

文章编号: 1007-4619 (2001) 01-0041-05

大型遥感图像处理系统中集成数据库设计及应用

李 军^{1,3}, 刘高焕¹, 迟耀斌², 朱重光³

(1. 中国科学院 地理研究所资源与环境信息系统国家重点实验室, 北京 100101;

2. 北京遥感信息研究所, 北京 100032; 3. 中国科学院 遥感应用研究所, 北京 100101)

摘 要: 大型遥感图像处理应用系统中, 往往需要实时获取各种背景或专题数据, 该过程即是数据动态集成过程。集成数据库是建立在各种专题数据库基础上的数据集成使用框架体系, 该文描述了集成数据库的结构及各类子库的组成, 根据项目的特殊需求提出了虚拟数据库概念, 并结合实例说明了集成数据库以元数据为链条的使用机制与方法。

关键词: 图像处理; 集成数据库; 元数据; 虚拟数据库

中图分类号: TP79/P208 **文献标识码:** A

1 引 言

相对于传统的数据获取方式, 时效性是遥感数据的最大优势。如何从遥感影像中快速获取需要的数据是图像处理的核心, 从大量遥感影像中快速获取特定的数据需要大量的背景数据及相关的知识与模型。大型遥感图像处理应用是一种在计算机软、硬件支持下, 动态地获取、管理和分析处理区域内地理环境信息及各类专题信息、服务于专向目的(如军事应用^[1]、环境变化监测等)的应用系统, 其主要功能是快速解译大量图像、解译产品制作等。

大型遥感图像处理项目需求的数据类型繁多, 数据量大, 数据尺度多样, 涉及的空间实体类型数据巨大, 软件系统多, 并有精度高、时效性强的需求。目前已出现的各种数据库技术及系统所处理的问题有一定的局限性^[2,3], 即对某些数据类型的数据库较适宜, 能高效管理包括空间数据的多类型数据的软件系统较少。大型遥感图像处理应用要求在图像实时处理过程中要用到各类信息以辅助确定地物, 系统成败的关键之一在于各种类型数据的组织和高效使用, 因而需要建立面向遥感图像处理的集成数据库以实现这些目的。

根据大型图像处理应用数据的特点和系统功能需求, 数据库设计为面向地物的多级管理模式。集

成数据库系统的主要功能是:(1)为影像判读提供相应的参考背景数据;(2)便于各种数据进行综合分析;(3)为影像判读、空间分析、产品制作提供模型;(4)生成各种产品。

2 集成数据库的设计

2.1 集成数据库设计基础

集成数据库是建立在各种专题数据库基础上, 能综合管理和操作各类数据的数据库框架体系, 其难点是多元数据的集成使用。它与数据仓库有一定区别, 数据仓库是支持管理决策过程的、面向主题的、集成的、随时间而变的、持久的数据集^[4], 是数据库的数据库, 而集成数据库是建立各种独立数据库管理系统之上的数据管理环境, 是各独立数据库与应用系统之间的中介。

集成数据库设计的原则是:(1)集成数据库的中心目的是服务于影像的快速解译和解译产品的表达, 数据库设计以地物解译为主线。(2)系统数据来源不同, 数据格式多样, 数据质量差异也大, 系统需要在统一的管理体系下实现快速调用、匹配、分析处理、更新, 并存贮及形成相应的元数据, 支持各种类型的检索查询, 各类数据灵活的集成机制保持数据的兼容性。(3)跨平台数据管理, 由于数据自身特点、软件功能限制及系统对各种数据的需求, 数据的

收稿日期: 1998-11-05; 修订日期: 2000-02-22

作者简介: 李 军(1968—), 男, 现为中国科学院遥感应用研究所博士后, 1998年在中国科学院地理研究所获得地图学与地理信息系统专业博士学位, 主要从事地球空间数据集成基础研究。

底层管理只能在不同的软件平台上进行,上层由集成数据库统一管理。

2.2 集成数据库结构

集成数据库有 3 部分组成(图 1),基础部分是各种独立的专题数据库,如地物图片库、地理信息库等;第二部分是集成数据库系统自身;第三部分元数据库系统。集成数据库在遥感图像处理应用系统中充

