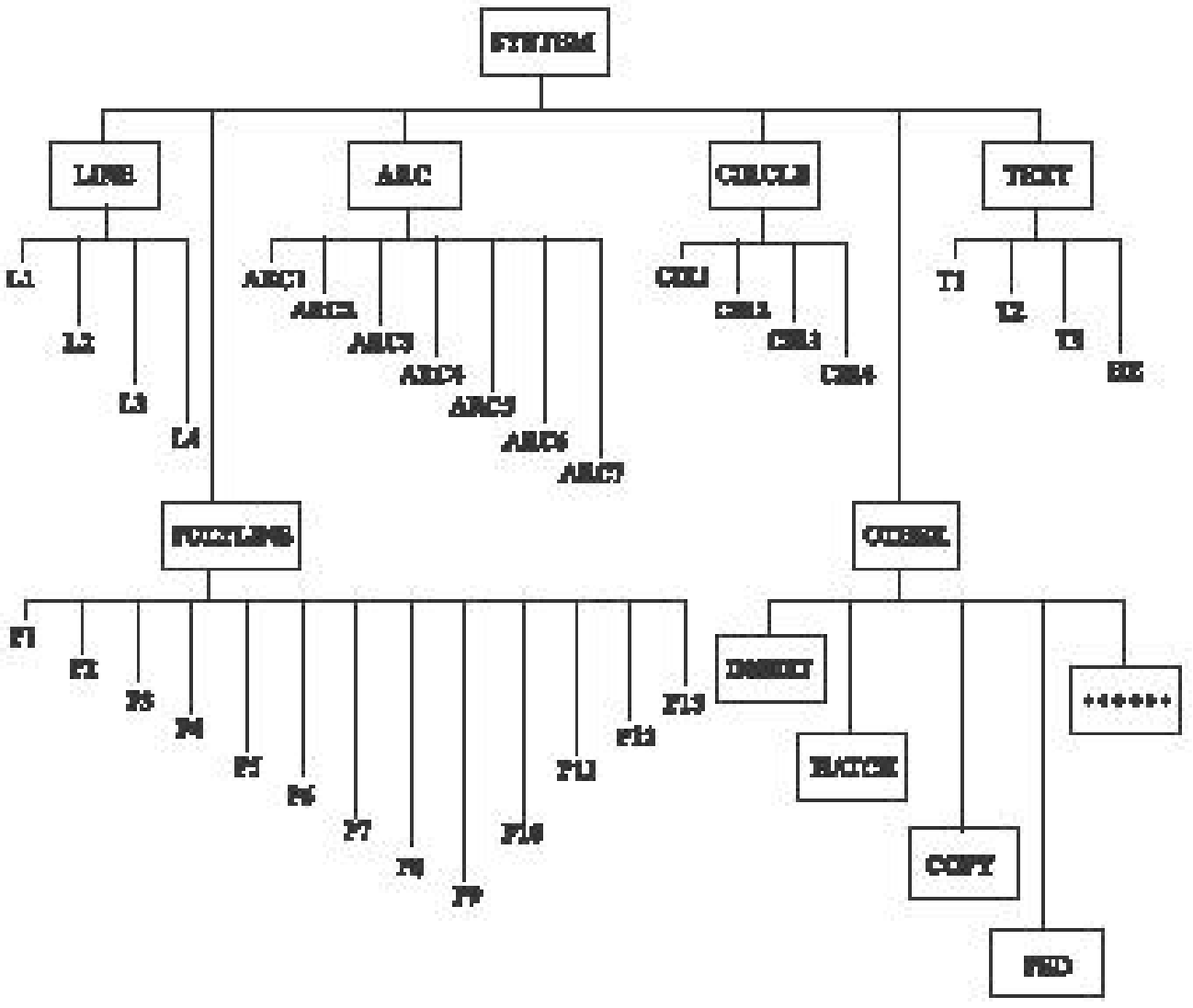


图 6-7　系统功能子程序

(1) L1 (X1, Y1, N1, X2, Y2, N2)

功能：过 (X1, Y1), (X2, Y2) 画直线。

说明：点的定义一般需要三个数据：(X, Y, N)。(X, Y) 表示该点的坐标值；N 为标示参数。N = 0 表示为绝对坐标，即相对于坐标原点的坐标；N = 1 表示相对坐标，即坐标由该点相对于前一点的位移确定。

(2) L2 (X, Y, N, D, A)

功能：以 (X, Y) 为起点，沿角度 A 方向作长度为 D 的直线。

说明：角度 A 以 X 轴正方向为零度，逆时针为正，以度为单位。

(3) L3 (X1, Y1, N1, X2, Y2, N2, W)

功能：过 (X1, Y1) 和 (X2, Y2) 画宽度为 W (mm) 的直线。

(4) L4 (X1, Y1, N1, X2, Y2, N2)

功能：**过**（X1，Y1）**和**（X2，Y2）**两点画点画线**。

(5) ARC1（X1，Y1，N1，X2，Y2，N2，X3，Y3，N3)

功能：**以**（X2，Y2）**为圆心**，（X1，Y1）**为起点**，（X3，Y3）**为终点逆时针画弧**。

(6) ARC2（X1，Y1，N1，X2，Y2，N2，D)

功能：**以**（X2，Y2）**为圆心**，（X1，Y1）**为起点**，D **为弦长逆时针画弧**。**当** D > 0 **时画小弧**，D < 0 **时画大弧**。

(7) ARC3（X1，Y1，N1，X2，Y2，N2，A)

功能：**以**（X2，Y2）**为圆心**，（X1，Y1）**为起点**，A **为圆心角逆时针画弧**。A **为负值时**，**顺时针画弧**。

(8) ARC4（X1，Y1，N1，X2，Y2，N2，X3，Y3，N3)

功能：**过**（X1，Y1)、(X2，Y2）**和**（X3，Y3）**三点画弧**。

(9) ARC5（X1，Y1，N1，X2，Y2，N2，D)

功能：**以**（X1，Y1）**为起点**，（X2，Y2）**为终点**，D **为直径逆时针画弧**。D > 0 **时画小弧**，D < 0 **时画大弧**。

(10) ARC6（X1，Y1，N1，X2，Y2，N2，R)

功能：**以**（X1，Y1）**为起点**，（X2，Y2）**为终点**，R **为半径逆时针画弧**。R > 0 **时画小弧**，R < 0 **时画大弧**。

(11) ARC7（X1，Y1，N1，X2，Y2，N2，A)

功能：**以**（X1，Y1）**为起点**，（X2，Y2）**为终点**，A **为圆心角逆时针画弧**。**当** A < 0 **时**，**顺时针画弧**。

(12) CIR1（X1，Y1，N1，X2，Y2，N2，X3，Y3，N3)

功能：**过**（X1，Y1)、(X2，Y2）**和**（X3，Y3）**三点画圆**。

(13) CIR2（X1，Y1，N1，X2，Y2，N2)

功能：**以**（X1，Y1）**和**（X2，Y2）**为直径两端点画圆**。

(14) CIR3（X，Y，N，D)

功能：**以**（X，Y）**为圆心**，D **为直径画圆**。

(15) CIR4（X，Y，N，R)

功能：**以**（X，Y）**为圆心**，R **为半径画圆**。

(16) T1（X，Y，N，H，RO，M，T)

功能：**以** H **为字高**，**以** RO **为旋转角**，**写字符串** T。

说明：M = 1，**指以**（X，Y）**为起点**，**即左对齐**；M = 2，**指以**（X，Y）**为中点**，**即中心对齐**；M = 3，**指以**（X，Y）**为终点**，**即右对齐**。

(17) T2（X1，Y1，N1，X2，Y2，N2，T)

功能：**在**（X1，Y1）**与**（X2，Y2）**之间写字符串** T。

(18) T3（X，Y，N，H，RO，M，DE，MM，NN)

功能：**以**（X，Y)、H、RO、M **确定的位置和大小**，**写由** DE、MM、NN **确定的数据**。

说明：DE **为欲填写的数据变量**，MM **为** DE **的小数点后欲取的位数**，NN **为** DE **的前缀特殊字符控制码**。

(19) HZ (X, Y, N, M, A, H, D, ST)

功能: 写汉字串。

说明: (X, Y) 为汉字的左下角起点; N = 0, (X, Y) 为点的坐标值; N = 1, (X, Y) 为对上一点的 X、Y 方向上的偏移量。

M 字数

A 汉字串

H 字高 (图上单位)

D 字间距 (图上单位)

ST 字体代码 (HZ0, HZ1, HZ2, HZ3 分别代表宋体、楷体、仿宋、黑体等)

(20) P1 (X, Y, N)

功能: 以 (X, Y) 开始画多段线。

说明: 多段线即一段由直线段或圆弧组成的折线, 也叫多义线。

(21) P2 (N)

功能: 结束多段线。

说明: N = 1 时, 将多线段封闭; N = 2 时, 不将多线段封闭;

(22) P3 (W)

功能: 设置多段线线宽 W (图上单位)。

说明: 线宽设置后, 以后的多段线均以此宽度作出, 直到遇到 P3 的另一次调用为止。

(23) P4 (X, Y, N)

功能: 以 (X, Y) 为末端点画直线。

(24) P5 (D)

功能: 按上一线段确定的方向, 画一长度为 D 的直线。

(25) P6 (D, A)

功能: 沿角度 A 方向画出一长度为 D 的直线。

(26) P7 (X, Y, N, A)

功能: 以 (X, Y) 为圆心, A 为圆心角逆时针画弧。

(27) P8 (A, R, D)

功能: 以 R 为半径, A 为圆心角画弧, D 为终点位置角。

说明: 终点位置角即终点与起点之连线与 X 轴正方向的夹角。

(28) P9 (A, X, Y, N)

功能: 以 (X, Y) 为终点, A 为圆心角逆时针画弧。

(29) P10 (D, X, Y, N)

功能: 以 (X, Y) 为圆心, D 为弦长, 逆时针画弧。

(30) P11 (R, X, Y, N)

功能: 以 (X, Y) 为终点, R 为半径逆时针画弧。

(31) P12 (X1, Y1, N1, X2, Y2, N2)

功能: 以 (X1, Y1) 和 (X2, Y2) 为第二、第三点画弧。

(32) P13 (X, Y, N)

功能: 过 (X, Y) 作弧与上一线段相切, 切点为上一线段之末端点。

(33) INSERT (X0, Y0, SX, SY, AN, NAME)

功能: 按要求在当前状态下插入一个已有图形。

说明: (X0, Y0)　　插入基点

SX　　X 方向插入比例

SY　　Y 方向插入比例

AN　　插入角度

NAME　　图形文件名

(34) HATCH (NAME, SC, AN, X1, Y1, X2, Y2)

功能: 以 (X1, Y1), (X2, Y2) 为上、下角点开窗口选目标进行花纹填充。

说明: NAME　　为花纹调用名

SC　　为填充比例因子

AN　　为填充花纹旋转角度

(35) COPY (X1, Y1, N1, X2, Y2, N2, X3, Y3, N3, X4, Y4, N4)

功能: 图形拷贝

说明: (X1, Y1) 和 (X2, Y2)　　窗口坐标

(X3, Y3)　　预拷贝的图形基础

(X4, Y4)　　图形移动后的基点

(36) PED (ST)

功能: 圆滑绘制的一条多义线。

说明: ST = 1, 指最后绘出的一条多义线; ST = 2, 指前一条多义线。

由于采用模块化和结构化方法，使得系统功能容易扩充和完善，程序设计效率和质量得到提高，系统开发的适应性和灵活性得到增强。

2. 曲线拟合的数学模型选择与线型制作子程序

曲线拟合是生成地质图件的基本内容之一，主要应用于等值线、线型制作及各种地质体边界线的追踪。在少数几个已知数据点的基础上进行曲线拟合，必然涉及数据点的外推计算，计算的精度取决于所使用的算法。

地矿资源勘查图件上的曲线具有多种类型，大多是多值函数，往往呈现大挠度、连续拐弯的图形特征。计算机绘制光滑曲线的基本思想，是把一条曲线看成是由按照某种数学模型排列的一系列密集点连接而成。如果已知曲线上有若干样点，就能通过一定的数学方法进行插值或逼近，得到这一系列密集点列的平面位置，并确保在这些点列上具有连续的一阶导数或连续的二阶导数，从而保证绘出的曲线是光滑的。由离散点生成光滑曲线并满足上述条件的数学方法很多，方法不同，生成的曲线也不相同。

(1) Bezier 曲线法

Bezier 曲线的数学基础是在第一个和最后一个端点之间进行插值的多项式函数。通常，将 Bezier 曲线段以参数方程表示如下:

$$P(t) = \sum_{i=0}^{n} P_i B_{i,n}(t) \qquad t \in [0,1]$$

这是一个 n 次多项式，具有 n +1 项，其中 P_i (i = 0, 1, …, n) 表示特征多边形 n +1 个顶点的位置向量。所谓特征多边形，即指依次用线段连接 P_i (i = 0, 1, …, n) 中

相邻两个向量的终点，这样组成的 n 边折线多边形称为 Bezier 多边形或特征多边形。$B_{i,n}(t)$ 是 Bernstein 多项式，称为基底函数，表示如下：

$$B_{i,n}(t) = \frac{n!}{i!(n-i)!}t^i(1-t)^{n-i} \qquad (i = 0,1,2,\cdots,n)$$

当 $i=0$ 且 $t=0$ 时，$t^i=1$，$0!=1$。

由上两式可得到一次、二次、三次 Bezier 曲线的数学表达式。

1）当 $n=1$ 时，则有

$$P(t) = \sum_{i=0}^{1} P_i B_{i,1}(t) = (1-t)P_0 + tP_1 \qquad (t \in [0,1])$$

表明一次 Bezier 曲线是连接起点 P_0 和终点 P_1 的直线段。

2）当 $n=2$ 时，则有

$$P(t) = \sum_{i=0}^{2} P_i B_{i,2}(t)$$

$$= (1-t)^2 P_0 + 2t(1-t)P_1 + t^2 P_2 \qquad (t \in [0,1])$$

二次 Bezier 曲线如图 6-8 所示。

3）当 $n=3$ 时，则有

$$P(t) = \sum_{i=0}^{3} P_i B_{i,3}(t)$$

$$= (1-t)^3 P_0 + 3t(1-t)^2 P_1 + 3t^2(1-t)P_2 + t^3 P_3 \qquad (t \in [0,1])$$

三次 Bezier 曲线如图 6-9 所示。

在实际构造各种资源勘查图形时，高于三次的 Bezier 曲线本身计算和数据处理都比较复杂，一般不常用。

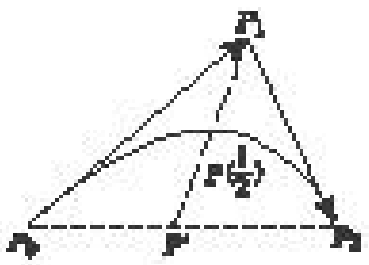


图 6-8　二次 Bezier 曲线

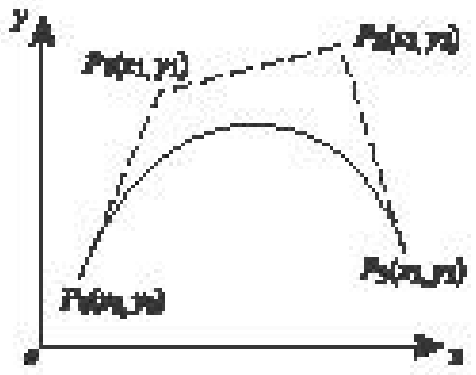


图 6-9　三次 Bezier 曲线

（2）B 样条函数法

Bezier 曲线的实质是：在给定了一串点列 P_i（$i=0$，1，…，n）后，选择函数族 $\{B_{i,n}(t)\}$ 作为基函数，且由此作出向量值形式的 Bernstein 逼近。后来，人们拓广了 Bezier 曲线，将基函数 $\{B_{i,n}(t)\}$ 换成 n 次 B 样条逼近，构造了等距节点 B 样条曲线。

若给定 $m+n+1$ 个空间向量 P_K（$K=0$，1，2，…，$m+n$）后，则称 n 次参数曲线

$$P_{i,n}(t) = \sum_{i=0}^{n} P_{i+1} F_{l,n}(t) \qquad (t \in [0,1])$$

为 n 次 B 样条的第 i 段曲线（i = 0，1，…，m），它的全体称为 n 次 B 样条曲线。相应地，如果依次用线段来连接 P_{i+l}（l = 0，1，…，n）中相邻两个向量的终点，那么，所形成的多边形称为样条的第 i 段的 B 特征多边形。

当 n = 2 时，则二次 B 样条曲线可以写成下列矩阵形式：

$$P(t) = \sum_{i=0}^{2} P_i F_{i,2}(t) = [t^2 \quad t \quad 1] \cdot \frac{1}{2} \cdot \begin{bmatrix} 1 & -2 & 1 \\ -2 & 2 & 0 \\ 1 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} P_0 \\ P_1 \\ P_2 \end{bmatrix} \qquad (t \in [0,1])$$

二次 B 样条曲线如图 6-10 所示。

当 n = 3 时，三次 B 样条曲线也可写成下面的矩阵形式：

$$P(t) = [t^3 \quad t^2 \quad t \quad 1] \cdot \frac{1}{6} \cdot \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 0 & 3 & 0 \\ 1 & 4 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} P_0 \\ P_1 \\ P_2 \\ P_3 \end{bmatrix} \qquad (t \in [0,1])$$

三次 B 样条曲线如图 6-11 所示。高于三次的 B 样条曲线情形比较复杂，一般不宜采用。与 Bezier 曲线相比，B 样条曲线具有更好的逼近性。

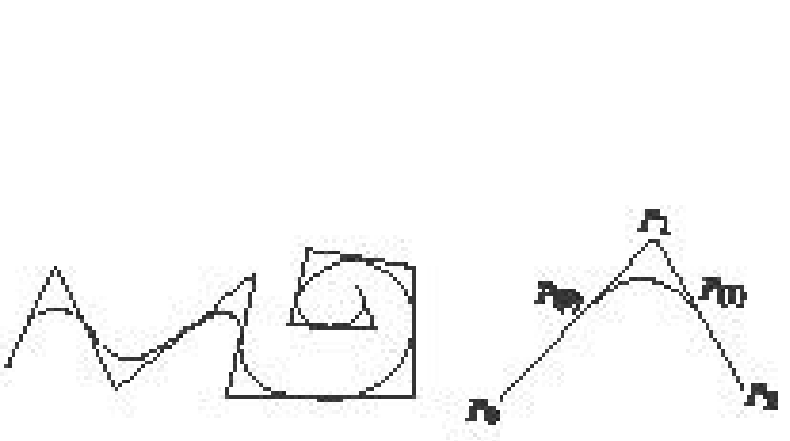


图 6-10　二次 B 样条曲线

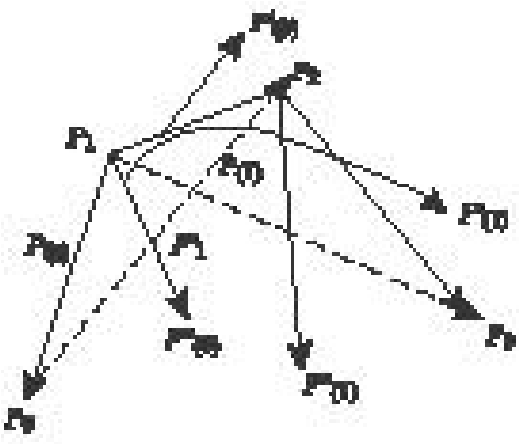


图 6-11　三次 B 样条曲线

(3) 三次参数样条函数法

三次样条函数是张力样条函数的特例。张力样条函数的显著特征是具有一个张力系数 σ。当 σ→0 时，张力样条函数就等同于三次样条函数，当 σ→∞时，则它将是化成分段线性函数，即从结点到结点是折线联结。因此，可以选择合适的参数 σ，使得点与点之间的曲线尽量缩短，好像在整条曲线的两端用一种作用力拉到合适的程度，既能消除可能出现的多余拐点，又保持了曲线的光滑性。图 6-12 给出了不同张力系数下产生的不同的图形。