当着桥梁的作用,其作用为向下传递信息及向上传递数据。前者指集成数据库把用户需求分解成可操作的步骤,并将它们(特别是数据需求)传递给各个数据库,根据元数据库系统寻找需求数据的具体位置;后者指集成数据库负责从各个数据库返回的数据打包返回给用户系统,并根据数据的类型启动相应的数据操作环境。

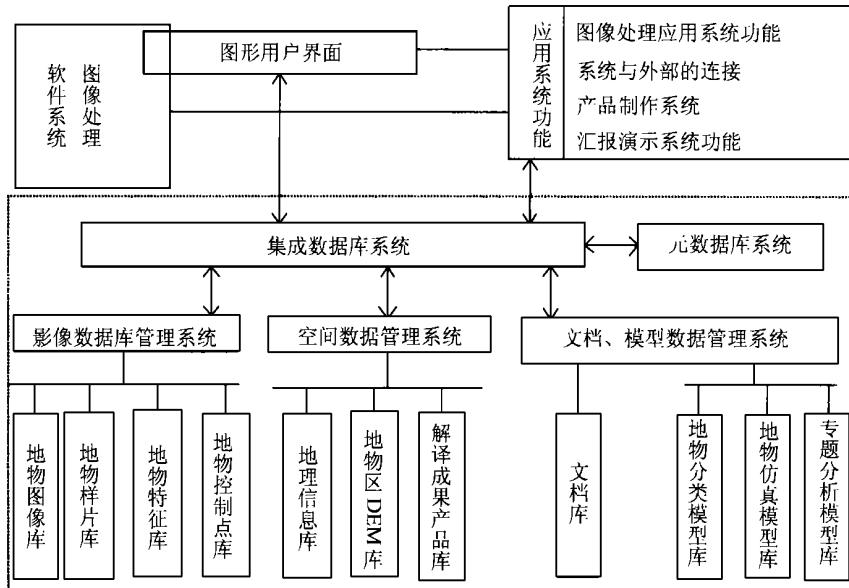


图 1 集成数据库系统结构及运行机制

Fig. 1 Structure of integrated database for large remote sensing applications

2.3 数据库组成

由于大型遥感图像处理项目本身的复杂性,导致集成数据库的复杂化,具体表现在数据库涉及的

内容广泛、数据的类型复杂多样、数据量庞大、功能多。从数据管理使用环境出发,集成数据库中把各类数据分别组织到不同的数据库中,各主要子数据库简述如表 1。

表 1 集成数据库的各子数据库系统

Table 1 Sub-database system of Integrated database in large remote sensing applications

数据库	数据库特征及实现过程描述
影像数据库	存贮不同尺度、精度的遥感影像。将地物图像文件与地物属性文件分别置于图像处理软件和文件管理系统(如 ORACLE)之下,通过元数据库系统,将地物图像及其属性相关联。其功能包括:多条件组合条件查询检索;报表生成、数据的输出、与 GIS 的接口、与产品输出模块相关联等。
地理信息数据库	地理信息数据库是遥感影像解译的辅助数据资料。该数据库的内容多数为矢量数据,其空间特征可以由通用的 GIS 商业软件系统(如 ARC/INFO 等)管理,属性数据由 GIS 软件或商用数据库软件(如 ORACLE)管理。地理信息数据库的关键技术在于:(1)如何实现数据高效查询检索、数据更新及数据在不同的软件平台和不同存贮物理地址之间的调用;(2)数据编码与系统中其它数据编码的兼容性;数据的快速输入和更新。
地物系统数据库	地物数据库系统包括:地物特征库、地物样片判读库、地物控制点库。地物特征库记录典型空间实体的波谱特征、空间几何特征及纹理特征,数据包括典型空间实体的形状、大小、位置等图文信息,主要为地物识别、波谱数据分析提供依据 ^[5] 。样片数据库为影像判读人员对地物性质进行判定提供参考影像。地物控制点库目的是卫星影像接收后将自动和控制点库中的控制点相匹配,然后进行几何精校正。
模型库系统	模型库作用主要是为影像判读、空间数据分析、解译结果产品制作提供各种分析与预测模型。内容包括:地物分类模型、典型地物选择、地物光谱特性模型、专用分析评价模型、地物地区的动态实时模拟。
产品库系统	产品数据库系统包括:数据库管理系统、数据处理功能块、成果输出系统。产品库的功能有:数据库管理功能、数据库查询检索及结果输出、数据打包。

2.4 集成数据库实现

集成数据库是各独立数据库的上层集成, 其实现采用自下而上的方法^[9], 集成数据库的界面嵌在应用系统界面中, 其功能用图像处理软件宏语言与高级语言(如 Visual Basic Visual C++ 等)编程实现。集成数据库功能的实现分为以下过程:

(1) 应用系统数据或功能请求, 数据库系统接受由应用系统界面发出的请求, 并将它们分解为数据、模型、功能等不同类型需求; (2) 由元数据库系统确定数据或模型文件所在的数据库, 并将需要的数据、模型通知具体的数据库, 对于功能需求集成数据库负责调用相应功能完成数据操作(如数据提取、选择); (3) 集成数据库将各子库中反馈的满足需求的数据打包, 并返回给应用系统。

2.5 数据命名方法

数据命名的目的是服务于数据集的检索与其他管理。对集成数据库下各数据库中数据文件的名称采用统一的命名准则。为便于以时间为条件的数据检索^[7], 影像数据命名中考虑了时间因素, 不同类型数据命名侧重不一致, 具体如下:

(1) 遥感影像数据。名称采用{影像来源卫星类型编号+覆盖空间区域+时间+处理状况}方案, 卫星根据其分辨率及其它综合参数分为几种类型, 每一类给一编号。覆盖区域采用内部密码, 可对每一种类型的卫星覆盖情况把地球表面分为若干区域, 每一区域用特定的符号表示。时间为影像数据的获取时间, 可以采用年(三位)、月(两位)、日(两位), 具体时间作为数据的元数据存放在特定文件中。处理

状况指数据处理的进程, 可根据处理状况分为几种类型: 如原始影像、初判、预判、详判、会判等类型。

(2) 矢量空间数据。名称采用{数据内容名称缩写+比例尺类型+覆盖区域编号+时间}方案。数据内容名称使用内容对应英文单词(或词的组合)的四字母缩写, 较长内容名称采用其关键词的缩写。精度类型, 根据系统中用到的各种类型数据精度情况, 把精度分为几种类型, 每一类型给一标识码。

(3) 文本数据。文本文档命名采用{数据名称+文件性质后缀}方式。数据名称采用上述的通用名称, 关于系统、应用程序的文档采用特殊表示方式。

3 元数据在系统中的应用

元数据是对数据的描述性数据, 它包含数据从空间到属性、从外部到内容的多种信息, 同时也包含数据集在数据库系统中使用存贮、数据获取等信息, 是数据检索查询的依据和桥梁^[9]。元数据的来源有两个途径, 静态描述性信息由数据生产者或仪器记录填写; 动态性信息是在数据库操作、数据处理、数据更新过程中的变化记录。考虑到集成数据库的复杂性, 把元数据组织到不同的层次中。

3.1 数据库元数据

大型遥感图像处理工程中数据库层次的元数据是指对集成数据库中各个子数据库的描述性信息, 目的是获取各子数据库的有关信息, 以服务于数据库的检索查询及其它数据库管理功能。数据库层次的元数据内容如表 2。

表 2 数据库层次元数据内容

Table 2 Contents of database level Metadata in large remote sensing application system

元数据项	内容描述	元数据类型
数据库的名称	用户或数据库管理系统定义的数据库名称	文本格式
存放数据类型	关键字描述, 地理数据、影像数据、样片数据、地物特征、地物控制点、模型、地物文档、产品成果	文本描述
数据库位置	数据库存放的逻辑或实际物理地址	字符串表达
数据源描述	数据库内容的主要来源	文本格式
数据量	数据库中数据集的数量	整数表达
元数据信息	数据库元数据形成及更新的时间	时间字符串
数据库管理	说明数据库管理的部门	文本描述