在图形效果上，张力样条函数可从理论上取得两点间最短曲线，而且节点间的距离不受限制，即无论整条曲线上相邻数据点间距差别多大，其曲线外形是一致的。如果数据点较多，可以分若干段处理，只要相邻端点条件一致，最终的整条曲线还是光滑的。

本方法的缺点是在某些情形下呈波浪形，有抖动现象，严重时致使曲线失真。

实用三次样条插值法（王来生等，1992）是三次样条函数的一种扩展形式，它提供

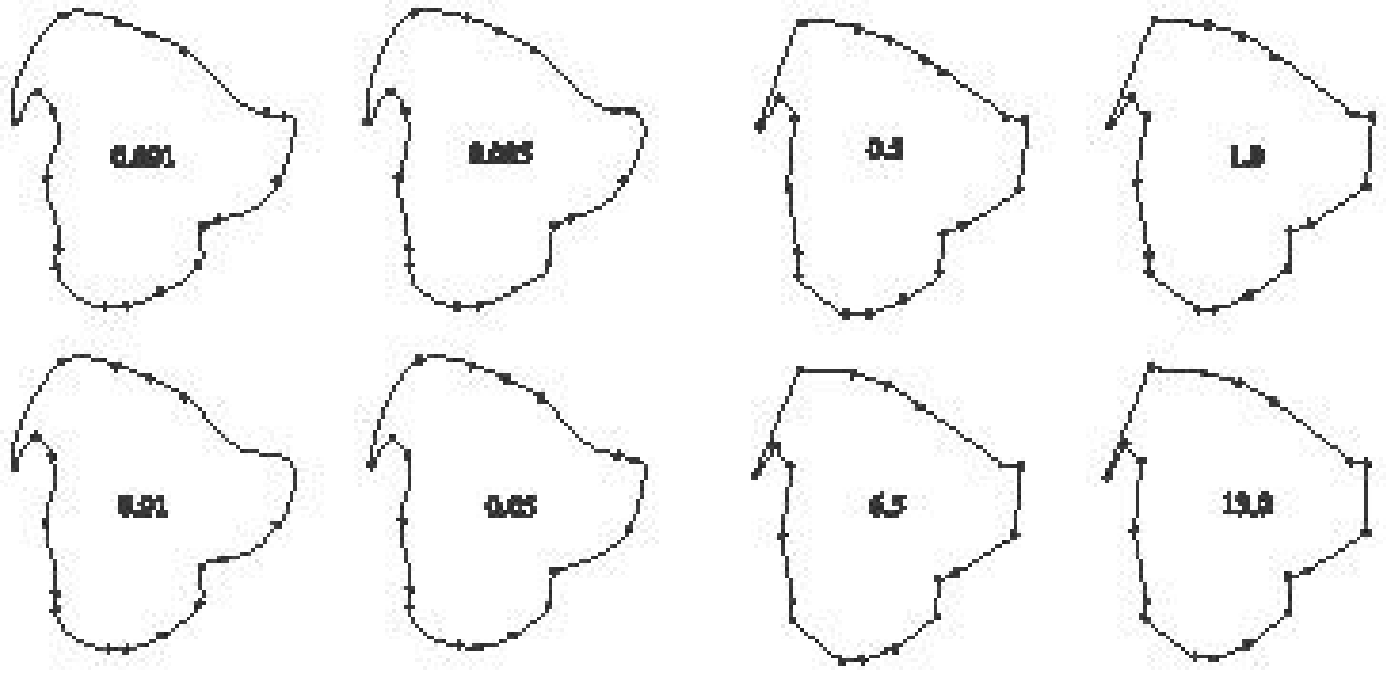


图 6-12　张力样条法中选择不同的张力系数对图形的影响

了一个调节曲线形状的参数，可以适用于不同的场合。其方法简述如下:

已知平面上 n 个点 P_1 (x_1， y_1)， P_2 (x_2， y_2)， …， P_n (x_n， y_n) 假设函数 f 是插值于 P_1， P_2， …， P_n 之间的三次样条插值函数，并设 V_i = (V_{ix}， V_{iy})， (i = 1， …， n) 是使 f 达到二次导数 (C^2) 连续的 P_i (i = 1， …， n) 处的切矢量。我们知道，给定两点 P_0、P_1 以及曲线在这两点的切矢量 V_0、V_1，可唯一地确定一段三次参数曲线。设平面上三次参数曲线段为:

$$\begin{cases} x(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 \\ y(t) = b_0 + b_1 t + b_2 t^2 + b_3 t^3 \end{cases} \qquad (0 \le t \le 1)$$

取曲线段上两端点 P_i (x_i， y_i)， P_{i+1} (x_{i+1}， y_{i+1}) 由假设知，过这两点的切矢量为 V_i，V_{i+1}。在此我们引入实参数 λ，对切矢量进行控制，则该曲线段的端点条件为:

$$\begin{cases} x(0) = x_i, y(0) = y_i \\ x(1) = x_{i+1}, y(1) = y_{i+1} \\ x'(0) = \lambda V_{ix}, y'(0) = \lambda V_{iy} \\ x'(1) = \lambda V_{(i+1)x}, y'(1) = \lambda V_{(i+1)y} \end{cases}$$

由上面两式可求出 a_0， a_1， a_2， a_3， b_0， b_1， b_2， b_3，从而得出实用三次样条曲线段方程:

$$\begin{cases} x(t) = x_i + \lambda V_{ix} t + [3(x_{i+1} - x_i) - 2\lambda V_{ix} - \lambda V_{(i+1)x}] t^2 \\ \qquad + [\lambda V_{ix} + \lambda V_{(i+1)x} - 2(x_{i+1} - x_i)] t^3 \\ y(t) = y_i + \lambda V_{iy} t + [3(y_{i+1} - y_i) - 2\lambda V_{iy} - \lambda V_{(i+1)y}] t^2 \\ \qquad + [\lambda V_{iy} + \lambda V_{(i+1)y} - 2(y_{i+1} - y_i)] t^3 \end{cases}$$

对于点列 P_i (i = 1， …， n)，上式可写成一般形式:

$$x_i(t) = x_i + \lambda V_{ix} t + [3(x_{i+1} - x_i) - 2\lambda V_{ix} - \lambda V_{(i+1)x}] t^2$$

$$
\left|\begin{aligned}
&{}_{ix} + \lambda V_{(i+1)x} - 2(x_{i+1} - x_i)]t^3 \\
y_i(t) &= y_i + \lambda V_{iy} t + [3(y_{i+1} - y_i) - 2\lambda V_{iy} - \lambda V_{(i+1)y}]t^2 \\
&\quad + [\lambda V_{iy} + \lambda V_{(i+1)y} - 2(y_{i+1} - y_i)]t^3
\end{aligned}\right.
$$

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

参数 λ的出现是实用三次样条插值函数的显著特征。该特征的图形效果表现在：当 $\lambda=1$ 时，即为满足 C^2 连续的三次样条函数；当 $\lambda=0$ 时，为线性插值函数（图 6-13）。从图上可以明显看出，实用三次样条曲线对应于不同的 λ值具有内收或外张的趋势，它的应用范围比较广。

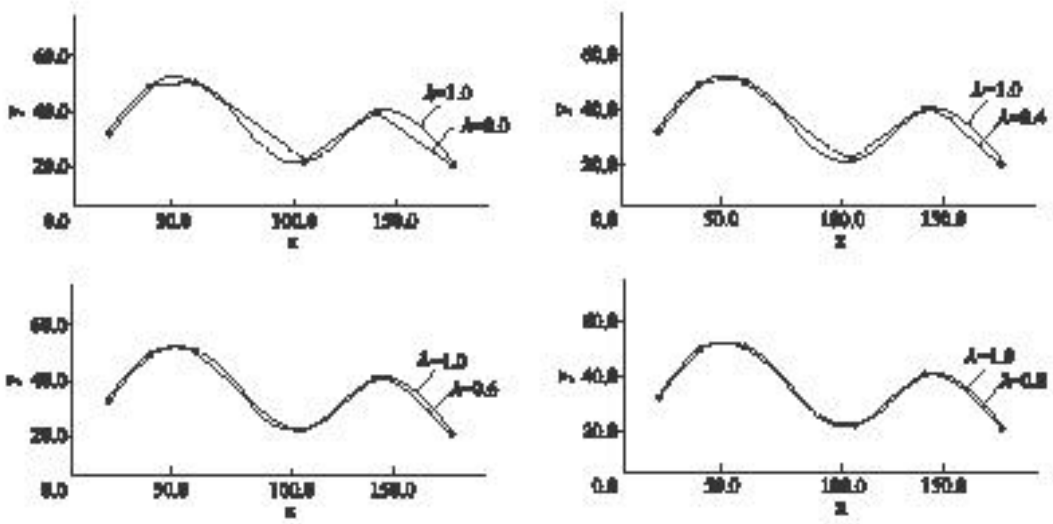


图 6-13　参数 λ取不同值时实用三次样条曲线的变化情况

(4) 几种拟合方法的比较

比较上述几种曲线拟合方法，可看到如下几点差别：

1) 对特征多边形的逼近性　二次 B 样条曲线优于三次 B 样条曲线，还可能优于实用三次样条曲线；实用三次样条曲线和三次 B 样条曲线优于三次 Bezier 曲线。

2) 对所逼近的相邻曲线段之间的连续性　二次 B 样条曲线只达到一阶导数连续；三次 B 样条曲线则达到二阶导数连续；实用三次样条曲线对应于不同的 λ值，可分别达到一阶和二阶导数连续。

3) 对特征多边形型值点的逼近性　实用三次样条曲线优于二次 B 样条曲线；二次 B 样条曲线优于三次 B 样条曲线；三次 B 样条曲线优于三次 Bezier 曲线。

图 6-14 表示这几种曲线对同一个特征多边形及其型值点 P_0，P_1，P_2，P_3 的逼近性。

根据这些差别可以认为，Bezier 曲线和 B 样条曲线适用于不要求通过型值点的曲线拟合，例如某些地物图示符号绘制时曲线的输出；实用三次样条曲线适用于要求通过型值点情形的曲线拟合，如等高线等的输出。张力样条函数产生的曲线有时会产生振荡现象(呈波浪形)，而实用三次样条函数在 $0 \le \lambda \le 1$ 时具有良好的性质。因此，选择实用三次样条函数，利用曲线拟合方式可以构造出多种美观适用的线型。

在确定了数学模型后，便可以着手线型制作。依参数化设计方法，必须进一步规定线型的适应范围和相应参数项。参数项包括：①实线段长度和虚线段长度——实线段和虚线段均可由多个重复的基本单元组成，可用于生成几种常用的点划线型，若虚线段长度为 0，则该线整条为实线；②线的宽度——规定线的宽度系数；③插入符号——可能是一种

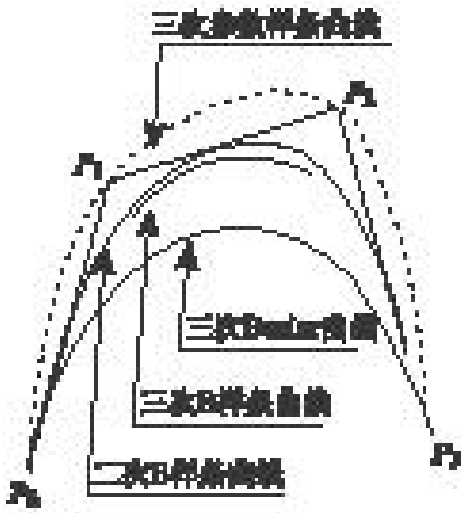


图 6-14　各种曲线对特征多边形的逼近性

图案，也可以是文字，应规定插入位置和方向；④线型缩放比例——默认值为 1。

以上①、②、③项基本确定了线的外形，其中插入符号可能较复杂，应作相应图案规定。第④项可在不同比例尺下调节使用。

在程序设计中，可建立线型编码（线型名）和线型库（图 6-15），以方便调用。图形实际显示和输出时，需通过解方程逐点精确计算并绘出线型。通过以上方式可以设计出大量实用的线型，满足日常工作的需要。

3. 花纹、图例制作子程序

任何资源图件都是由符号、线条、花纹与颜色



图 6-15　实用线型举例

组成的，而图例是各种符号、线条、花纹与颜色的详细说明。这里着重介绍岩石花纹及其图例的制作。

(1) 花纹、图例的制作原则

资源图件中岩石花纹占有极为重要的地位，地矿部门所颁布的岩石花纹直观地反映了地质体的成分、构造、历史演变等。绚丽的花纹还有装饰、美化图面的作用。迄今为止，我国复杂的岩石花纹系统都是通过手工方式设计和描绘出来的，要将其往计算机上直接移植，有一定的难度，需要建造专门的地矿花纹图例库。为满足地质专业工作和计算机处理两个方面的要求，在建造地矿花纹图例库时，应当考虑以下几点:

通用性。应能适用于各类地质岩性花纹，不能局限于某一类岩石；能适用各类地质图件，而不能仅适用于某一种图件。

独立性。岩性花纹图例库应独立于程序之外，各种程序都能方便、快速地调用。

开放性。应完全对用户开放，允许用户任意地增删、修改、编辑，甚至重建。

单一性。岩石花纹和岩石名称具有明确的对应关系，每一个岩石花纹都有一个岩石名称，每一个岩石名称都要有对应的岩石花纹。为了使用方便，同一岩性花纹可以有两种以上的岩石名称或代号，但一个岩石名称或代号决不能对应一种以上的岩性花纹。

可编辑性。每一个岩石花纹都允许阅读、查看、修改、增删和编辑。

(2) 花纹、图例的制作方法

1) 基本图元法。对于规则边界的岩石花纹绘制，采用基本图元法能收到很好效果，

例如钻孔柱状图等。其基本思想是将岩石花纹划分成基本单元块，然后由基本单元按一定规律重复排列组合形成各类岩石花纹（吴冲龙等，1987），如图 6-16 所示。



图 6-16　用基本图元法形成柱状图

基本图元法的实现有两种途径，一种为计算方式，一种为组合方式。前者是根据边界的宽度进行函数计算，直接确定组成直线段或圆弧的花纹的坐标；后者是在充分分析各岩石花纹组成的情况下，建立各单一花纹的图形块，再通过图形块的排列组合来构成各种岩性花纹。计算方式灵活多变、适应性强，组合方式则快速简便、容易实现。

2）辅助设计法。在不规则边界区域内填充花纹是一个常见而复杂的问题。复性杂不在于花纹填充的算法上，而在于花纹图案的设计。为了提高效率，可编制专门的机助设计子程序来实现。下面以阴影图案为例，对图案的计算机辅助设计加以说明。

阴影图案（岩石花纹或晕线）是由一些基本图形单元纵横有规律地组合而成的，允许图案整体以任意角度旋转。在阴影图案中，每一根图案线的生成语句格式说明如下：

$$(\text{Angle},\ X_0,\ Y_0,\ X_d,\ Y_d,\ XL_1,\ XL_2)$$

其中：Angle，指画线的角度，在 - 360 °～360 °之间，一般取较特殊角度如 30 °，45 °，90 °等；X_0，画线起点的横坐标；Y_0，画线起点的纵坐标；X_0，该线延伸方向上的偏移；Y_d，线族内线与线之间的垂直距离；XL_1，XL_2 两参数说明该线线型，正值为实线段，负值为虚线段，零值表示一个点。例如（45，0，0，0，0.125）表示线的偏转角为 45 °，图案起点为（0，0），阴影线自身方向的偏移量为 0，垂直向上阴影线之间的距离为 0.125 单位；XL_1，XL_2 参数缺失，表明该线为一条实线。其图形如图 6-17 所示。

根据以上基本原理，可以编制出阴影线图案的计算机辅助设计程序，用于编辑花纹（点线绘制、删除等）、显示花纹（不同类型、不同比例尺）。

4. 图件整饰子程序

地矿资源图件涉及的专业范围较广泛，图件整饰子程序的编制应参照相应专业的规范要求，而程序设计上要力求灵活。

（1）图框整饰

大部分地矿勘查图件，特别是各种平面图和剖面图都要求有内图框或外图框（图 6-18）。一般规定，在 1∶25 000 至 1∶100 000 的图内只绘方里网，不绘经纬线，但应当在四图角注出经纬度数字，并且在内外图廓的粗细线间按经纬差各 1 绘出分度带。

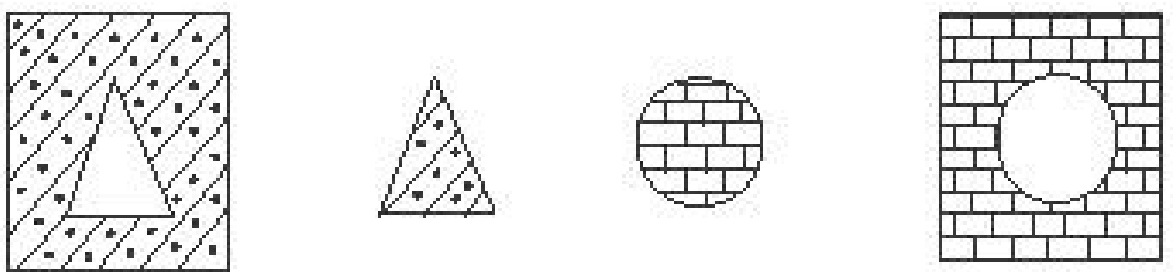


图 6-17　阴影线充填

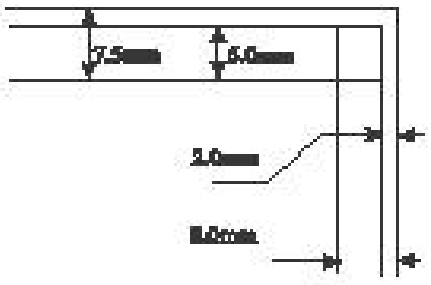


图 6-18　内外图框示意图

在不同用途的图件上图框常有变化，例如有时将图框简化为单线图框，而有时在一个图框中包含两个以上的小图，子程序调用应当同时给出大小和型号。

(2) 坐标网整饰

在平面图上，方里网通常按图幅和比例尺要求绘制；在剖面图上，则要求将平面的方里网进行投影，并同时绘出高差，高程标注于图框两侧。由于每个图幅的坐标系事先给定，只要绘图范围明确，其坐标网就可利用通用子程序绘制出来。

(3) 分度带整饰

按规定，比例尺小于 1∶ 25 000 万的图件用 6 °分带，等于或大于 1∶ 10 000 的图件用 3 °分带。这些分度带都可按实际图形自动绘制。

(4) 图幅和分幅整饰

关于国际分幅，有一系列具体规定与原则，但勘探报告中所附的各种平面图件，除区域地质图和地形地质图需考虑国际分幅外，其余图件均以满足图幅能包括全部内容，而又出现最少空白为原则，不必考虑国际分幅。图幅的大小，应同时考虑绘图仪的大小和便于折叠保存，尽量为标准 16 开纸的整数倍。如图幅过大需分为数幅时，在每幅图廓外侧的右上方绘出接图表，表明各图的相对位置，并且在该图幅所在位置处打上阴影。

(5) 图名整饰

资源勘查图件的名称由省、县 (市)、矿区名三部分及图名排列组成。图名全部用汉字，写在图的上方并居中。系统中配有标宋、楷体、黑体、仿宋、隶书和魏碑等多种字体。

(6) 比例尺整饰

1∶ 50 000 及更小的各类平面图，应兼有数字比例尺和直线比例尺。1∶ 10 000 及更大比例尺的平面图只画数字比例尺。标准分幅图件的数字比例尺和直线比例尺绘于图名下方并居中，任意分幅比例尺绘于图的上方图名之下。凡未按正上方表示正北方向绘制的平面图，均应绘出表明真北或磁北方向的指北针注记。

(7) 图签整饰

每一幅勘查图件都要求在右下角附有图签。图签的大小，通常根据图幅大小而定，一般以 9 cm× 5. 5 cm ～10 cm× 6 cm 左右为宜。图签位置应有利于在图纸折叠时，露出图签的全部或大部。图签的内容如图 6-19 所示。

(8) 图例整饰

×× 省×× 局×× 队			
××××× 图			
拟编		顺序号	
审核		图号	
清绘		比例尺	
技术负责		日期	
队长		资料来源	

图 6-19　图签范例

图件中所绘各种图形符号、文字符号、花纹及颜色都必须全部给出图例，地形底图的某些惯用符号可以例外。图例排列的顺序为：地层系统 (自新而老)、侵入岩、岩相、构造、矿产、探矿工程等。图例一般绘在右图廓外，但图面空白较多时也可绘于图框内。成套的图件，如同一地区的多幅剖面图，可单独编一张统一图例。

各种岩性的图例可在程序处理过程中自动处理排序。各类图例可事先做成标准子模

块，在程序处理或人机交互制作中调用。

第三节　专题图件机助编绘的程序设计

专题图件编绘程序设计，是指各种专用资源勘查图件的综合生成程序设计与开发。专题图件的计算机编绘大致包括如下几个步骤：①建立图件模型（包括专业模型和几何模型）。依照行业规范和专家经验，确定图件格式，详细指明图件的各部分组成和要求，以及相互关系。②建立图形计算模型。在地质图件中，数据按其来源和性质可分为两类：一类为详实数据，可依此类数据直接进行空间点的定位输出，如地理坐标，高程等；另一类为逻辑数据或推测数据，需要在处理过程中进行逻辑判断与计算。因此在图形数据可视化的过程中，应当针对不同数据特点建立相应的计算模型。③给定数据格式要求。依照所建立的图件模型和图形计算模型，制定合理的数据格式，再交由点源数据库检索生成。④生成图形交换数据或文件。应用绘图程序按计算模型处理数据，并且按照标准格式生成图形交换数据或文件。⑤生成正式图形文件。在基础绘图系统中生成和显示图形，然后通过人机交互手段对图形进行修改加工，对整幅图进行整饰，生成正式图形文件。⑥输出图形进行硬拷贝。

资源勘查图件种类繁多，下面仅以三种最具代表性的常用图件为例加以说明。

一、钻孔柱状图编绘程序

钻探是地质勘探工程最主要的手段。因此，作为钻探成果的钻孔柱状图是地质勘探报告中最基本的、数量最多的图件。

1. 图件格式规定

钻孔柱状图是一种表格式图件，从计算机程序设计的角度可将其视为表格。

钻孔柱状图可划分为如下几个基本单元：表名（图名）、表眉（图眉）、表头（图头）、岩层表、物性参数表、样品表、责任表及其他附表，如图 6-20 所示。有些钻孔综合柱状图的图式比较复杂，例如岩层表和物性参数表部分都含有多项栏目，但总体结构是一致的。

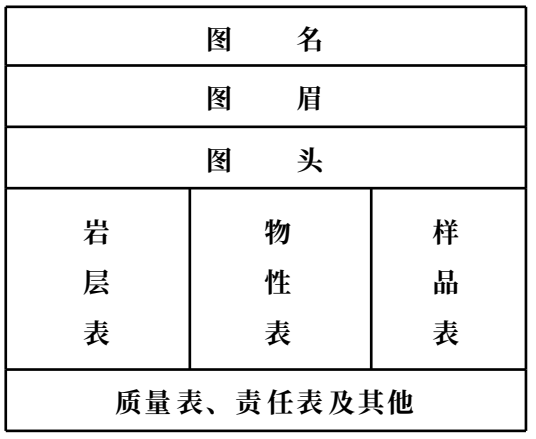


图 6-20　钻孔柱状图结构分割图

2. 机助编绘方法概要

图头文字部分按规范规定格式，将其框架作为一个整体进行单独处理，预先形成构件，在生成图形时作为图块插入，然后填入可变的数据部分。这样可减少许多重复工作并提高成图速度。

钻孔柱状图的长度由钻孔深度和比例尺决定。岩芯柱状、分层界线、测井曲线和深度标尺，只需按比例尺要求换算长度并定位即可生成。在柱状图图面上，每一岩层的厚度为：

$$层厚 = （本层底板深度 - 上层底板深度）\times 比例尺$$

当分层厚度太薄或其他原因导致岩性描述栏容纳不下所有文字时，分层底线能自动下移至合适位置；而当底线下移仍不能解决问题时，分层顶线将自动上移，或者分层顶线和底线同时自动上移和下移。线条下移或上移距离用以下公式计算：

$$\begin{cases} S_2 = S_1 + (H - W) & \text{当 } H \geq W + S_1 \\ S_2 = 0 & \text{当 } H \leq W + S_1 \\ S_1 = S_2 \end{cases}$$

式中：S_1，S_2 分别为上层底线和本层底线下移距离；H 为本层描述栏所需的顶底线之间最小距离；W 为本层的厚度。描述栏顶底线间最小距离 H 由描述内容的长度 L、每行最多描述字符数 N 和行距 D 决定，即

$$H = D \cdot INT(L/N) + 1$$

若考虑到屏幕左右长、上下短的特点，可将图形置为左右延伸状态，做逆时针旋转 90°。应用旋转公式并令：

$$\begin{cases} x' = -y \\ y' = x \end{cases}$$

其中，(x，y) 为原坐标，(x′，y′) 为旋转后的坐标（公式可参照本章前述的有关公式)。

3. 数据准备

数据是程序处理的核心。数据格式是依据图件内容和专家分析结果制定的。在本系统中，数据文件可直接从点源数据库中按钻孔检索生成。一个钻孔对应一个数据文件。

下面是南水北调中线工程地质勘察信息系统中的一个数据文件格式实例：

(1) 勘测单位全称；
(2) 工程名称；
(3) 钻孔编号；
(4) 钻孔位置；
(5) 地面高程，单位：米 (m)；
(6) 孔口坐标，x 值，y 值 (地理坐标)；
(7) 钻探目的；
(8) 钻孔深度，单位：米 (m)；
(9) 覆盖层厚度，单位：米 (m)；
(10) 地下水位，单位：米 (m)；
(11) 钻孔角度，单位：度 (°)；
(12) 岩芯平均获得率；
(13) 风化带厚度，单位：米 (m)；
(14) 开孔日期；
(15) 终孔日期；
(16) 钻探单位；
(17) 机长姓名；
(18) 制图者姓名；
(19) 校核者姓名；
(20) 记录者姓名；
(21) 图号；
(22) 地层时代：a. 起始符号 $1

b. 划分时代的个数 m

c. 下面顺序排列 m 个数据，每组如下：底深度，时代名（由浅至深)

(23) 钻孔结构：a. 起始符号 $2

b. 钻孔变径数 m

c. 由浅至深排列 m 个数据组：底深，钻孔直径

(24) 柱状图钻孔数据：a. 起始符号 $3

b. 钻孔分层数 m

c. 顺序由浅至深排列 m 个数据组：底深度，岩性，层序号

(25) 岩性描述：a. 起始符号 $4

b. 描述内容总行数 m

c. 顺序排列 m 行文字说明

(26) 单位吸水量：a. 起始符号 $5

b. 给出数据点个数 m

c. 顺序给出各点数据：深度，数据（字符型)

(27) 风化程度：a. 起始符号 $6

b. 给出数据点个数 m

c. 顺序给出各点数据：深度，数据（字符型)

(28) 岩芯获得率：a. 起始符号 $7

b. 给出数据个数 m

c. 由浅至深顺序给出 m 个数据组：起始深，底深，百分数值（%)

(29) RQD 值：a. 起始符号 $8

b. 给出数据个数 m

c. 由浅至深顺序给出数据组：起始深，底深，百分数值（%)

(30) 声波波速：a. 起始符号 $9

b. 给出数据个数 m

c. 由浅至深顺序给出 m 个数据组：深度，波速值（km/s)

(31) 点荷载：a. 起始符号 $10

b. 给出数据个数 m

c. 依次给出各点数据：深度，点荷载值

(32) 采样位置及编号：a. 起始符号 $11

b. 给出采样个数 m

c. 依次给出各采样

点的数据：深度，编号

(33) 钻孔小结：a. 起始符号 $12

b. 给出钻孔小结总行数 m

c. 顺序给出 m 行文字小结

内容

(34) 全文结束符：$ $ $

4. 程序功能与程序设计框图

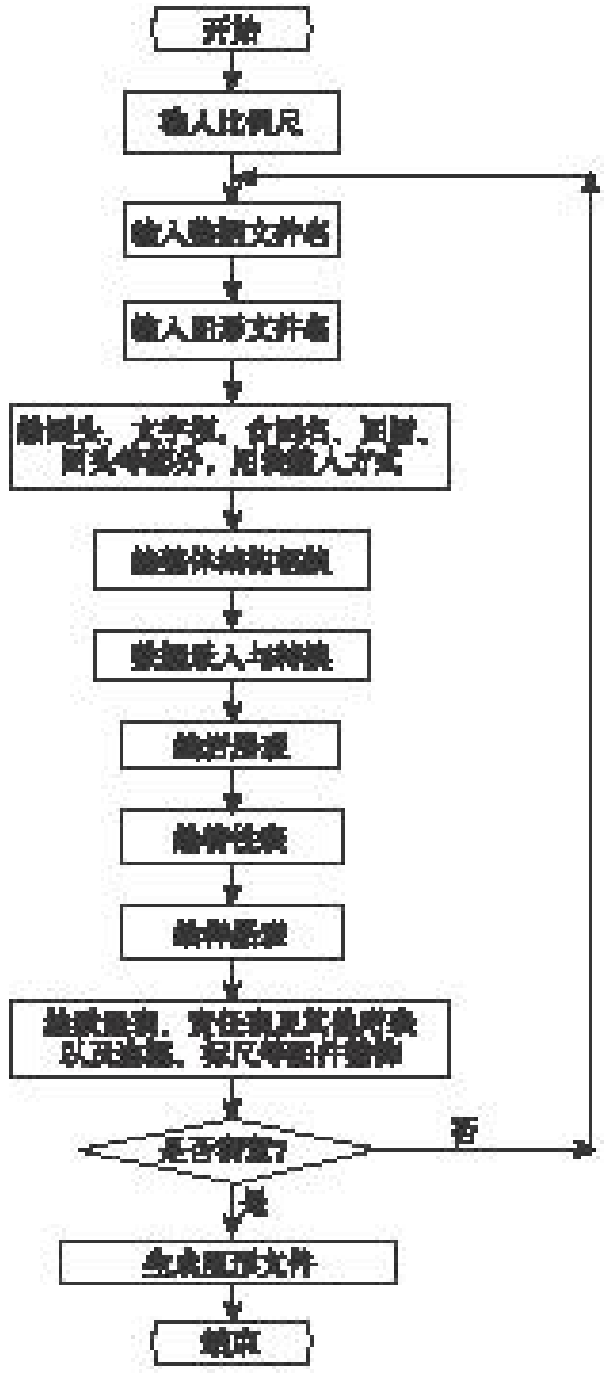


图 6-21　钻孔柱状图程序处理框图

本教程所附的实习程序所带的原始数据可由数据库检索生成，也可由用户自行建立。程序可以按照格式要求，自动完成柱状图的绝大部分绘制工作，能自动充填花纹、标注文字说明，能按比例尺要求自动形成各岩层界限及边框标尺等。在一般情况下，使用喷墨绘图仪等绘图设备就能满足用户的精度要求。

程序框图如图 6-21 所示。

二、勘探线剖面图编绘程序

勘探线剖面图是勘探报告中的另一类基本图件。它可直观地显示出勘探区的某一方向上（勘探线）地层、构造和矿体的变化情况，由多条勘探线剖面图组成的网络，还可以揭示各种地质体在三度空间中的分布特征。因此，勘探线剖面图具有指导地质勘探、成矿预测、矿山设计与生产决策的作用。

1. 图件内容分析

勘探线剖面图的主要内容包括:

①图名，比例尺，图廓，坐标网，责任表，图例;

②地形线，方位角，勘探线上的工程位置及编号，地物标志等;

③钻孔岩性柱状图及花纹填充;

④地层、岩体、断层、褶皱、破碎带等地质体的界线与产状，以及地质符号;

⑤不同地质专业有不同的特殊要求，例如工程地质勘探线剖面图的下部为标注栏，需要标明钻孔间距、桩号、钻孔编号、孔口高程、地质结构分段等参数;

⑥采样位置标注、特殊曲线表达以及其他附表。

2. 数据准备

地质剖面图内容丰富，钻孔地层数据在其中仍占主要地位。编绘勘探剖面图涉及多个钻孔的岩层数据，所需要的数据文件可通过数据库的相应检索专题来生成。

勘探线剖面图的专题检索结果是有序的，因为钻孔的地层数据由浅至深排列，地层时代的给出有先后，而且钻孔在勘探线上的定位也是有方向性和次序的。目前，系统在数据库专题检索中已做到钻孔在勘探线上的自动搜索定位，即给出勘探线方位以及起始点坐标、终点坐标后，计算机可以对各钻孔（井）进行自动投影定位。

下面是南水北调工程地质勘察信息系统中勘探线剖面图程序所用的专题数据示例，包含有 5 个数据文件，涵盖了剖面绘制的各种主要信息。

(1) 钻孔数据文件（.DAT)

按钻孔分别建立，便于查询及修改。格式如下（注意: ①岩层由浅至深顺序排列; ②后缀名定为“.DAT”) :

钻孔孔号，桩号，孔口高程，钻孔深度，覆盖层厚度，地下水位，基岩顶面高程，弱风化带底板高程。

岩层层数 N

岩层深度 1，标准代码 1，岩性名 1

岩层深度 2，标准代码 2，岩性名 2

… … …

岩层深度 n，标准代码 n，岩性名 n

(2) 地形线数据文件（.DX)

数据组组数 m
桩号 1，高程值 1
桩号 2，高程值 2
…，…
桩号 m，高程值 m
$$$，$$$（多条地形线可放于同一文件顺序排列）
(3) 地物说明文件（.DW)
数据组组数 m
桩号 1，高程值 1，文字说明 1
桩号 2，高程值 2，文字说明 2
… … …
桩号 m，高程值 m，文字说明 m
(4) 方位角说明文件（.FW)
数据组组数 m
桩号 1，高程值 1，方位角 1
桩号 2，高程值 2，方位角 2
… … …
桩号 m，高程值 m，方位角 m
(5) 钻孔名汇总文件（.NAM)
钻孔个数 n
钻孔数据文件名 1（不加后缀），钻孔类型 1，桩号 1
钻孔数据文件名 2（不加后缀），钻孔类型 2，桩号 2
… … …
钻孔数据文件名 m（不加后缀），钻孔类型 m，桩号 m

注意：可以使用文件名的后缀来区分各类数据，以方便程序处理，但这种处理方法缺乏灵活性，所以程序中实现直接读库和直接检索信息是较好的途径。

3. 程序设计与处理

勘探剖面图机助编绘原理与常规手工编绘类似，但借鉴了人工智能方式，工作效率高，绘出的图件精度也高。

(1) 基本规则

1) 图框按照地质专业规范绘制;

2) 坐标网按测绘要求绘制，并标注有关参数;

3) 曲线的拟合可采用实用三次样条函数插值，兼顾美观与精度;

4) 岩性符号的充填应事先建库，然后由计算机自动调用快速充填。

(2) 处理功能

本教程所附的实习程序具有编制地矿资源图件的一些主要功能。

1) 基本的自动成图功能。能按所要求的比例尺，自动生成各钻孔外形和岩层界线、填充岩性花纹，自动标注孔口标高、孔号及方位角等参数，自动生成地形线、等深线及边框标尺，而且能绘制多种线型。

2）孔斜校正和钻孔地质点投影计算。该模块能够面对孔深较大、孔斜也较大的钻孔，从而扩大了程序的适用范围。

3）图形编辑。考虑到剖面钻孔间的岩层对比需要地质工作者干预，系统选用了 CAD 软件系统的图形编辑界面。用户可采取人机对话方式进行对比和连线操作，还可对计算机自动生成的图形进行人机交互式补充和修改。

4）线型、符号和属性的分层存放。为了方便编辑操作和图形处理，系统事先安排一些图层（Layer）来存放各种线型、断层符号等所需图元和各种地质体的属性，在需要时可灵活地调用。

5）图形的链式生成。程序将图形生成过程表达为一个链式过程，其优越性在于：若由于外界因素造成某一钻孔的生成中断，不必从头开始再生成一遍，只需调用该孔相应的生成段即可。

(3) 程序框图

编绘系统的勘探线剖面图编绘流程如图6-22所示。

三、等值线图编绘程序

等值线图是在二维平面域内表现三维空间物体特征的图件。无论是地形地貌，还是地球化学场和地球物理场，只要有数值量度如高程值、分析值和场强参数值，均可用二维平面上的等值线图来表示。

1. 等值线图编绘的方法模型

等值线图的编绘过程，实际上可看成是一种由数据点数值到图形符号（等值线）的变换过程，这种变换不可避免地包含数据点之间的内插与外推。由于各种内插和外推方法都有其各自的假设成分，因而等值线图也就具有它所采用的插值法所强加的特性（Harbaugh et al.，1968）。显而易见，为了能充分反映数据点原始数据所包含的信息，必须选择合理的插值与作图方法。

等值线的自动绘制，按其插值时所使用的网格形状，可分为规则方格网法与不规则三角形网法两种。前者是将规则方形网格点数据，或者将随机分布数据经网格化处理所得的规则方形网格点数据作为原始作图数据，然后在网格边上内插等值点，并追踪连接这些等值点而成光滑的曲线；后者是直接根据随机分布的数据点来建立不规则形状的三角形网，然后在三角形边上内插等值点并追踪，最后连结成光滑曲线。可以看出，无论是哪种方法，绘制等值线都有如下的处理过程：①离散数据点的网格化（Grid）；②依网格内插等值点（Interpolation）；③追踪等值点并形成光滑等值曲线（Contour smoothing）；④在曲线上注记等值线值，进行图件整饰（Contour label）。

不规则三角形网法比较适合于地矿资源勘查平面图的编绘。这是因为地矿资源勘查工作的实际观测点（如露头，钻孔等），往往带有随机分布的特征，对有些图件（如矿层和岩层底板等高线图、等厚度图等）需要较高的精度，以便更好地指导钻探、储量计算、工程施工和矿产开采（吴冲龙等，1992；刘刚等，1998）。不规则三角网法能直接利用随机数据点的数据进行等值点的内插计算，而无需像规则方格网法那样进行二次插值，因而等值点的定位精度较高，能充分保留原数据点的信息，而且三角网法能相对容易地处理构造线、断层线等附加信息（Petrie et al.，1987）。为此，本节重点介绍不规则三角形网法原理和程序设计方法。

2．三角网绘制等值线的原理与方法

(1) 三角网的自动联结

原则上，对于一任意随机分布的平面点集，其三角剖分不是唯一的。但为了保证等值点的内插精度，要求所形成的每个三角形都尽可能地等边并且有最小边长，避免狭长三角形的出现，此外，还要保证互不相交、互不重复。满足这一要求的三角剖分，比较典型的有以下两种：①Delaunay 三角剖分；②经向扫描算法的三角剖分。前者被国际上确认为“最佳三角剖分方案”，有较可靠的理论依据。