3.2 数据集元数据

关于空间数据集的元数据应包括的内容,国际上已有一系列的规范和标准^[9],其包含的内容有一定的普遍意义。但对于特定的应用系统而言,数据集层次的元数据所要求的内容又有一定的特殊性,一般比通用的数据集元数据内容更为丰富,它与数据使用的硬件和软件环境条件有一定的联系。大型遥感图像处理系统中数据集元数据的主要功能是服务于数据检索、数据集成、数据变动记录等,因而数据集的元数据内容包括:数据标识信息、数据质量信息、空间数据生产者描述、数据空间参考消息、地理实体及属性信息、数据使用信息、元数据参考信息、其它信息。

数据集元数据的管理以数据库为单位,即每一个数据库有一个元数据文件,每一个数据集元数据内容是该元数据文件中的一个记录,该表格型元数据文件存放为关系数据库数据文件,对它的操作在

关系数据库系统中用 SQL 实现。对数据集内容操作引起数据集元数据的变化,首先由数据处理应用系统将数据变化传递给元数据管理系统,然后由元数据信息系统调用相应的元数据文件并对之进行相应的操作,以实现元数据的更新。

4 虚拟数据库

简单地讲虚拟数据库是一种不包含实际内容的逻辑数据库,它用以解决用户的特殊需求:从形式上需要这样的数据库,而其内容又是其它数据库的内容。在设计上虚拟数据库的实际内容是数据具体存放的路径,虚拟数据库自身提供一种数据管理和使用机制,即当用户对虚拟数据库操作时,虚拟数据库能够将用户需求经过后台处理反馈给用户。

4.1 虚拟数据库结构

集成数据库中虚拟数据库的结构如图 2,系统

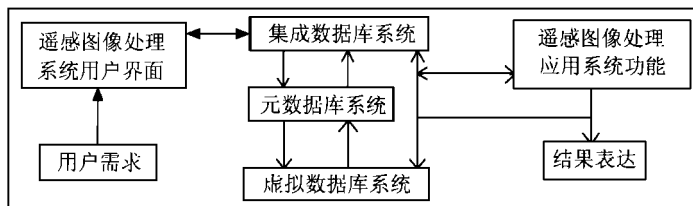


图 2 大型遥感图像处理项目中虚拟数据库机制

Fig. 2 Virtual database design in large remote sensing applications

用户通过图像处理系统用户界面将对虚拟数据库的需求传递给集成数据库管理系统,集成数据库将需求分解后传递给虚拟数据库,虚拟数据库将用户需求分解为数据需求和功能需求,然后用元数据系统获取需求数据和功能的位置信息,并将它反馈给集成数据库,由集成数据库将虚拟数据库反馈的信息分解后再由其它各个独立的数据库提供数据和功能模型,最后由集成数据库负责调用相应的功能以用户指定的方式把结果表达出来。

4.2 虚拟数据库应用

在大型遥感图像处理系统中影像解译成果产品数据库即可设计为虚拟数据库,按照数据统一管理及其分布式存贮结构的需要,影像解译的结果应根据结果的数据类型分别存放到不同的数据库中。但由于影像解译结果数据量巨大,不可能把解译结果建立一个独立的数据库,而是建立一个逻辑虚拟成果产品数据库,其中各数据的实际存放地址有数据

库管理系统统一负责,其有关的信息写入相应的元数据文件中。

产品库的查询机制同其它数据库的查询机制一样(图 3),查询指令传给 DBMS,由 DBMS 索引相应的元数据,然后将反馈信息经由 DBMS 到不同的系统功能块中提取对应的数据。如果其中需要对原有的数据进行进一步的加工处理,则有 DBMS 传达指令到相应的系统功能块进行一定的操作,然后将结果返回给数据用户。