1) Delaunay 三角剖分的基本概念与性质。与 Delaunay 三角剖分密切相关的概念是点集的 Voronoi 图。Voronoi 图在地质、地理等科学中有着广泛的应用，有时被直观地称为“影响区域图”。

一个 $N \geq 3$ 平面互异随机点集 $\{P_i\}$ $(i = 1, 2, \cdots, N)$ 的 Voronoi 图是一个满足以下条件的凸多边形集合 $\{V_i\}$：$V_i =$

$$\{(x, y) \mid d_i(x, y) \leq d_j(x, y)\} \qquad (1 \leq j \leq N, i \neq j)$$

式中 d_i (x, y) 代表平面上任意点 (x, y) 与 P_i 点的欧氏距离。可称这样的 V_i 为 Voronoi 多边形，其顶点称为它的 Voronoi 点，P_i 称为它的 Voronoi 中心。显见，在 Voronoi 多边形 V_i 中包含的所有点 (x, y)，与 P_i 的距离比到其他 P_j $(j \neq i)$ 的距离要近，这就是说 V_i 是 P_i 的一个“影响区域”或“载负”。

如图 6-23 所示，V_i 与 P_i 是一一对应的。在一个 V_i 中，包含且仅包含 $\{P_i\}$ 中的一个点。在一般情况下，任一 Voronoi 坐标点 (x_v, y_v) 至少被三个 Voronoi 多边形所共用，并且与点集 $\{P_i\}$ 中的某三个点的距离相等，即存在 P_i，P_j，P_k $(i \neq j \neq k)$ 三点，有

$$d_i(x_v, y_v) = d_j(x_v, y_v) = d_k(x_v, y_v)$$

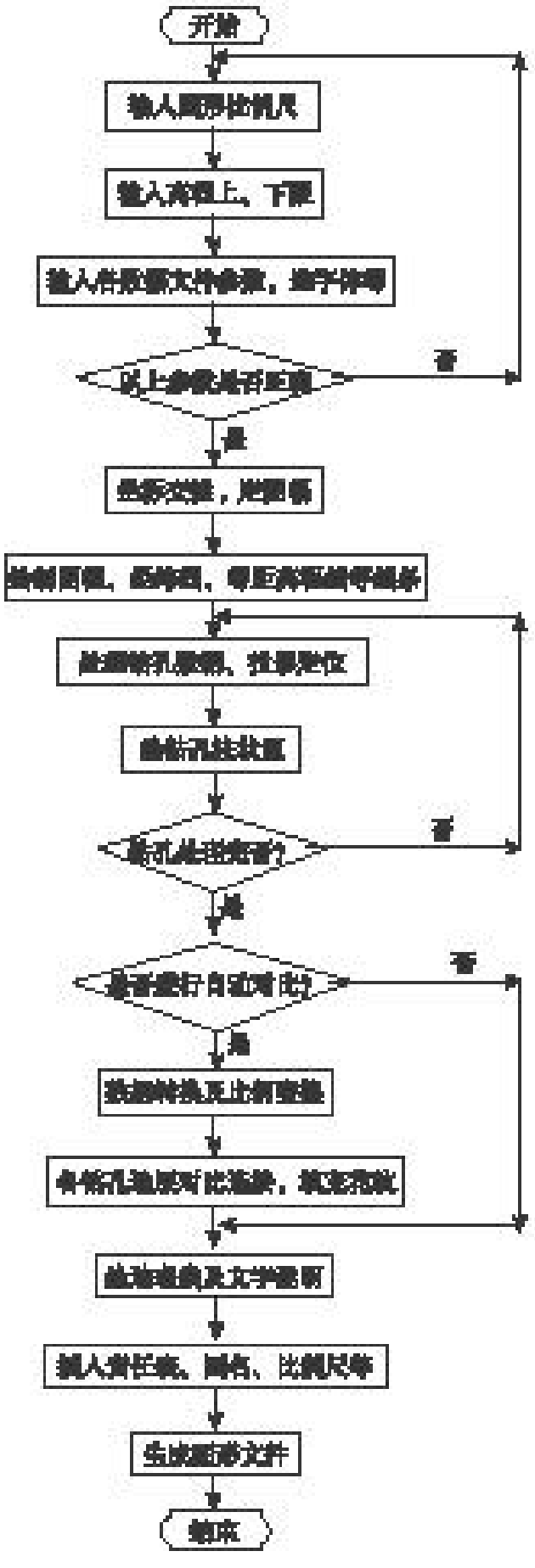


图 6-22　勘探线剖面图编绘程序框图

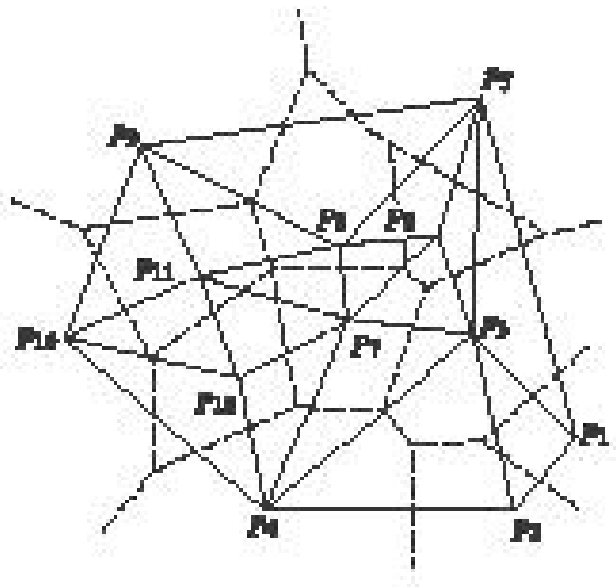


图 6-23　平面随机点集 $\{P_i\}$ 的 · Voronoi 图及其 Delaunay 三角剖分

虚线—Voronoi 图；实线—Delaunay 剖分；

“. P_i”—数据点 ($i=1, 2, \cdots$)

成立。这就是说，(x_v, y_v) 是$\triangle P_iP_jP_k$的外接圆圆心。反过来说，若 $\{P_i\}$ 中存在某三个点 P_i，P_j，P_k ($i\neq j\neq k$)，在它们所决定的$\triangle P_iP_jP_k$的外接圆内，不再含有 $\{P_i\}$ 中的其他点，则该圆心必定是一个 Voronoi 点。在退化情况下，一个 Voronoi 点被多于三个多边形所共用，并且是多边形的外接圆圆心。把相邻的 Voronoi 多边形中所对应的 $\{P_i\}$ 中的一点 P_i 用直线连接起来，并对退化情况作适当处理，则就形成了 $\{P_i\}$ 的 Delaunay 三角剖分，如图 6-23，虚线是点集的 Voronoi 图；实线是相应的 Delaunay 三角剖分。显见，在这样的三角形边上内插等值点是比较合理的，它满足插值原理中最邻近点的要求。根据 Voronoi 图与 Delaunay 三角剖分的定义，我们可利用 Voronoi 点与 $\{P_i\}$ 的关系来构造 Delaunay 三角剖分。Lawson 曾提出一种方便的 Delarnay 三角剖分算法，即最大化所有三角形的最小内角（简称 max-min 准则），并证明了在 max-min 准则下的全局最优三角剖分，就是 Delaunay 三角剖分。

值得指出的是，上述 Delaunay 剖分没有考虑数据点集所隐含的结构线信息。在隐含有结构线信息的情况下，上述 Delaunay 三角剖分就不一定合理了（张克权等，1982）。例如在图 6-24 中，若 a、c 是山谷线的两个控制点，b、d 是两个山头，则图 6-24（a）的连线方式是合理的，而图 6-24（b）的连网方式就不合理了。因为每个三角形在二维平面上形成后，就表示一个空间平面，反映了一定的地理特征（图 6-25）。由此而引出约束的 Delaunay 三角剖分方法（Auerbach et al.，1990；Falcidino et al.，1990）。

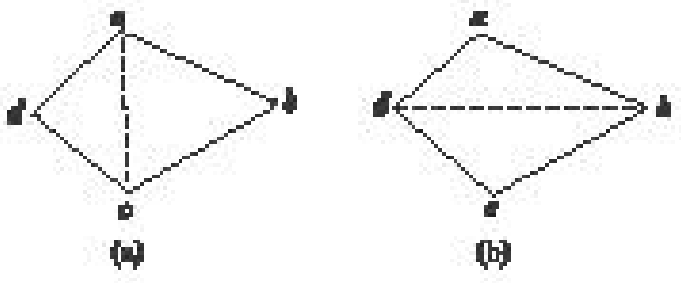


图 6-24　四个观测点的两种构网方式

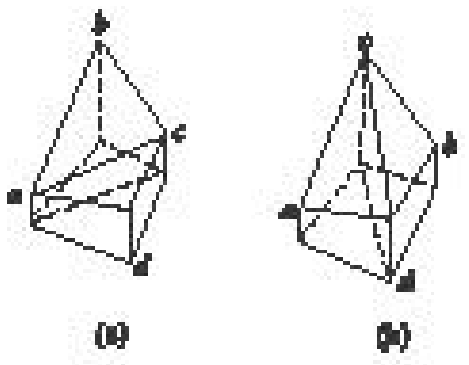


图 6-25　不同构网方式对插值的影响

约束的 Delaunay 三角剖分，就是在三角剖分的结果中，保留那些预先给定的反映结构信息的预定边 (P_i，P_j) ($i, j = 1, 2, \cdots, N, i\neq j$)，它与无约束 Delaunay 三角剖分只在与 (P_i，P_j) 边可能相交的部分有不同。据此可提供生成约束 Delaunay 剖分的两种方案：其一是在无约束 Delaunay 三角剖分过程中，掺进一定的约束算法；其二是对无约

束 Delaunay 三角剖分的结果，进行后续的局部调整修改，使三角剖分中没有边与（P_i，P_j）相交。

2）三角剖分算法。假设在 x - y 平面上，已知观测点的坐标为（x_i，y_i），（i = 1，2，…，n）；z_i 为各点的观测值，建立相应的原始数据文件（如 333DATA）；若有结构信息，则建立相应的预定义边信息。

设 L 为自动形成的三角形的顺序编号，k 为每次用来扩展的三角形编号（初值为 0），LB_1（L）、LB_2（L）、LB_3（L）分别存放第 L 号三角形的三个顶点编号。

首先，确定第一个三角形（L = 1），以相邻距离最短的任意两点连线为第一个三角形的一条边（当有约束边时，可简单地约束为第一个三角形的第一边）。其次，找出与该两点所夹角最大的点作为第三点（在有约束边的情况下，需建立一个一定容量的队列结构，按夹角大小顺序排列，当最大夹角的点所构成的三角形的边与约束边相交时，进行回溯，找出所成三角形不与约束边相交的点作为第三点）。至此已形成了第一个三角形，把三个顶点存入 LB_1（L）、LB_2（L）、LB_3（L）中。然后，令 k = 1，依次对 k 号三角形的三条边进行扩展（当 k≠1 时，第一条边无需进行扩展，因为此边已共用过两次）。当 k 号三角形的某一条边，例如第二边向外扩展时，为了避免三角形的交叉，位于 IB_1（k）同侧的点号必须被排除，可利用直线判别正负区的原理来实现这一目的。为此，建立直线方程：

$$f(x, y) = \begin{cases} y - ax - b & x_1 \neq x_2 \\ x - x_1 & x_1 = x_2 \end{cases}$$

式中：$a = (y_2 - y_1) / (x_2 - x_1)$；$b = (y_1 x_2 - y_2 x_1) / (x_2 - x_1)$。$(x_1, y_1)$、$(x_2, y_2)$ 分别表示三角形扩展边两端点的坐标。设 (x_3, y_3) 表示另一个顶点的坐标，则当 $f(x, y) f(x_3, y_3) \geq 0$ 时，(x, y) 点与 (x_3, y_3) 点同侧或在 (x_1, y_1)、(x_2, y_2) 两点所在的直线上；当 $f(x, y) f(x_3, y_3) < 0$ 时，点 (x, y) 点与 (x_3, y_3) 异侧，可考虑扩展，但能否真正成为扩展点，仍按最大夹角原理进行取舍（当有约束边的情况下，与前面类似处理）。同时，为了避免三角形的重复，还需判断此三角形是否与以前所形成的三角形发生重复。若不重复则令 L = L + 1，把该三角形的三个顶点依次存入 $IB_1(L)$、$IB_2(L)$、$IB_3(L)$ 中。此时，该条边已扩展完毕，继续对下一条边进行扩展，直至 k 号三角形的三条边均已扩展完毕，转入 k + 1 号三角形的扩展工作，直至 k≥L 且 L≠1 时为止，表示三角形的扩展工作全部结束。

注意在“退化”情况下，即当某边的一侧有多个点满足最大夹角时，可按顺时针或逆时针方向直接顺次相连（如图 6-26）。

经过上述计算所连成的三角形网，其边界线形成点集的凸包，往往在边界附近形成又窄又长的三角形，一方面人为地扩大了绘图区域，另一方面也为分片数据联网结果的拼接带来麻烦（如图 6-27）。为此，可以赋予边界点数据一定的特征值，将三个顶点均具有边界特征的三角形删除。这种方法可以把分片数据的联网结果拼接在一起，用于编绘带断层的等值线图。另外，也可采用简单限定三角形最大边长的方法，当某三角形的某边长大于限定值时，舍去该三角形。此法可以限制边界连线，也可以限制点集内部大片空白区的连线问题。

（2）内插等值点并追踪等值线

当制图区域内所有的离散点都连结成三角形网络后，即可内插等值点及追踪等值线。等值线的内插既可以是简单的线性内插，也可以是曲面函数内插，如 Bernatein Bezier 多项

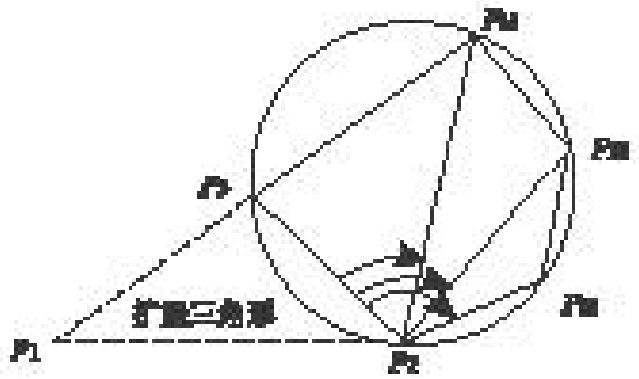


图 6-26　退化情况的处理

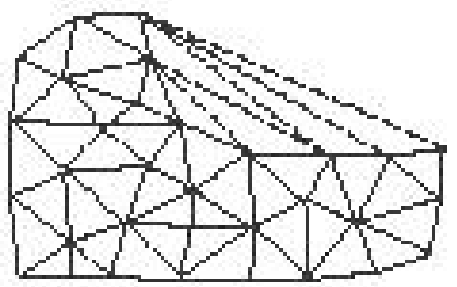


图 6-27　最大边长的限制

式、B 样条（Auerbach et al.，1990）、克立格方法以及分形几何方法等。

在三角形边上内插等值点时，为避免等值点刚好落在三角形角点上，需要对那些数值等于等值线值的数据点作微量修改，使它与等值线值有所差别。

等值点的线性内插方法较为简单，其公式是:

$$x_d = x_0 + (x_b - x_a) \times (z_d - z_a) / (z_b - z_a)$$

$$y_d = y_0 + (y_b - y_a) \times (z_d - z_a) / (z_b - z_a)$$

其中：(x_d, y_d) 为等值点坐标；z_d 为内插值即等值线值；(x_a, y_a)，(x_b, y_b) 和 z_a，z_b 分别表示三角形插值边两端点的坐标及相应观测值。

等值线有闭曲线与开曲线之分，闭曲线必定分布在区域内部，开曲线则必定始于边界并终于边界。追踪等值线时，首先从开曲线开始，即先在边界三角形的边界上，用 $(z_d - z_a) / (z_d - z_b)$ 的正负来判定有无等值点，若有，则用线性内插公式算出等值点坐标，作为开曲线的起始点，继而对该三角形的其他两条边进行有无等值线点的判断，并插值计算出等值点坐标，作为第二个等值点位置。然后，在与该边相邻接的三角形的其他两条边上找等值点，如此下去，直至追到终点为止。在此过程中，为避免已被追踪过的三角形被重新追踪，应抹去该三角形。一旦追完所有的开曲线，便转入闭曲线的追踪，在剩下的三角形中，以所找到的任何一个等值点作为闭曲线的起点，采用类似于开曲线的追踪方法进行追踪，直到又回到该点。当某一数值的等值线追踪完毕，便转入下一数值的等值线的追踪，直到全部等值线追踪结束。

为了获得比较光滑的等值线，同时避免三角形内两等值点间的长距离直线连接，可以把三角形进一步细分为许多小三角形（如图 6-28），每个小三角形顶点的数值可用一定的面插值方法求出，如距离平方倒数加权法:

$$z_m = \frac{\sum_{i=1}^{n} z_i d_i^{-2}}{\sum_{i=1}^{n} d_i^{-2}}$$

其中 d_i 为参与插值计算的点位置与插值点的距离;

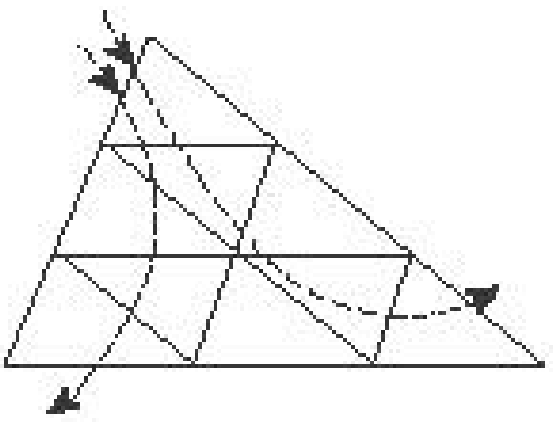


图 6-28　大三角形的细分

z_m 为插值点的值。一般地说，参与插值计算的点可以是大三角形的 3 个角点，也可以是与大三角形相邻接的三个三角形的 6 个角点。小三角形内的等值线追踪方法与大三角形的追踪类似。

(3) 曲线圆滑

经过上述等值点的追踪，等值线的分布特征已基本确定。为了进一步获得比较圆滑的等值线形态，可以再对等值点序列进行曲线逼近或拟合。不过，在采用曲线逼近或拟合方法时，可能导致等值线间出现相交的现象，需要做一些特殊处理。下面仅对抹角法（或称线性迭代法）和分段三次多项式拟合法（或称 AKIMA 法）作些简单介绍。

1) 抹角法（线性迭代法)。抹角法即是把等值线折角的两边，按等比分点公式:

$$\begin{cases} x = \dfrac{(m-1)x_2 + x_1}{2m} \\ y = \dfrac{(m-1)y_2 + y_1}{2m} \end{cases}$$

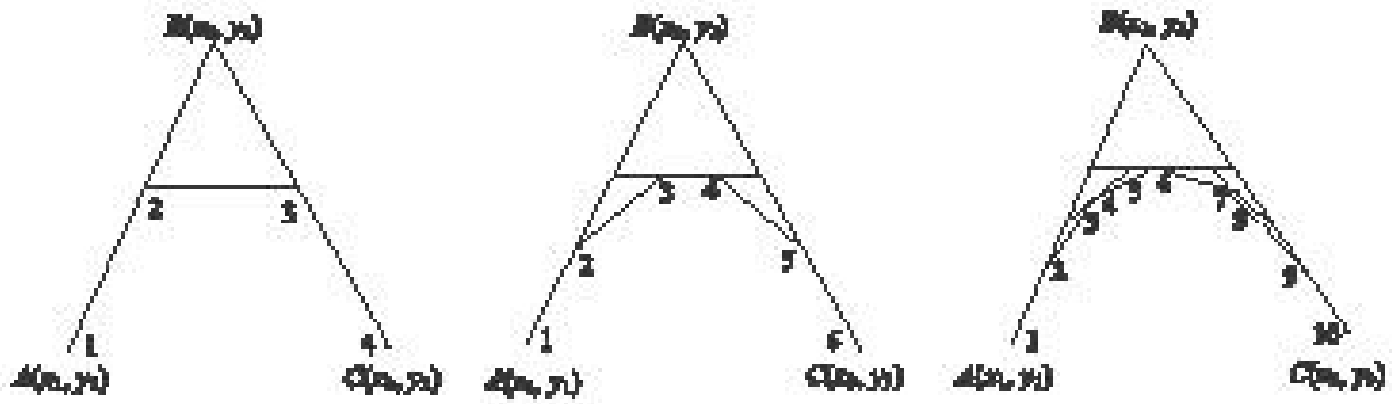


图 6-29　抹角法示意图

分别计算出各边插值点的坐标，并连接两边上最靠近折角的插值点，抹去折角，便完成了第一次迭代。这一线性插补过程可重复进行，直至曲线达到所需的光滑度（图 6-29)。

在启动圆滑程序时，可由人工输入迭代次数 N_1 和等分数 M。N_1 越大，则迭代次数越多，图形圆滑，但绘图速度慢。M 越大，光滑度越差，但等值线与实际观测值的偏差小。因此，必须根据需要合理地选择 N_1 和 M。

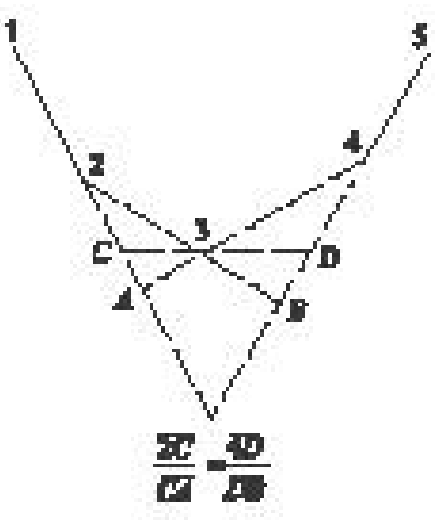


图 6-30　三次多项式曲线上各点求一阶导数的几何条件

经过抹角法处理后，组成线段的点增多了，曲线圆滑程度好，并保证曲线不会相交。增多的点数与迭代数有关(即点数 $N = 3 + \sum_{i=1}^{N_1} 2^{i-1}$)。但是，抹角法也存在精度较差的问题。采用三次多项式分段拟合的方法对等值线进行圆滑则可以提高其精度。

2) 分段三次多项式拟合法 (AKIMA 法)。三次多项式拟合的基本原理，是在两个相邻数据点之间拟合一条具有连续一阶导数的三次多项式曲线，这就需要在两点之间建立一个三次多项式方程。在三次多项式曲线上各点的一阶导数，是以该点为中心及两边分别相邻的两点，共五个点的坐标所确定的，其几何条件如图 6-30 所示。由上述几何条件可以推导曲线上任一点的导数为 $\mathrm{tg}\theta$。为了便于计算，可写成:

$$\sin\theta = \frac{b_0}{\sqrt{a_0^2 + b_0^2}}, \quad \cos\theta = \frac{a_0}{\sqrt{a_0^2 + b_0^2}}$$

这里,

$$a_0 = w_2 a_2 + w_3 a_3; \quad b_0 = w_2 b_2 + w_3 b_3;$$

$$w_2 = |\, a_3 b_4 - a_4 b_3 \,|; \quad w_3 = |\, a_1 b_2 - a_2 b_1 \,|;$$

$$a_i = x_{i+1} - x_i; b_i = y_{i+1} - y_i \quad (i = 1, 2, 3, 4)$$

由此可建立相邻两数据点即 (x_i, y_i) 和 (x_{i+1}, y_{i+1}) 之间的三次曲线方程:

$$x = p_0 + p_1 z + p_2 z^2 + p_3 z^3$$

$$y = q_0 + q_1 z + q_2 z^2 + q_3 z^3$$

其中: p, q 均为常数，z 为参变量。曲线由 (x_i, y_i) 变到 (x_{i+1}, y_{i+1}) 时，z 从 0 变到 1，故假设:

$$\text{当 } z = 0 \text{ 时}, \; x = x_j, y = y_j, \; \frac{d_x}{d_z} = r\cos\theta_j, \; \frac{d_y}{d_z} = r\sin\theta_j$$

$$\text{当 } z = 1 \text{ 时}, \; x = x_{j+1}, \; y = y_{j+1}, \frac{d_x}{d_z} = r\cos\theta_{j+1}, \; \frac{d_y}{d_z} = r\sin\theta_{j+1}$$

其中，$r = [(x_{j+1} - x_j)^2 + (y_{j+1} - y_j)^2]^{1/2}$

因此，可以推导出系数 p、q:

$$p_0 = x_j$$

$$p_1 = r\cos\theta_j$$

$$p_2 = 3(x_{j+1} - x_j) - r(\cos\theta_{j+1} + 2\cos\theta_j)$$

$$p_3 = -2(x_{j+1} - x_j) - r(\cos\theta_{j+1} + \cos\theta_j)$$

$$q_0 = y_j$$

$$q_1 = r\sin\theta_j$$

$$q_2 = 3(y_{j+1} - y_j) - r(\sin\theta_{j+1} + 2\sin\theta_j)$$

$$q_3 = -2(y_{j+1} - y_j) - r(\sin\theta_{j+1} + \sin\theta_j)$$

采用这种方法圆滑的等值线是严格通过各实测数据点的，因此精度较高。但在等值线密集区，两条弯曲方向相反的等值线有可能交叉。当出现这种情况时可以改变 z 值使等值线不再相交，此时等值线的光滑度将变差。在程序中 z 值的大小可由一控制参数来标识。

(4) 注记等值线值并输出图形

为了便于等值线的判读与使用，最后还需要对等值线的值进行标注。标注的具体位置通常是在曲线曲率较小处，标注字符的方向在 - 90 °到 90 °之间，字型的大小可根据图件的大小自由选择。当曲线较密集时，可以间隔几条等值线标注；而当图幅较大、曲线较长时，一条等值线上也可按一定间隔多处标注。

有时候，人们要求所绘制的等值线图具有经纬网和边框，甚至是套合地理底图。若借助于 CAD 的图形编辑功能，可很方便地实现。

3. 数据预处理

在进行等值线编绘时，应先进行数据的预处理，即直接用编辑程序建立如下格式顺序的原始数据文件:

x_1，　y_1，　z_1

x_2，　y_2，　z_2

…，　…，　…

也可以在每个数据点前赋予点号，即:

No1，　x_1，　y_1，　z_1

No2，　x_2，　y_2，　z_2

…，　…，　…，　…

该文件定名为 333DATA。

假若用户具有一些三角联网的“结构”控制信息，可建立一个名为 STRU 的文件，其格式如下:

line1，n，No_{i1}，No_{i2}，No_{i3}，……No_{in}

line2，n，No_{j1}，No_{j2}，No_{j3}，……No_{jn}

… … … … … … …

其中: line1，line2 表示特征线名，n，m 分别为每条特征线所具有的点数，No_{i1}，No_{i2}，No_{i3}，…… No_{in}，No_{j1}，No_{j2}，No_{j3}，……No_{jn} 为数据点点号。

编绘图件所需的数据文件通过数据库来建立。用户只需在数据库中选择相应的应用专题，即可获得与该图件有关的全部数据。

4. 程序功能与程序框图

本教程所附的实习程序可以处理随机分布的原始数据，适用于不同勘探网密度、线距和孔距。图形计算子模块有: ①三角网的自动联结；②内插等值点并

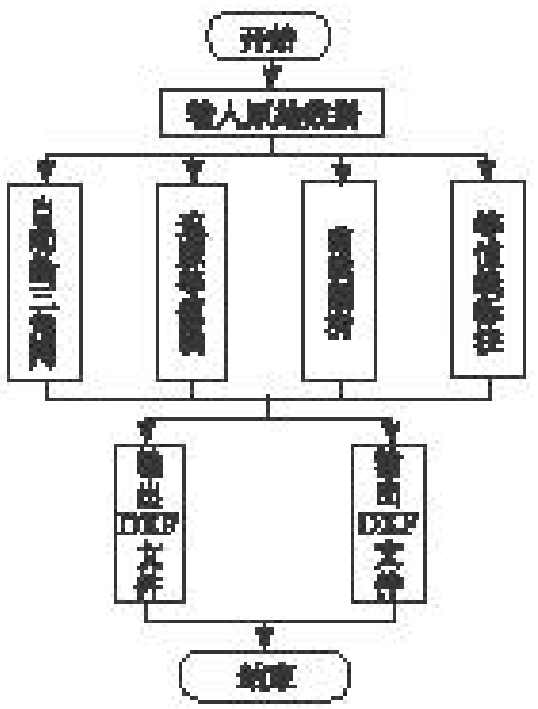


图 6-31　三角剖分法绘等值线主框图

追踪等值线；③曲线圆滑；④等值线注记子程序。

针对某些地质平面图需要分块绘制等值线，例如按不同级别的储量块段计算绘制等值线，系统中设置了程序的“批处理”功能。用户可以根据需要选择使用。

该程序的主体框图如图 6-31 所示，绘制（追踪）等值线程序的框图见图 6-32。

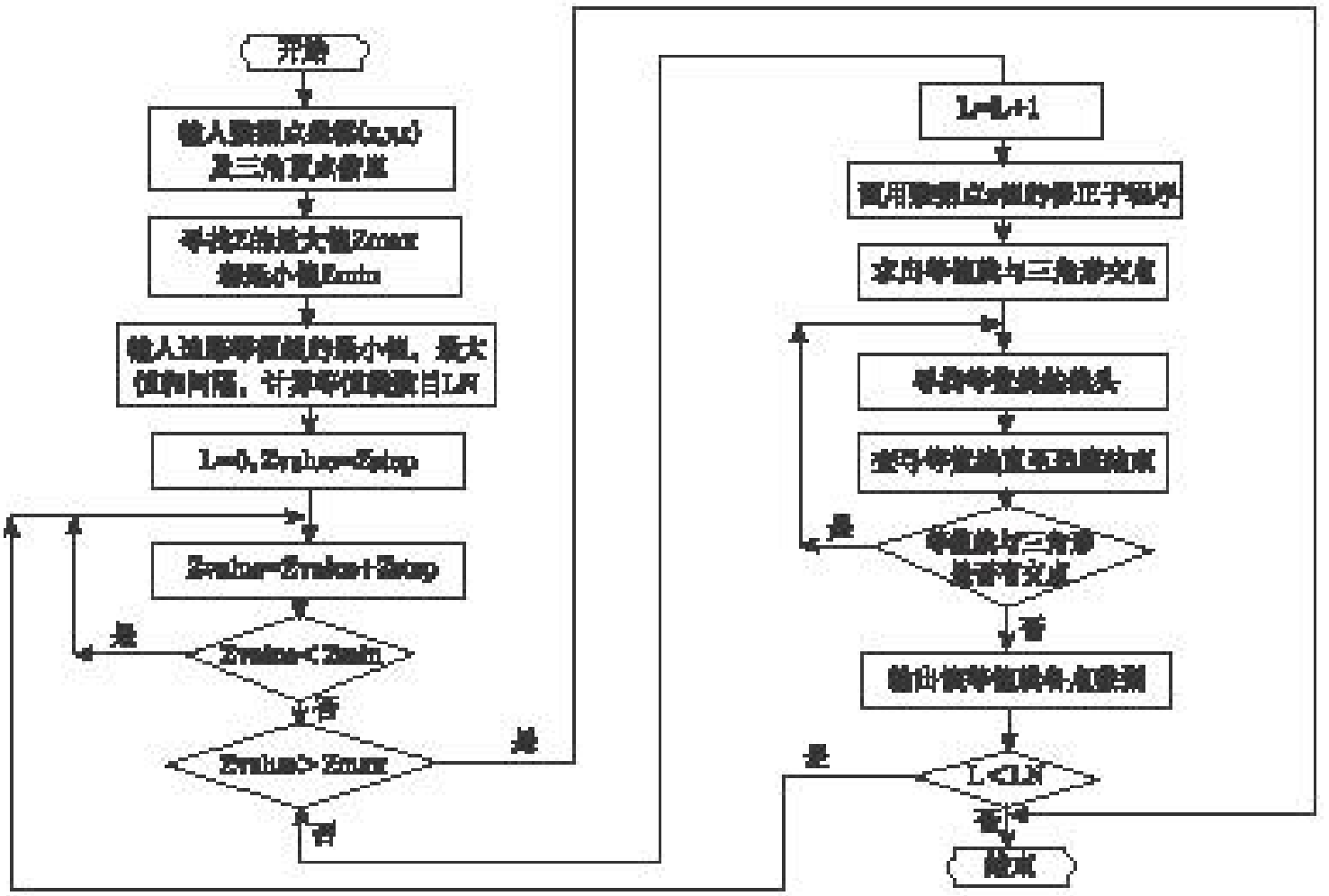


图 6-32　等值线追踪子程序框图

第四节　地矿资源图件机助编绘子系统的应用

本节主要讲述应用资源图件机助编绘子系统的一般过程和方法，其中包括基础图形系统的选择、应用图形系统的建立和应用。

一、基础图形系统的选择

基础图形系统的选择关系到二次开发工作的开展和开发成果的推广应用。下面介绍三种目前国内外流行的微机通用绘图系统及其选择原则。

1. 通用绘图软件 AutoCAD 简介

(1) 主要功能特色

当前在我国流传较广且深受用户欢迎的一种通用的 CAD 软件是 AutoCAD。它是美国 AutoDesk 公司于 1982 年推出的，现已发展到第 15 版。AutoCAD 的坐标系统是基于向量浮点运算的坐标系统。由于采用浮点算法，精度可达到 $10^{15}\times 10^{15}$ 以上。从软件功能上看，它提供了一组输入基本图形元以构成各种图形的命令，如直线、圆、圆弧、线条、文字等；提供了一些可供修改、编辑图形的命令，如擦除、移动、缩放、拷贝等；还可以测量

某些参数如面积、周长等。它可以用不同层次、不同颜色的线型和字型，在同一画面上形成二维平面图形及三维立体图形，将屏幕上画好的图形用不同比例在绘图机或图形打印机上输出，并具有菜单功能和制作各种用户图库的功能。它可以与高级语言对话，可运行DOS的内部和外部命令，还配置了可供用户编程的AutoLISP语言及与C++语言的接口，供用户进行二次开发以扩充AutoCAD软件包的功能。

(2) AutoCAD的多层次二次开发工具

AutoCAD的二次开发体系包括其内部三代开发环境和外部应用程序开发，如图6-33所示。其中包括：①第一代开发环境：内嵌的解释执行且交互性好的AutoLISP语言；②第二代开发环境：基于C语言的ADS/ADSRX结构化的面向过程的开发工具；③第三代开发环境：基于C++的ObjectARX面向对象的开发工具及可编译的Visual AutoLISP开发工具；④基于ActiveX Automation的面向对象的解释型开发工具；⑤基于Script文件和DXF文件的开发接口；⑥基于DDE的进程通讯(IPC)开发接口。

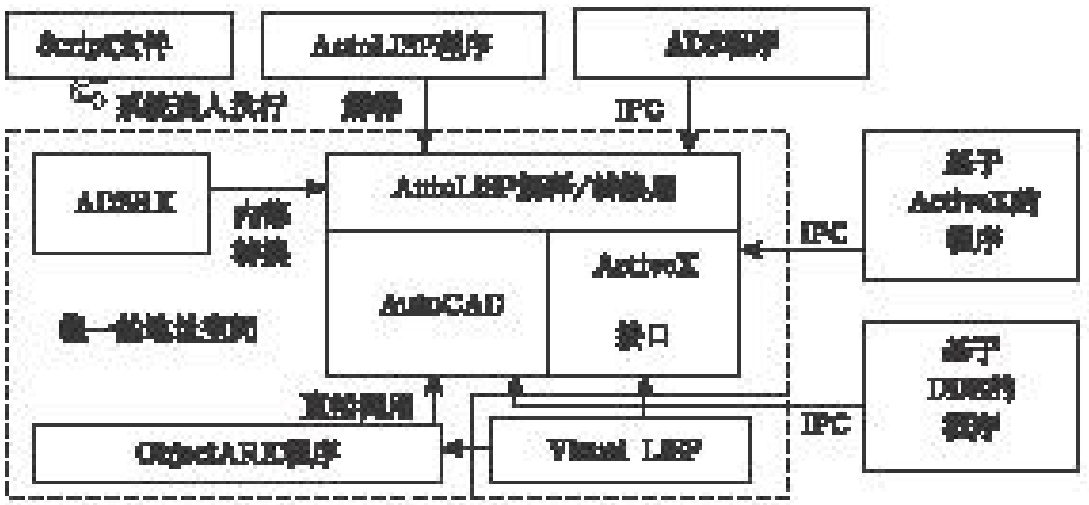


图6-33 AutoCAD的多层次二次开发工具(1999)

2. MicroStation计算机辅助绘图系统

(1) 功能特色

MicroStation是著名的美国Intergraph公司的CAD软件产品。MicroStation的前身是IGDS(Interactive Graphics Design System)，是Intergraph公司研制的大型CAD系统，主要运行于以VAX为主机的工作站上。在20世纪90年代后期，Intergraph公司将IGDS移植到微机上，命名为MicroStation。现在的MicroStation已能在各种软硬件平台上运行，如DOS、Windows、Macintosh及SUN Sparc、HP RS6000、Intergraph CLIX等。

MicroStation与AutoCAD在功能和用户界面上具有很多相似之处，这里仅介绍其较之AutoCAD更具特点的几个方面。以5.0版的MicroStation为例，其主要特点是：

◆ 全面、完善的图形式用户界面

虽然AutoCAD在12版之后也提供了图形式用户界面(GUI—Graphic User Interface)，但是，MicroStation的图形式用户界面更加全面和完善。MicroStation通过主工具板(Main Pallettes)和子工具板以图形方式向用户展现几乎所有的绘图功能。

◆ 工业标准的数据格式

MicroStation可实现DWG、DXF、IGES和CGM等图形数据格式的互换；可以直接从

dBASE Ⅲ/Ⅳ、Foxbase、Oracle 等数据库中读取数据作为参数供图形设计使用。

Windows 版本的 MicroStation 还可以利用剪贴板功能从 Windows 的某个应用程序中拷贝或粘贴数据。这一动态数据交换特性使得应用开发者不仅可以在图形设计文件中粘贴数据，而且还可以增加语音注释。

◆ 强大的光栅文件处理能力

MicroStation 除了能观察并绘制全部 Intergraph 支持的光栅文件（BUMP、COT、RGB 和 RLE）外，还可以接受其他格式的光栅文件，如 IMG、JPEG、JFIF、TIFF、PCX、GIF、Targa 、Sun Raster 和 Windows 的 BMP 文件。利用 Raster 文件转换实用程序还可以方便且有效地将上述文件转换成更为通用的光栅文件，如 PICT、PostScript 和 WordPerfect WPG 格式。