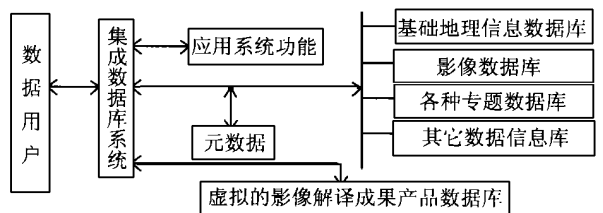


图 3 虚拟影像解译成果产品数据库的检索机制

Fig. 3 Querying pattern in virtual product database in large remote sensing applications

5 结 论

大型遥感图像处理系统中数据库的设计与建设涉及有相当多的问题, 本文仅从整体上对其中涉及的几个主要问题进行了论述, 作者认为元数据库系统是图像处理系统的核心纽带, 虚拟数据库体现了大型数据库的灵活性。

参 考 文 献 (References)

- [1] Jiang Nan, Hua Yixin. Application of remote sensing technique on military geographical information system [J]. *Remote Sensing Information*. 1997, (2) 25—27. [江南, 华一新. 遥感技术在军事地理信息系统中的应用研究[J]. 遥感信息, 1997, (2) 25—27.]
- [2] Li Zhoyuan. Latest progress in database technology [M]. Beijing: Tsinghua University Press, 1997. [李昭原. 数据库技术新进展[M]. 北京: 清华大学出版社, 1997.]
- [3] Bhavani M. Thuraisingham, Data Management System. CRC Press, New York, 1997, 1—14.
- [4] Hajinder S. Gill, Prakash C. Rao. The Official Client/Server Computing Guide to Data Warehousing [M]. 1996. [Hajinder S. Gill 著. 王促谋等译. 数据仓库——客户/服务器计算指南[M]. 北京: 清华大学出版社, 1997.]
- [5] Mai Airai. Remote sensing information models [M]. Beijing: Peking University Press, 1997. [马蕊乃. 遥感信息模型[M]. 北京: 北京大学出版社, 1997.]
- [6] Zhang Haifan. Guide to software project [M]. Beijing: Tsinghua University Press, 1998. [张海藩. 软件工程导论[M]. 北京: 清华大学出版社, 1998.]
- [7] Elisa Bertino, Beng Chin Ooi, Ron Sacks-Davis, et al. Indexing techniques for advanced database systems [M]. Kluwer Academic Publishers, 1997.
- [8] Li Jun, Chen Chongcheng. Studies on Metadata of Geo-spatial data [J]. *Geography Research*, 1997, 16(1): 31—38. [李军, 陈崇成. 地球空间数据的元数据研究[J]. 地理研究, 1997, 16(1): 31—38.]
- [9] Li Jun, Zhou Chenghu. Overview on Metadata standards of Geo-spatial data [J]. *Progress in Geography*, 17(4): 55—63. [李军, 周成虎. 地球空间元数据标准初探[J]. 地理科学进展, 17(4): 55—63.]

Design and Application of Integrated Database for a Large Remote Sensing Processing System

LI Jun^{1,3}, LIU Gao-huan¹, CHI Yao-bin², ZHU Chong-guang³

(1. State Key Laboratory of Resources & Environment Information System, Institute of Geography, CAS, Beijing 100101, China;

2. Beijing Remote Sensing Research Institute, Beijing 100032, China;

3. Institute of Remote Sensing Applications,
Chinese Academy of Sciences, Beijing 100101, China)

Abstract: It is necessary to provide any essential background data and thematic data timely in image processing and applications. In fact, it is very difficult to integrate different kinds of data into one database that is managed by commercial GIS or image processing software such as ARC/INFO or ERDAS. In this paper, the author describes an integrated database management system which is a framework based on different kinds of database, such as image database, vector spatial database, spatial entity spectrum characteristics database, spatial entity image sample database, control point (tics) database, documents database, models database, and product database. The querying and retrieving system, which are basic functions of integrated database management system, depend on metadata being divided into three parts: database metadata, dataset metadata and attribute field metadata. Finally, the author introduces the concept of virtual database that is a logical database based on other practical databases, and describes its structure and application in product making system for a large remote sensing application in detail.

Key words: image processing; integrated database; metadata; virtual database