◆ 智能式图形关联

AutoCAD 中有一个图形关联方法，需要将重复使用于绘图中的某些图形元素定义成“块”(block)，如果要利用修改过的“块”来替换图形中的所有的“块”，必须用“insert”命令来实现。MicroStation 则是通过定义图形的关联关系来实现。当某个图形元件修改或更新之后，与之关联的图形自动更新，因此“智能”化程度更高。

◆ 可以引用更广泛的字体

MicroStation 不仅支持自己和 Intergraph IGDS 的各种字体，而且可以安装并引用 AutoCAD、Postscript、Type 1、TrueType 等格式的字体。

◆“在线”式例子文件使用户能更快、更正确地掌握 MicroStation

“在线”式帮助是很多软件都提供的功能，但大部分软件的帮助只显示命令的格式和功能，而 MicroStation 的例子文件非常丰富而且实用。

◆ 为 AutoCAD 用户提供移植工具

MicroStation 提供了一套移植工具，使得用户能够在它与 AutoCAD 上所作的应用开发均可以互相移植，包括图形文件、字体、线型、块及其属性等。

(2) 开发语言

MicroStation 可以直接用于绘制图形，同时又是一个 CAD 支撑软件。MicroStation 用于应用软件开发的程序设计语言是 MDL (MicroStation Development Language)。

MDL 属于伪编译类型的 C 语言，它具有完善的 Motif 菜单制作功能和强有力的图形函数。MDL 语言的语法与标准 C 的语法完全相同。MDL 500 多个函数中的大部分与 ANSI C 语言的标准函数完全兼容。在许多情况下，只需要将在 MicroStation 平台上开发的原文件传送到目标开发平台上，再重新编译，生成可执行文件即可完成移植。

3. 国产绘图与地理信息系统软件系统 MapGIS

(1) 功能特色

MapGIS 微机地理信息系统是中国地质大学（武汉）中地公司开发的工具型地理信息系统软件。它是一个集当代先进的图形、图像、地质、地理、遥感、测绘、人工智能、计算机科学等于一体的高效全汉字大型智能软件系统，是集地图输入、数据库管理及空间数据分析为一体的空间信息系统。MapGIS 的前身是 MapCAD，是从矢量图形处理逐步发展起来的，其图形的输入、编辑处理、管理和输出（印刷出版）的能力很强，尤其对彩色地图的面色处理和全中文化的操作方式适合于我国的用户。

MapGIS 软件的图形辅助设计部分由以下子系统组成:

1) 数字化子系统。实现空间信息从模拟式到数字式的转换，可通过手扶跟踪数字化，输入点、线和区域的空间位置，也可通过扫描仪扫描输入，然后自动或半自动矢量化确定实体的空间位置。

2) 拓扑处理子系统。可对图形中的位置结构建立拓扑关系，从而使搜区、检索、造区更加快速、方便、简捷，并提供自动、半自动或手动生成、检查和校正拓扑关系的工具。

3) 图形编辑子系统。用来编辑、修改矢量结构的点、线、区域的空间位置及其图形属性，增加或删除点、线、区域边界，并适时自动校正拓扑关系，可对图形数据库中的图形进行编辑、修改、检索、造区等，从而使输入的图形更准确。

4) 数据交换子系统。该子系统为 MAPGIS 系统与其他 CAD/CAM 软件系统间架设了一道桥梁，实现了不同系统间所用数据文件的转换，从而达到数据共享的目的。

5) 数字高程模型子系统。该子系统可将某一区域的大量单变量值，如重力异常值、磁异常值、元素含量等离散数据进行网格化、插值处理，变成覆盖工作区域的栅格数据。

6) 地图数据库管理子系统。该子系统用于存贮和管理地图信息，可形成并输出绘图文件或各类地理数据文件。

7) 图像分析子系统。该子系统提供处理和分析图像数据的能力，可实现对遥感影像、航片影像及各种专业图像的处理分析。

8) 地图投影变换。提供从一种地图投影系统变换为另一种投影系统的手段和方法。

(2) 二次开发接口

该系统的二次开发接口提供了输入、编辑、输出、地图库、属性库、投影变换、图像处理、网络分析等 API 函数库和动态连接库。目前所适用的语言是微软的 Visual C^{++}。

4. 基础软件平台选择的基本原则

对于地矿资源图件编绘，基础软件平台的选择需要考虑以下几方面因素: ①硬件平台的适应性: 我们的专业应用环境 (如地质队) 一般是微型机，因此要选择基于微机的 CAD 软件; ②软件平台的功能适应性: 各种 CAD 软件系统都有各自的特色，在开发应用系统时必须考虑到相应软件平台的基础功能是否具有拓展能力，例如系统在汉字处理、线型制作、花纹填充、面色处理、图形数据交换、数据库连接和界面设计等; ③系统二次开发工具的针对性: 一般 CAD 软件系统都提供了多种开发接口，其中有属于程序语言型的，也有数据交换型的，选择系统时要考虑到自己的开发能力。要选择开发能力强、集成度高、交互性和通用性好的开发工具，以利于应用程序开发的连续性、适应性和先进性。④软件平台的可持续发展性: 软件平台是地矿资源图件机助编绘系统的基础，选择时特别要考虑基础平台的升级和可持续发展特性。

二、实用图形系统的建立与应用

实用图形系统的建立，是根据用户需求利用基础图形软件开发编图程序，通常也称为应用开发或二次开发。目前，实用图形系统在地矿调查与勘探中的应用已经十分广泛。本节以钻孔柱状图、勘探剖面图和储量计算平面图为例，介绍一个实用图形系统的应用。

1. 实用图形系统的建立

在地矿资源信息系统的框架下，大量复杂的基础数据由数据库子系统负责分类录入和

管理，而图形数据的显示和输出由基础图形系统完成，实用编图系统主要负责实现数据的分析、计算，以及与其他子系统的协调。

(1) 与数据库实现连接

与数据库连接的方式有两种，一种是由数据库应用专题完成数据检索，交由编图程序处理；另一种则由编图应用程序通过数据库连接技术（如 ODBC 等）直接完成读库和检索操作。两种方式各有所长，前者数据管理检索方便，容易进行复杂的数据提取；后者编图应用程序开发方便，集成度高。

(2) 建立专业化的系统库

系统库主要包括线型库、花纹库、图例库和操作工具库等。

(3) 改造基础图形系统的操作界面

对基础图形系统的操作界面进行改造，以满足普通专业应用技术人员的需求。

(4) 将数据转换为图形

进行数据的分类计算，按规范实现由数据到图形的转换，包括比例转换、曲线圆滑、图形优化和显示控制等，将图形显示在屏幕上。

(5) 实施图形计算机辅助编绘

对自动生成的地矿资源基础图形进行人机交互编辑、处理和输出绘制，同时还可通过对数据的综合分析和处理，提取深层次的图形信息，编绘出各种分析图件。

(6) 图件综合管理和应用

对已有图件进行综合管理和应用，如快速查询、检索和显示等。这项工作涉及生成图形的前期、中期和后期管理，属于高层应用。

2. 钻孔柱状图编绘模块的应用

按照前面所述的设计方法，可以设计出各种类型资源勘查钻孔柱状图编绘模块。

(1) 一般应用步骤

在实际应用过程中，一般有以下几个步骤：①从数据库检索得到所需的原始数据，并对数据进行正确性检查；②在应用程序中设置有关图形处理参数，如比例尺、文字字高及图形格式等；③调用基础图形系统生成并显示图形；④检查生成的图形是否满意，否则修改处理参数，并重新生成；⑤最终图形存盘。

(2) 综合柱状图的编绘

1) 分栏绘制。一幅钻孔综合柱状图由地层时代、钻孔结构和岩心柱状图、孔深、底面高程、厚度、岩性描述、岩心采取率、岩石风化程度、工程质量等级、测井曲线、备注等许多栏目组成。有时我们会在最终成果图中发现某些栏目需要修改。本综合柱状图编绘模块提供了分栏成图与绘制的功能，以便原始数据修正后重新生成该栏目，不影响图中其他部分。

2) 分幅绘制。有的钻孔孔深比较大，本综合柱状图编绘模块提供了解决图形和文字的按给定长度分幅绘制的算法，同时图形整体格式的设置也做了相应的调整（图 6-34）。

3) 操作步骤简介。钻孔柱状图编绘的主要步骤如图 6-35 所示。其中，(a) 是柱状图编绘模块的主界面；(b) 是绘图内容的选项操作界面；(c) 是图头样式举例；(d) 是分幅绘制的钻孔柱状图。

3. 勘探剖面图编绘模块的应用

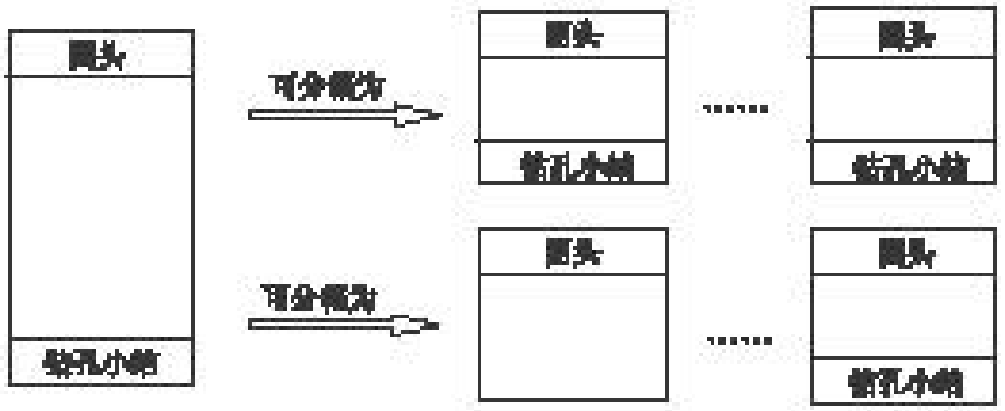


图 6-34　钻孔柱状图分幅绘制图示

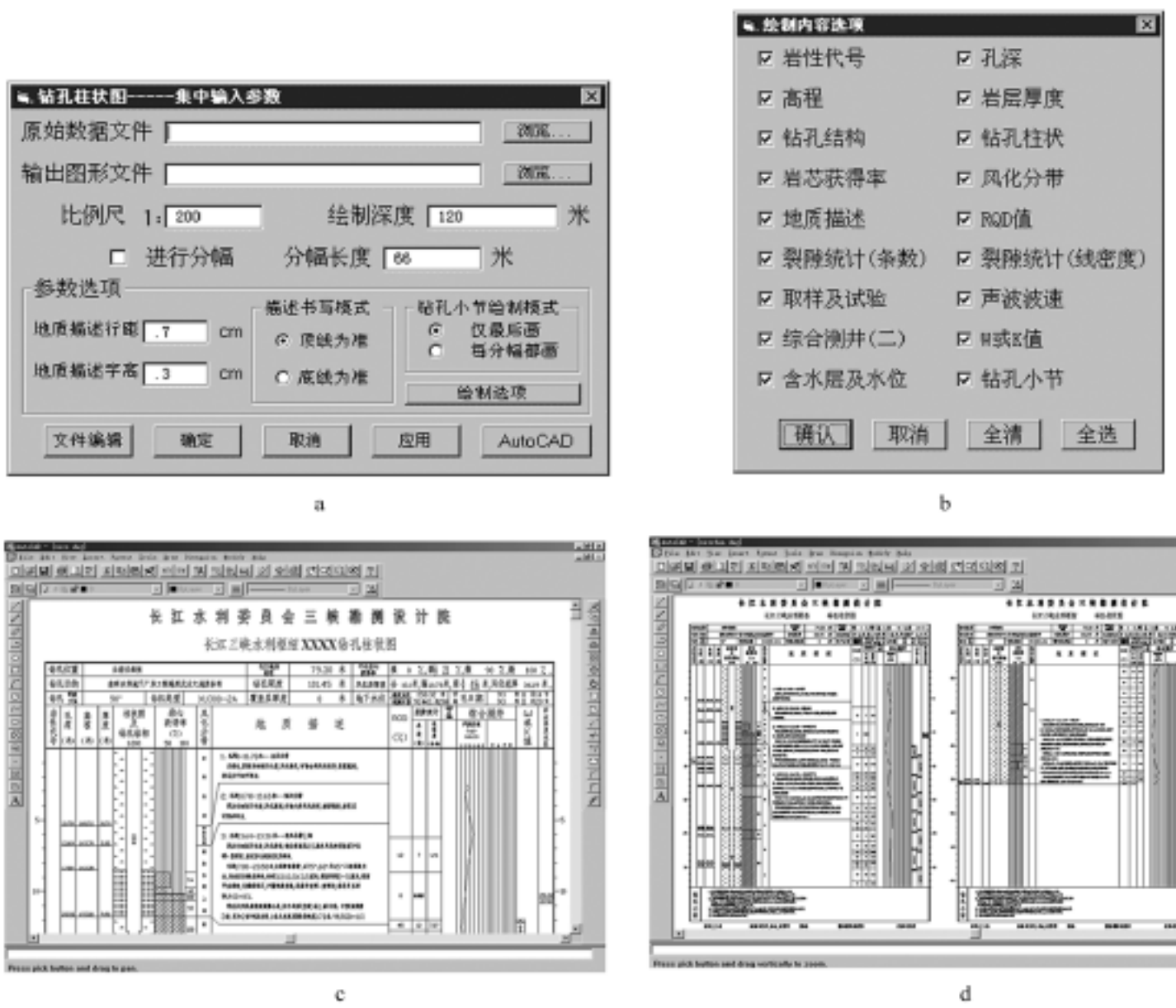


图 6-35　钻孔柱状图编绘操作图示

在地矿勘查与地学研究中，常用的剖面图有地形剖面图、地表实测地质剖面图、各种矿产地质勘探剖面图（包括工程地质勘探剖面图、水文地质勘探剖面图、金属非金属地质勘探剖面图、煤田地质勘探剖面图、石油地质勘探剖面图、地震勘探剖面图等）、构造剖面图和沉积剖面图。勘探剖面图编绘模块的应用步骤与钻孔柱状图的编绘基本一致。

（1）地质勘探剖面的内容和图面结构

地质勘探剖面的编绘内容包括：①确定所使用的坐标系统，根据剖面的起终点位置和钻孔资料确定剖面横向、纵向范围，并绘制相应的坐标网和标尺；②将地表资料在剖面上进行定位，如地形线、地表地质界线及地质现象、地表工程、地下工程在地表的出口等；③将地下资料在剖面上进行定位和标注；④地质界线及矿体边界等地质特征界线的联接；⑤在图的一侧编制取样及分析结果等附表。最后还要标上图名、比例尺、图签和图例等。

(2) 勘探剖面图编绘模块的实际应用

在绘制剖面图的过程中，有几个问题需要解决：一是剖面所需钻孔数据的自动搜索问题，特别是可视化搜索；二是钻孔投影和孔斜校正的问题；三是过钻孔的断层表达、岩层的自动对比连线和岩性花纹的自动填充问题。第一个问题在数据库应用一章中已经有所介绍，这里不再重复。下面着重介绍第二个问题和第三个问题的解决办法。

1) 钻孔投影和孔斜校正。勘探设计要求钻孔沿着勘探线剖面施工，但实际施工的结果，钻孔总是或多或少偏离了勘探线。因此，在制作综合图件时，要求将钻孔资料投影到剖面图和平面图上，以便正确地进行地质界线联接和推断。

实际的钻孔中心线是一条十分复杂的空间曲线。通过间隔的钻孔倾角、方位角测量得出一条折线是实际钻孔中心线的近似表达。测量的间隔越小，角度测量越准确，两条线就越接近，误差越小。钻孔投影可以分解为：①钻孔中心线分段，②钻孔分段投影及各段联接，③钻孔中地质界线的投影。这里讨论一下钻孔中心线分段问题。

A. 根据间隔测量得到的钻孔顶角（或倾角）和方位角的数据绘制钻孔投影

首先根据测点深度将钻孔中心线分段。方法有将测点深度作为分段节点和将相邻测点中点作为分段节点两种。在后一种情况中，得到测点深度和测量结果后，首先计算控制深度。设在 i 点及 i + 1 点的测点深度分别为 h_i 和 h_{i+1}，在这两点之间的控制深度 h_i' 为

$$h_i' = h_i + (h_{i+1} - h_i)/2$$

控制深度包括开孔深度及终孔深度。然后求出控制距离。控制距离为相邻控制深度之差。

B. 线段投影及各段联接

了解线段投影原理是钻孔投影的基础。由于勘探线地质剖面垂直于研究地质体走向布置，所以最常用的是垂直投影。下面介绍垂直投影的方法原理。如图 6-36 所示，AB 是空间线段，A 点在垂直投影面 P 上（投影面 P 与勘探线剖面平行），B 点在水平投影面 Q 上。从 A 点向 Q 平面投影，其投影点为 O。线段 OB 为 AB 的水平投影。OB 方向为线段 AB 的方位，令线段方位角为 ω。线段 AB 的倾角为 β（∠OBA），顶角为 α（∠OAB）。过 B 点向 P、Q 两平面的交线引垂线交于 C 点。OC 方向为投影面 P 的方位，也就是剖面的方位，令剖面方位角为 ε。角 （∠BOC）为剖面方位与线段方位的夹角。AC 为 AB 在 P 平面上的投影。BC 的长度即为 B 点偏离剖面的距离。这样一来，将线段 AB 分解为在垂直方向上的分量 Δz（AO）。在剖面方向上的分量 Δx（OC）

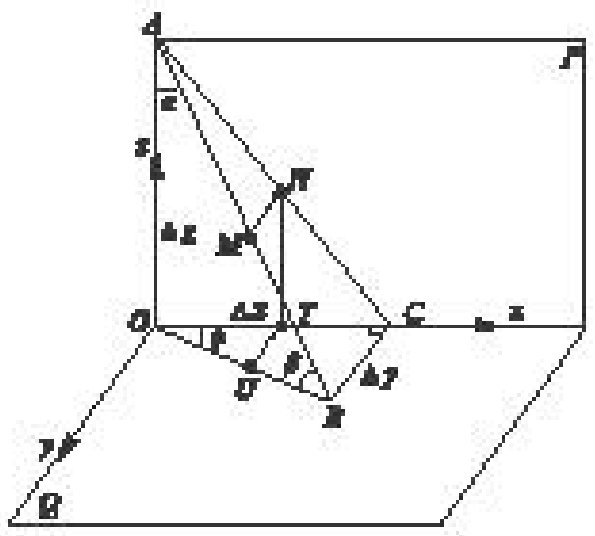


图 6-36　钻孔投影计算

和偏离剖面的分量 Δy（BC）。线段投影就归结于这三个分量的求取。求取的方法有计算法，投影制图法和量板法。这里简单介绍一下计算法。

令 AB 长度为 l，如图 6-36 所示，可得：

$$\Delta z = l \cdot \sin\beta$$

$$OB = l \cdot \cos\beta$$

$$\Delta x = OB \cdot \cos = l\cos\beta \cdot \cos$$

$$\Delta y = OB \cdot \sin = l\cos\beta \cdot \sin$$

式中：　= ω - ε。

在图 6-36 中，x 为剖面方向；y 为垂直于 x 的方向（即偏离剖面方向），在 x 方向左侧为正；z 为铅直方向，向上为正。剖面左端点 x = 0，y = 0，z 为高程。并令孔口坐标为（x_0、y_0、z_0），z_0 为孔口高程，x_0，y_0 为在上述坐标系的坐标。

令第 i 个控制深度对应点的坐标为 x_i、y_i、z_i，则与第 i + 1 个控制深度对应的坐标为：

$$x_{i+1} = x_i + \Delta x$$

$$y_{i+1} = y_i + \Delta y$$

$$z_{i+1} = z_i - \Delta z$$

若孔口坐标已测定，求得各段的坐标增量 Δx、Δy、Δz 以后，就可依次求得各控制深度对应点的坐标。将这些点投在剖面图上并联接起来，就得到钻孔中心线的剖面投影。

钻孔投影和孔斜校正技术较为成熟，有现成的子程序供调用。

2）过钻孔断层的表达。这个问题涉及岩性盲区的镂空技术。地质矿产勘探剖面上所标绘的钻孔宽度是超比例的。设 AB 为断层线，E 为 AB 与钻孔中心线的交点，CD 过 E 点且垂直于钻孔中心线，则当有断层线通过时，△ACE 和△EBD 内的岩性实际上是不知道的，可称为岩性盲区。最好利用闭合区删除技术，将其挖去，实现过孔断层的正确表示（图 6-37）。

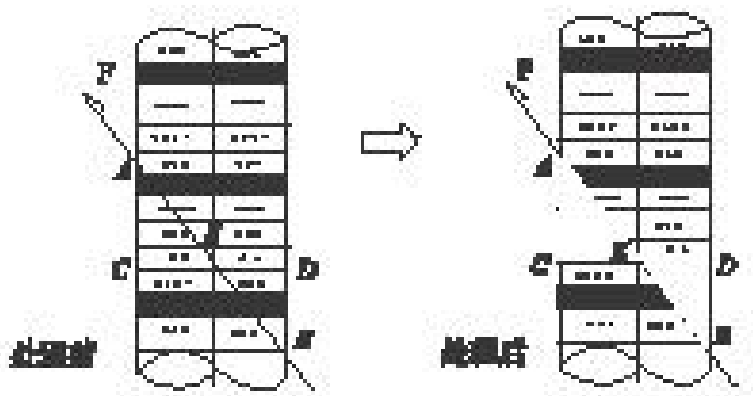


图 6-37　钻遇断层表示法

3）岩层自动对比连线和岩性花纹自动填充。岩层自动对比连线和岩性花纹自动填充涉及复杂的分析判断，需要采用人工智能技术。这里介绍一种适用于第四系的工程地质剖面自动对比连线和岩性花纹填充技术。

该项岩层对比技术要求遵循以下原则：①最小误差原则：本着减少错误的原则，先将第四系岩层按不同岩性组合进行划分，再在组合内进行自动对比；②平等原则：各钻孔的每一岩层都具有相同参与对比的机会，实施对比时，两钻孔应相互参照；③瓦尔特相律：参照瓦尔特相律和实际资料总结的“相变”、“粒序”规律，如粗砂、中砂、细砂或砾石出现的先后次序或包含关系，来修正对比结果；④松散必平缓原则：第四系松散沉积地层通常未经变形，在多数情况下不可能太陡，对比中应予以充分考虑；⑤厚度兼顾原则：在考虑岩性的同时兼顾地层厚度因素，即除了出现透镜状岩层，同一地层的厚度在相邻钻孔中相近。

由于地质现象的复杂多变，系统的自动对比功能只能说是给出了图形的一个基本框

架。系统开发者的重要任务是给地质工作者提供一个方便灵活的图形编辑环境，例如尖灭点的灵活移动，花纹的随意修改和替换等等。

4）勘探剖面图的编辑和输出。通常，只要用户提供比例尺、剖面范围以及钻孔资料数据等参数后，剖面图编绘程序可以自动生成剖面图的主体框架和钻孔岩性花纹的填充，而用户的主要工作是进行交互式剖面对比连线和部分图件整饰工作。主要步骤如图 6-38 所示：图（a）进入剖面图编绘程序主界面；图（b）分层显示与分层操作；图（c）地质界线交互联接和花纹填充；图（d）成果图打印输出。

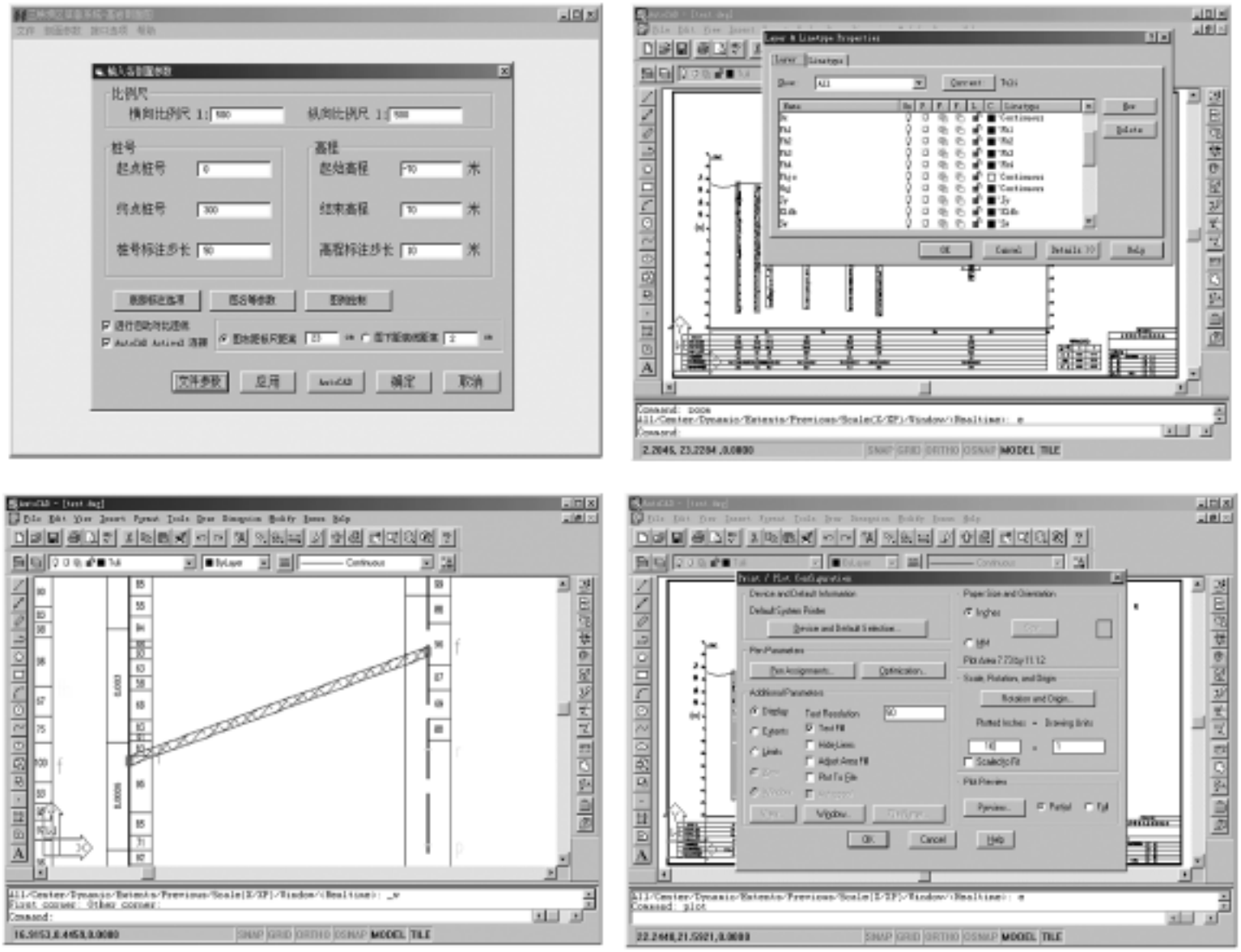


图 6-38　勘探剖面图计算机辅助编绘过程

4. 储量计算图编绘程序的应用

储量计算图是最具代表性的矿产地质勘探平面图，往往由被断层错断的不连续等值线图与钻孔分布图、钻孔矿层（煤层）结构图、储量计算参数、地理底图等组合而成。因此，在编绘储量计算平面图时，要解决不连续型等值线图的编绘问题，其次要解决钻孔矿层（煤层）结构图（俗称矿层小柱状图）的定位编绘问题。

（1）等值线图及其应用

等值线图在地学领域有着非常广泛的应用。例如地形等值线图能很好地反映出地表高程的变化；地球化学领域用某种化学元素含量等值线图可以很清楚地看出研究区内元素分

布状况；在地球物理领域可用等值线图来表示大地电阻的分布或磁场强度的变化等等。

等值线图编绘模块的应用步骤与钻孔柱状图和地质剖面图的编绘基本一致。等值线图往往作为某个综合图件的基本图来使用，因此还要能够与研究区平面图方便地叠合。

现在，有许多大型的专业图形系统都包含有等值线处理模块，如 ARC/INFO 软件包可以连接不规则三角网绘等值线模块（TIN），也有一些成熟的专用绘等值线商业软件，如美国 GOLDENSOFTWARE 公司的 SURFER 软件包。

(2) 储量计算图编绘

不同的储量计算方法通常采用不同的图件来表达。这里以常用的块段法为例，介绍储量计算图编绘的原理和工作流程。

1) 储量计算图的内容与图面结构。储量计算图的编绘内容与勘探剖面图大体一致，包括：①要确定所使用的坐标系统和图幅范围；②将地表资料在平面图上进行定位和分层处理；③将地下资料在平面上进行定位，并标注钻孔编号、孔口高程及孔深等；④地质界线及矿体边界等地质特征界线的联接；⑤在煤田储量计算图中，有时需要绘制重点钻孔标志层的煤层小柱状图；⑥勾绘块段边界，分等级标注块段符号，并分块绘制研究矿层的底板或顶板等值线；⑦在图的一侧标上图名、比例尺、图签和图例绘制等图件整饰。图名和图例的位置可以依照图面负载均衡的原则进行调整。

2) 断层助判子系统的方法原理。把等值线图编绘技术用于储量计算平面图编绘，主要难点在于实现非连续型等值线图的编绘。其关键技术问题是如何提取断层与分块输出等值线。

三角网法绘制等值线是模拟人工插值方式进行的，除了可以满足矿（岩）层底板等高线图的精度要求外，还为建立断层助判系统并实现等高线在断层两侧自动错开，提供了可能和方便。将三角网法绘制等值线与断层助判、等值线错移等技术结合起来，便可建立一个完整的矿（岩）层底板等高线图编绘程序（刘刚等，1999）。

为了在矿（岩）层底板等高线图中插入断层，可以在程序中引进二维散布点趋势分析和趋势剩余分析方法。其方法原理是：

以多重回归分析的理论为基础，把地质特征的观测值分解为趋势部分和剩余部分。前者用以了解该地质特征观测值的总体变化规律，后者用以分析地质特征的局部异常。

断层作为一种构造上的局部因素，在矿（煤、岩）层底板高程趋势图上，通常表现为等值线密集带；而在趋势剩余图上，通常表现为等值线密集带和正负剩余转换带。

二维散布点趋势分析与剩余值求法简述如下：

设有 n 个样品，横纵坐标值分别为 x_i，y_i（i = 1，2，…，n）观测值为 z_i（i = 1，2，…，n）。可建立回归方程：

$$Z = G \cdot A + \varepsilon$$

其中：G 为 x，y 坐标多项式矩阵（n× k），若多项式次数为 s，则 k = (s + 1) ·(s + 2) /2，而 ε为偏差。

按最小二乘准则，求出系数 A 的估计量：

$$A = (G' \cdot G)^{-1} G' \cdot Z$$

则趋势值为 $\hat{Z} = G \cdot A$，剩余值为 $\Delta Z = Z - \hat{Z}$，

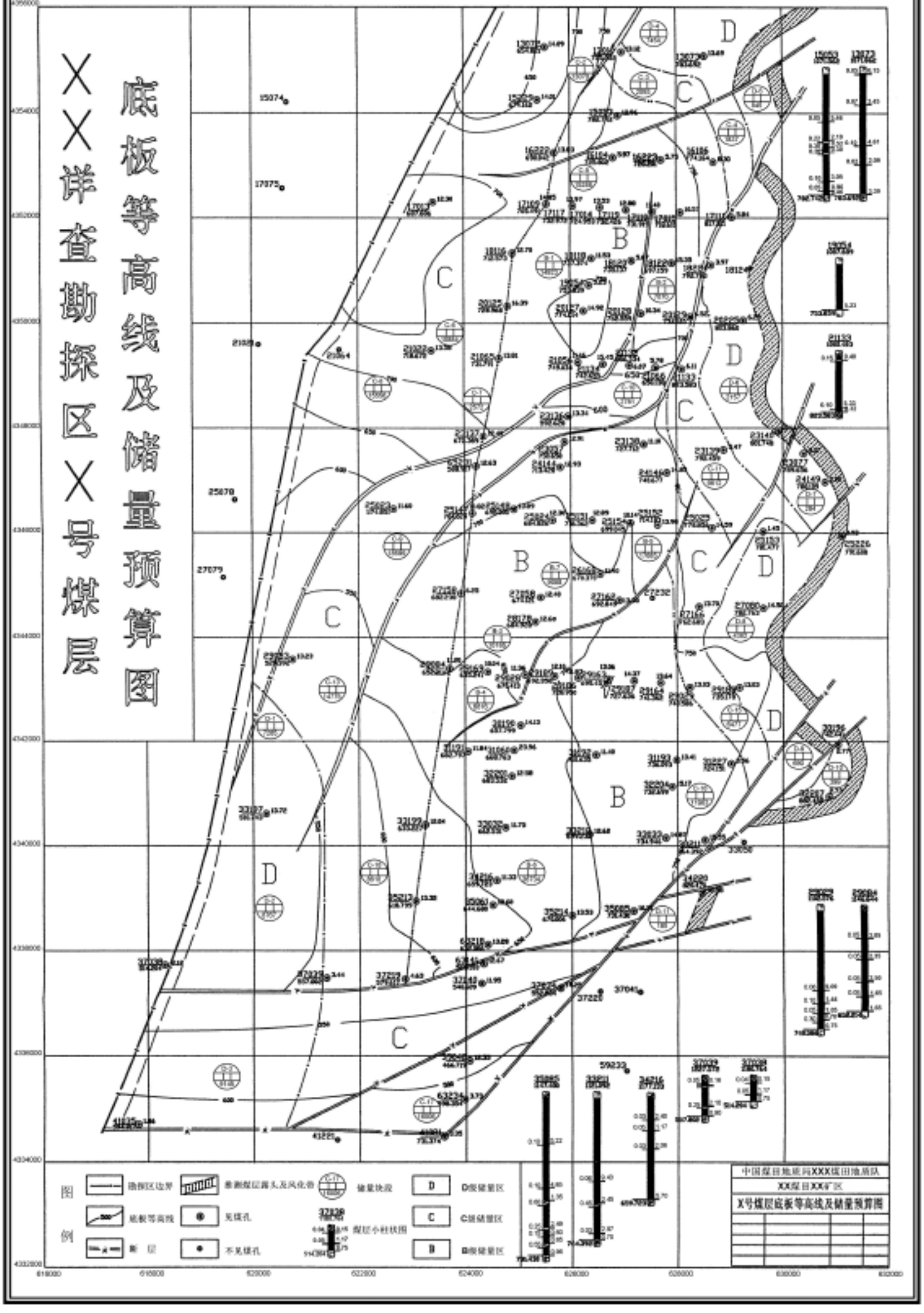


图 6-39　煤层底板等高线图并煤炭储量计算图示例

拟合度为 $C=\left[1-\frac{\sum_{i=1}^{n}(Z_i-\hat{Z}_i)^2}{\sum(Z_i-Z)^2}\right]\times 100\%$，式中：$Z=\left(\sum_{i=1}^{n}Z_i\right)/n$

3）断层助判的实现步骤。

①进行一次趋势面拟合。将矿（岩）层底板高程看作连续变化量，生成连续的底板等高线图，对底板高程值求一次趋势。一次趋势面拟合度的大小，能大致反映构造的复杂程度。若一次趋势面的拟合度低于50%，说明此区构造较复杂。

②绘制趋势异常图及剩余异常图。调入岩层露头及钻遇断层的情况，并且在图上自动标注出来。

③提取断层信息。对几种图所反映的情况进行人工分析、综合，判别勘探区内断层存在与否，进而确定断层分布及走向，在连续的底板等高线图上勾画出来。

④提取所勾画的断层线，并投影到新的目标底图上。

4）矿（岩）层底板等高线图的分块生成。根据断层的分布，分块生成等值线，经修整后输出正式图件。

5）矿产储量计算图的编辑与输出。通常，只要用户提供比例尺、图幅范围以及钻孔资料数据等参数后，编绘程序可以自动生成平面图的主体框架和钻孔标注，而用户的主要工作是进行交互式块段划分连线和部分图件整饰工作。需强调的是，矿产储量计算的方法很多。

将矿（岩）层底板等高线图与地理底图、矿产勘查块段图、矿（岩）层结构柱状图及储量参数图等叠合起来，便可形成矿产储量计算图（图6-39）。

思 考 题

1. 在传统资源勘查与开发工作中应用计算机辅助设计技术有何现实意义?
2. CADS与GIS的区别与联系是什么?
3. 资源图件机助编绘系统的逻辑结构和组成是怎样的?
4. 机助编绘系统的设计目标是什么?
5. 曲线的描述与生成有哪几种方式?
6. 图例花纹制作的基本原则是什么？图件整饰的基本内容是什么?
7. 等值线分析的实际应用意义有那些？试举例说明。
8. 试说明地质分析方法与图件编绘的结合途径，举例说明具体方法。

参 考 文 献

边馥苓等 . 1996. 地理信息系统原理和方法 . 北京: 测绘出版社

陈佳 . 1998. 信息系统开发方法教程 . 北京: 清华大学出版社

陈建平 . 矿产资源勘查与评价学科发展动态——记第五届国际矿产资源学术讨论会, 地质科技情报, 18 (3): 47 ~50

陈金水 . 1998. 微机 CAD 方法与技术 . 南京: 东南大学出版社

陈述彭 . 1995. 资源环境问题与系统信息工程, 见: 杜道生等主编 GPS、RS、GIS 的集成与应用 . 武汉: 武汉测绘科技大学出版社, 9 ~19

陈宗希 . 1996. 沿海地区的地下海水入侵模拟系统 . 武汉: 中国地质大学出版社

池顺都等 . 1992. 矿产勘查与评价实习指导书 . 北京: 地质出版社

池顺都, 吴新林 . 1998. 云南元江地区铜矿 GIS 预测时的找矿有利度和空间相关性分析 . 地球科学, 23 (1): 75 ~78

池顺都, 赵鹏大 . 1998. 应用 GIS 圈定找矿可行地段和有利地段——以云南元江地区大红山群铜矿床预测为例 . 地球科学, Vol. 23 No. 2, 125 ~128

迟忠先等 . 1994. 实用图形处理技术 . 大连: 大连理工大学出版社

戴同等 . 1989. 微计算机辅助设计 . 北京: 科学出版社

丁贵明 . 1996. 油气勘探工程及实践, 勘探家 . No. 1: 1 ~5.

董国臣等 . 1997. 计算机辅助 1∶ 5 万区调工作方法, 北京: 地质出版社

恩卡纳萨奥 . J 等 . 1988. 计算机辅助设计基本原理与系统结构 . 北京: 国防工业出版社

范玉青等 . 1996. CAD 软件设计 . 北京: 北京航空航天大学出版社

樊文有等 . 1998. GIS 制图中标准图框的机助生成 . 地球科学, 23 (4): 379 ~382

冯玉才 . 1993. 数据库系统基础 (第二版) . 武汉: 华中理工大学出版社

哈博 . J. W., 梅里亚姆 . D. F. . 1978. 电子计算机在地层分析中的应用 . 北京: 科学出版社

韩志军等 . 1999. 数据库系统数据字典的设计与实现, 微机发展, 2: 30 ~32

韩志军等 . 1999. 地质矿产信息系统开发的标准化 . 中国标准化, (11): 7 ~8

韩志军等 . 2000. 野外数据采集系统数据字典的研制, 地球科学, 24 (5): 539 ~541

何建邦 . 1995. 我国地理信息系统发展的回顾及其在第七、八个五年计划科技攻关中的技术进步, 见于杜道生等主编的《GPS、RS、GIS 的集成与应用》. 武汉: 武汉测绘科技大学出版社, 44 ~50

黄建全等 . 1998. 实用计算机地质制图 . 北京: 地质出版社

洪昌松等 . 1993. 地图制图学 . 武汉: 中国地质大学出版社

侯景儒等 . 1990. 地质统计学的理论与方法, 北京: 地质出版社

侯景儒等 . 1993. 矿床统计预测及地质统计学的理论与应用, 北京: 冶金出版社

胡安国等 . 1992. 微机在矿床储量计算与资料整理中的应用 . 北京: 地质出版社

胡朝元等 . 1990. 油气田勘探及实例分析 . 北京: 石油工业出版社

胡友元等 . 1987. 计算机地图制图 . 北京: 测绘出版社

高富 . 1993. 石油勘探信息浅论 . 大庆石油地质与开发 . 12 (3)

高复先等 . 1989. 信息工程与数据总体规划 . 北京: 人民交通出版社

高瑞祺等 . 中油集团油气勘探形势与技术发展, 石油勘探与开发 . 26 (2) 1999: 1 ~6

郭朝勇等 . 1998. AutoCAD 的定制与开发 . 北京: 人民邮电出版社

国家测绘总局测绘科学研究所译 . 1981. 地图制图自动化译文专辑 (第二集) . 北京: 测绘出版社

蒋景瞳 . 1991. 全国性资源与环境信息系统研究 . 北京: 测绘出版社

谯章明等 . 1984. 地质图绘制 . 北京: 原子能出版社

金泽兰 . 1982. 地质图编绘法 . 北京: 地质出版社

克拉里西 E, 欧文 D B, 杜鲁肯 S (魏一鸣译) . 1995. 用神经网络估算可采储量 . 国外金属矿山 . (2): 1 ~7

李德仁 . 1993. 地理信息系统导论 . 北京: 测绘出版社

李德仁 . 1995. 当前国际 GIS 的研究和应用现状 . 见杜道生等主编 GPS、RS、GIS 的集成与应用 . 武汉: 武汉测绘科技大学出版社, 24 ~43

李德仁 . 1996. 论地理信息科学的形成及其在跨世纪中的发展 . 世界科技研究与发展 . No. 5: 1 ~8

李德仁等 . 1997. GPS 辅助全自动空中三角测量 . 遥感学报, 11 (4): 306 ~310

李德仁等 . 1998. 论地球空间信息科学的形成 . 地球科学进展 . 13 (4): 319 ~326

李鸿吉等 . 1981. 电子计算机制图方法及应用 . 北京: 地质出版社

李建威 . 1995. 矿产资源定量预测方法综述 . 地质科技情报 . 14 (4): 57 ~64

李裕伟 . 1998. 空间信息技术的发展及其在地球科学中的应用 . 地学前缘 . 5 (1 ~2): 335 ~341

李晓池等 . 1995. 地理信息系统 (GIS) 在地质学中的应用 . 地学前缘, 2 (1 ~2): 22 ~25

李新中等 . 1995. 基于规则知识表示的模型单元选择专家系统的实现 . 地球科学 . 20 (2): 173 ~178

李之棠等 . 1997. 信息系统工程原理、方法与实践 . 武汉: 华中理工大学出版社

林龙震工作室 . 1999. AutoCAD R14 程序设计基础教程 . 青岛: 青岛出版社

刘发中 . 1987. 信息唯物论与科学学体论 . 武汉: 湖北人民出版社

刘刚等 . 1999. 参数化设计方法在地矿图件计算机辅助编绘中的应用 . 地质科技情报 . 18 (1): 93 ~96

刘刚等 . 2001. 资源勘查信息系统中参数化图形设计方法的应用框架研究 . 地球科学 . 26 (2), 197 ~200

刘刚等 . 2002. 资源勘查图件计算机辅助编绘系统的结构分析与开发策略研究 . 地质与勘探 [J], 38 (4): 60 ~63

刘刚等 . 2003. 计算机辅助区域地质调查野外工作系统研究进展 . 地球科学进展 . 18 (1): 77 ~84

刘刚等 . 2003. 工程勘察信息系统中多 S 集成应用研究 . 计算机工程 . 29 (3): 19 ~21

刘晶珠 . 1990. 决策支持系统导论 . 哈尔滨: 哈尔滨工业大学出版社

刘鲁主编 . 1995. 信息系统设计原理与应用 . 北京: 北京航空航天大学出版社

刘修国等 . 1998. MAPGIS 扫描矢量化及其相关技术 . 地球科学, 23 (4): 341 ~343

刘云生等 . 1992. 数据库系统概论 . 武汉: 华中理工大学出版社

卢振荣 . 1985. 计算机绘图初步 . 西安: 交通大学出版社

吕新彪 . 1996. 人工神经网络在地球科学中的应用综述 . 地质科技情报, 15 (3): 93 ~98

其和日格等 . 中国地质调查局地质调查信息化建设, 2003/6, http: //www. cgs. gov. cn/ TASK/xinxi/index. htm

其和日格等 . 1997. 加速我国 1∶ 5 万区调现代化、信息化进程, 中国地质, (10): 40 ~41

齐景顺等 . 1995. 浅谈油气地质勘探信息和信息管理 . 地质科技管理, 3

钱汉臣 . 1996. 计算机实用图形学与 CAD 技术 . 重庆: 西南师范大学出版社

任国林等译 . 1989. 工程地质编图指南 . 北京: 地质出版社

邵维忠等 . 1998. 面向对象的系统分析 . 北京: 清华大学出版社, 1 ~41

石广仁 . 1994. 含油气系统模拟方法 . 北京: 石油工业出版社

孙家广等 . 1997. 计算机图形学 . 北京: 清华大学出版社

唐荣锡 . 1994. CAD/CAM 技术 . 北京: 北京航空航天大学出版社

王宏琳 . 2001. 油气勘探开发信息化—从数据处理到数字油藏 . 北京: 石油工业出版社

王来生等 . 1993. 大比例尺地形图机助绘图算法及程序 . 北京: 测绘出版社

王通讯 . 1986. 论知识结构 . 北京: 北京出版社, p. 64

王伟元等 . 1995. ProBases 二维盆地模拟系统及其在崖 13 ~1 气田的应用, 中国海上油气地质, 9 (3): 179 ~190

王润生 . 1994. GIS 的应用与发展 . 地质信息技术, 10 (4): 1 ~22

汪厚祥等 . 1996. 交互式计算机图形学 . 武汉: 华中理工大学出版社

汪新庆等 . 1998a. 地质矿产点源数据库系统的模型库及其分类体系 . 地球科学, 23 (2): 199 ~204

汪新庆等 . 1998b. 岩土工程勘察点源程序信息系统的开发 . 岩土工程青年专家学术论坛文集, 73 ~78

汪新庆等 . 1999. 地质矿产术语分类代码在地矿点源信息系统中的应用 . 地球科学, 24 (5): 529 ~532

汪新庆等 . 2001. 数据字典在 ArcView GIS 中的实现与应用 . 计算机与信息技术, 4

王世称等 . 1989. 综合信息解释原理 . 长春, 吉林大学出版社

王世称等 . 1990. 矿产资源评价方法 . 长春, 吉林科学技术出版社

王世称等 . 1995. 金矿综合信息成矿系列预测理论 . 黄金地质, (1): 1 ~7
汪昭等 .1987. 信息处理概论 (第一、二卷) . 北京: 中国人民大学出版社
王益清等 . 1998. 油气勘探与石油地质综合研究—理论方法和程序 . 武汉: 中国地质大学出版社
吴冲龙等 . 1987. 煤炭资源点源信息系统研制的设想 . 地质科技情报, 11 (2): 160 ~164
吴冲龙等 . 1987. PC1500 袖珍计算机地质常用程序 . 武汉: 武汉地质学院出版社
吴冲龙等 . 1992. 聚煤盆地地质信息计算机处理的途径与方法 . 北京: 地质出版社
吴冲龙等 . 1995. 论地质矿产点源信息系统 . 载于中国地质学会数学地质专业委员会主编的《中国数学地质 (6)》, 120 ~125
吴冲龙等 . 1996. 地矿点源信息系统设计原理及应用 . 武汉: 中国地质大学出版社
吴冲龙 .1998. 计算机技术与地矿工作信息化 . 地学前缘, Vol. 5: (1 - 2)
吴冲龙 .1998. 地矿点源信息系统的开发与应用 . 地球科学, 23 (2): 193 ~198
吴冲龙、刘刚 .2002. 中国“数字国土”工程的方法论研究 . 地球科学, 27 (5): 605 ~609
武晓丽等 . 1997. 计算机绘图实用教程 . 北京: 电子工业出版社
吴信才等 . 2000. 地理信息系统原理、方法及应用 . 武汉: 中国地质大学信息工程学院
徐冠华、孙枢等 . 1999. 迎接“数字地球”的挑战 . 遥感学报, 3 (2): 85 ~89
许友志等 . 1994. 定量煤田地质学 . 徐州: 中国矿业大学出版社
杨昂岳 .1995. 微机实用绘图方法与技巧 . 长沙: 国防科技大学出版社
於崇文等 . 1998. 成矿作用动力学 . 北京: 地质出版社
俞全宏 .1994. 数字制图与地学数据可视化的新进展 . 地质信息技术, 10 (4): 31 ~36
袁艳斌等 . 2001. 面向野外地质填图的空间实体对象表达 . 地球科学, 26 (2), 192 ~196
岳永霞 .1998. 勘探数据库的建立与发展 . 石油地球物理勘探 .8, 33 (1)
叶飞跃等 . 1998. 石油勘探开发工程数据库模型设计 . 石油大学学报 (自然科学版), 22 (1)
燕汉业等 . 1998. 油气勘探科研数据库设计与建立 . 大庆石油地质与开发, 17 (5)
张大顺等 . 1994. 地理信息系统技术及其在煤矿水害预测中应用 . 徐州: 中国矿业大学出版社
张洪涛, 2001. 服务国家目标, 体现科技创新, 中国地质, 28 (1): 4 ~8
张洪涛、庄育勋等, 2002. 区调提速的紧迫性和可能性, 中国地质, (1): 1 ~5
张夏林等 . 2001. 计算机辅助地质填图属性数据采集子系统的动态数据模型 . 地球科学, 26 (2), 201 ~204
张夏林等 . 2001. 复杂野外地质属性数据采集子系统的动态数据模型 . 地质科技情报, 20 (3), 93 ~96
赵鹏大等 . 1983. 矿床统计预测 . 北京: 地质出版社, 10 ~12
赵鹏大等 . 1991. 矿产勘查与评价 . 北京: 地质出版社
赵鹏大等 . 1992. 地质学的定量化问题 . 地球科学, 17 (增刊): 51 ~56
赵鹏大等 . 1995. 中国地质异常 . 地球科学, 20 (2): 117 ~127
赵锐 . 1990. 区域开发信息系统研究 . 北京: 测绘出版社
郑贵州 .1998. 地理信息系统 (GIS) 在地质学中的应用 . 地球科学, 23 (4)
钟义信 .1988. 信息的科学 . 北京: 光明日报出版社, 39 ~49
中国石油天然气总公司勘探局 .1999. 油气资源评价技术 . 北京: 石油工业出版社
周顺平等 . 1998. MAPGIS 二次开发库的设计与实现简介 . 地球科学, Vol. 23 (4): 337 ~340
朱章森等 . 1991. 求异理论与无模型预测 . 中国数学地质 (3), 北京: 地质出版社
朱章森 .1992. 矿产资源无模型预测刍议 . 物探化探计算机技术, (1): 60 ~62
Andrew P. Sage, 1991. Decision Support Systems Engineering, New York: John Wiey
Auerbach, S., 1990. Schaeben, H. Surface Representation Reproducing Given Digitized Contour Lines. Mathmatical Geology, 22 (6)
Axon Allan G.; 1996. Crowell D. L.: Implementation of Geographic-information-system Technology for Use in Coal geology investigations at the Ohio Department of Natural Resources, Division of Geological Survey. AAPG Bulletin. 80; 9, PP 1519.
Berg Richard C.; Soller David R.; 1995. Kempton John P; Weibel C. Pius: Three-dimensional geologic mapping of Quaternary deposits in east-central Illinois. Abstracts with Programs-Geological Society of America. Geological Society of America (GSA)

. Boulder, CO, United States. 27 (3), P. 39

Brodaric B. 1997. Field data capture and manipulation using GSC Fieldlog v3. 0. Proceedings of a workshop on digital mapping techniques: methods for geologic map data capture, management, and publication; by Soller, D R; United States, Geological Survey, 77 ~81

Bruce R. R, 1996. Three ~dimensional selection and its potential application in graphics and image processing of geoscience data. Computers & Geosciences. 22 (10): 1187 ~1189

Burrough P A. 1996. Principles of Geological information System for Land Resources Assessment. Clarcndon press

Cherry M. 1993 - 94. Focus on NATMAP: Canada's National Geoscience Mapping Program. Geological Survey of Canada: New direction for geoscience, 12 ~13

Falcidieno B., Pienovi C.. 1990. Natural surface approximation by constrained stochastic interpolation. Computer aided design, 11 (3)

FGDC. 1997, Content standards for digital geo-spatial metadata. Federal Geographic Data Committee, 6

Fyfe W S. 1997. Earth Sciences for the 21st Century: Evolve or die. GFF, 119 part 2: 85 ~90

Harbaugh J. W., Merriam D. F.. 1968. Computer Application in Stratigraphic Analysis. New York: John Wiley and Sons

Harbaugh J. W., Bonham-Carter G., 1980. Computer simulation in geology. New York, Wiley-Interscience

Hughes J. D. 1996.: Optimizing resource analysis of stratified sequences with 3D modeling, GIS, remote sensing and 3D visualization. Kennard-John-M (editor): Abstracts-Geological Society of Australia. 41; Geological Society of Australia. Sydney, N. S. W., Australia. p. 203

Inmon W H. 1992. Building the Data Warehouse [M]. New York: John Wiley & Scons Inc

James M., 1977. Computer Database Organizations. Second Edition, Prentice-Hall

Johnson B. D., Bradbury. R., 1991. The National Resource Information Centre; its role in the identification, access and integration of resource information in support of government decision making processes and in geoscience research. in: Abstracts-Geological Society of Australia. Sydney, N. S. W., Australia. 30: 109 ~110

Kurt Hampe, Jim Boyce. 1994. AutoCAD 应用开发工具大全. 北京: 清华大学出版社

Leonard J. E., 1989. Basin modeling economic tool for locating hydrocarbons, Amer. oil gas repoter. 32 (11): 11 ~15

List Franz K.; Squyres Coy H. 1996.: Geologic mapping from satellite data; contribution of remote sensing and GIS technology to hydrocarbon and mineral exploration. Proceedings of the Thematic Conference on Geologic Remote Sensing. 11; Pages I. 309 ~ I. 320

Franz L. K., Coy S. H; 1996. Geologic mapping from satellite data; contribution of remote sensing and GIS technology to hydrocarbon and mineral exploration. Proceedings of the Thematic Conference on Geologic Remote Sensing. Ann Arbor, MI, United States. 11: I. 309 ~320

Martin J. 1985. Haigh. An Introduction to Computer Aided Design and Manufacture. Blackwell scientific publications

Michael. K. B., 刘子明、肖唐金 编译. 1994. 信息与信息系统. 广州: 中山大学出版社

Miller B. M., 1991. Integration of geographic information systems and expert systems technology for resource management. In: Proceedings; Resource Technology 90; American Society for Photogrammetry and Remote Sensing. Bethesda, Second international symposium on Advanced technology in natural resource management. Bethesda, MD, United States. 727 ~736

Moore R. F., 1991. GIS in the minerals and oil exploration industry. Abstracts-Geological Society of Australia. Sydney, N. S. W., Australia. 30: 81 ~82

Morgan K. 1990. The assessment of hydrocarbon resource in Shannon basin, U S. 1990 Conference on Artificial Intelligence in Petroleum Exploration & Production. 50 ~57

Mark R., 1992. The changing face of geological maps and mapping. Geological Society of New Zealand Miscellaneous Publication. Christchurch, New Zealand. 63A: 130

Partt P. J. 1997. and Adamski J. J, The concepts of database Management, Second Edition, Course Technology, a Division of International Thomson Publishing, Cambridge, U. S. A.

Petrie G. 1987. and Kennie . J. M.. Terrian modelling in surveying and civil engineering. Computer aided design, Vol. 19 (4)

Power W. L.; Lamb P.; 1995. Horowitz F. G.: From databases to visualization; data transfer standards and data structures for

3D geological modeling. Publication Series-Australasian Institute of Mining and Metallurgy. 4 /95; Parkville, Victoria, Australia. p. 65 ～70

Rattenbury Mark: 1992. The changing face of geological maps and mapping. Geological Society of New Zealand Miscellaneous Publication. 63A; p. 130

Reddy R. K. T, Bonham C. G. F., Wright D. F.; GIS for mapping and assessing mineral resource potential; 1990. an application to basemetal resource evaluation in Manitoba. In: GIS for the 1990s, national conference; proceedings—SIG, la prochaine decennie, conference nationale; actes. Canada. 384 ～400

Rhind D., 1987. Recent developments in geografical informaition systems in UK: INT. J. Geographical information systems. 11 (3): 229 ～241

Richard B. C., David S. R., 1995. John K. P. et al.: Three-dimensional geologic mapping of Quaternary deposits in east-central Illinois. Abstracts with Programs-Geological Society of America. Geological Society of America (GSA). Boulder, CO, United States. 27 (3): p. 39

Rogers N, Brodaric B. 1996. Spatially linked relational database management of petrology and geochemistry using Fieldlog v3. 0: a worked example from the Bathurst mining camp, New Brunswick. Current Research/Recherches en cours; by Geological Survey of Canada; Geological Survey of Canada, Current Research, no. 1996-E, 255 ～260

Russell H A J, Logan C, Brennand T A, 1996. et al. Regional geoscience database for the Oak Ridges Moraine project (southern Ontario). Current Recherches en cours; by Geological Survey of Canada; Geological Survey of Canada, Current Research, no. 1996-E, 191 ～200

Simon. H, 1982. **管理决策新科学. 北京: 中国社会科学出版社**

Struik L C, Atrens A, 1994. Haynes A. Computer geological field notebook system and its use with GIS. Decision making with GIS, the fourth dimension: symposium proceedings, Volume 2; by Polaris Conferences Inc.; FRDA Report, 529 ～537

Su C B, Black J. 1996. Practical, Scalable, Modern 3 ～D Seismic Processing. The beading Edge, 15 (12)

Talat Q. M., Khalid K. A; 1996. Computer aided graphics simulation modelling using seismogeologic approach in sequence stratigraphy of Early Cretaceous Punjab Platform, central Indus Basin, Pakistan. AAPG Bulletin. 80 (8): 1326 ～1327

Verma Ashok, Northcott W J, 1998. Mitchell J Kent. Object ～oriented database management systems (OODBMS) for watershed hydrologic modeling. American Water Resources Association Technical Publication Series TPS. TPS; 98 ～3, p. 201

William J., 1997. Mitchell, Malcolm McCullough. **数字设计媒体 (第二版). 北京: 清华大学出版社**

Williams F. N., 1996. Three-dimensional subsurface characterization using geoscientific information systems in the Springfield-Harrison 1∶250, 000 map quadrangles, Missouri and Arkansas. Golden, CO, United States. 200

B. A. **Чфвяков** 1996. **等值线模拟的理论、应用与展望. 地图**, No. 3