

对建立海洋基础数据管理系统设计方案的探讨

栾振东, 范奉鑫, 阎 军

(中国科学院海洋研究所, 山东 青岛 266071)

摘要: 论述了从浩瀚的资料中快速提取信息而建立海洋基础数据管理系统的必要性, 并根据地理信息系统 (GIS) 技术及空间数据仓库技术提出了建立海洋基础数据管理系统的基本思路。

关键词: 地理信息系统 (GIS); 数字仓库; 数据库结构设计

中图分类号: P315.69 文献标识码: A 文章编号: 1000-3096(2004)10-0006-04

1 建立海洋基础数据管理系统的必要性

近年来, 随着信息技术的飞速发展, 数字地球的观念深入人心, 各发达国家纷纷竞相投入巨大的人力、财力进行开发和研究。数字地球基础平台是多维、多种分辨率, 能嵌入巨大数量的地理数据、属性数据和多媒体数据等。构成数字地球的数据的基石是具有空间信息属性的空间数据库, 数字海洋则是数字地球的重要组成部分。建国以来, 各相关单位在中国海域和大洋地区相继开展了多次大规模的海洋调查研究工作, 积累了丰富的资料。随着以信息化技术为代表的高新技术的应用, 例如遥感、多波束探测等, 使得海洋研究数据呈指数增长。海洋学是一门综合性极强的学科, 涉及众多的领域, 包括沉积学、矿物学、地球化学与海水化学、天体物理、地球物理学、古海洋、海洋物理、遥感、生物学等, 而以上每一领域又包括了众多的研究方向和研究手段, 在海洋科学研究领域中, 从资料采集、数据处理到研究方式都发生了巨大的变化, 如何从浩瀚的资料中快速提取所需的信息用于特定研究目的, 是摆在科研人员面前的重要问题。地理信息系统 (GIS) 技术作为一种利用计算机对地球空间数据进行采集、存储、检索、查询、分析、建模和表示的有效工具, 不仅能够管理属性数据及空间数据, 而且能够进行各种空间分析。自 20 世纪 60 年代中期, 加拿大建立起世界上第一个地理信息系统 (CGIS, 主要用于土地资源管理) 以来, GIS 在应用和技术方面都有了飞速的发展, 在测绘、制图、环境及资源管理等传统领域有了更深入、更广泛的应用。随着海洋科学的不断发展, 信息的大量增加, GIS 技术已成为海

洋科学研究中不可缺少的组成部分, 在海洋科学研究与管理的各个领域发挥愈来愈重要的作用^[1]。空间数据仓库技术是近几年才兴起的一门新技术。空间数据仓库就是面向主题的、集成的、稳定的、不同时间的空间数据集合, 用以支持空间信息管理中的决策制定过程。当代海洋学问题的提出和解决常常要求多学科多部门之间的协作。因此, 空间信息的共享与互操作、分析与综合对于海洋学问题的综合研究便显得日益重要。

2 海洋基础数据管理系统的内容

海洋信息有明显的空间信息特点和多用途的服务对象。利用数据库和地理信息系统 (GIS) 技术建立空间数据库是解决这一问题的途径之一。健全的海洋基础信息系统可以迅速准确地对某一海域的资源开发、利用管理, 环境监测与保护, 油气生产宏观调控及港口建设等提供基础数据, 为各种方案的分析比较和重大战略决策提供信息服务。其信息具有类型多、数据量大、来源多、获取代价大等特点。

目前各研究、生产和其它相关机构基于不同的目的, 已经和正在建立各种各样的海洋空间数据库, 这为历史资料的查询和资料存储、更新提供了很大的方

收稿日期: 2003-03-13; 修回日期: 2003-05-20

作者简介: 栾振东 (1976-), 男, 山东青岛人, 在职硕士, 助理研究员, 参加了中国科学院重大科学问题研究海洋特殊环境及其特性研究课题。电话: 0532-2898523, E-mail: luan@ms.qdio.ac.cn

便。建设与海洋信息相关的主要空间数据库应包括:海洋及海岸带基础数据库、环境物理参数数据库、海洋遥感数据库、海洋生物数据库等。

由于现在的 GIS 技术是一种面向空间数据管理及处理的技术,因此这就要求资源与环境信息系统设计中首先建立资源环境信息向 GIS 空间概念的影射关系,其空间特征的正确表达是最重要的^[2]。海洋基础数据管理系统对数据的管理是基于图层的,并通过关键字段将空间数据和属性数据连接起来^[3]。为了能对数据合理地分层,首先必须进行合理地分类,设立主题、亚层主题等。其次,应对每一层的空间符号(点、线、面)进行定义,制定出合理分类、分层体系。属性数据通常用来反映与空间实体对应的属性,一般是通过分类、量算、命名、统计等方法得到的。海洋基础信息系统主要内容包括栅格数据和矢量数据的输入、存储、地理数据的操作和分析(矢量数据空间分析和栅格数据空间分析)、将用户查询的结果或是数据分析的结果以合适的形式输出^[4]。

海洋基础信息系统栅格数据和矢量数据的质量控制方法及标准:各专业子库都包括一个质量标准表,含质量标志码、质量级别、数据精度、说明等信息,数据库中每条数据记录都含有质量标志描述,具体标准将根据各专业历史及现状,用户需求而定。质量控制过程采用分阶段(粗筛与精筛)自动与人机交互相结合,以确保数据入库的质量与速度。

海洋基础数据管理系统的进一步开发,就是将空间数据仓库技术与空间数据挖掘技术运用到海洋科学信息系统中来,通过空间数据仓库技术、空间数据挖掘技术、元数据管理技术、网络技术以及 GIS 技术,将不同来源、不同格式的空间数据库资源有机结合起来,为用户提供统一、友好的界面,以人机交互方式实现海量数据的海洋科学信息分析系统。

3 海洋基础数据管理系统的基本规则

3.1 主题内容

海洋基础数据管理系统的记录格式应规定海底地形、沉积物、海流、地貌、生物等特征分布,信息系统的图层划分及各图层的属性文件格式,并对图层属性文件名及图元编码、数据项编号的结构做出规定。

3.2 适用范围

海洋基础数据管理系统的记录格式适用于海底地形、沉积物、海流、地貌、生物等特征分布信息系统的信息采集数据交换、存储管理、检索、空间分析、输出和共享。

3.3 引用标准

引用标准见表 1。

表 1 引用标准

Tab.1 Citing standard

标准编号	标准名称
GB 958-89	区域地质图图例(1:50000)
GB 2808-81	全数学式日期表示法
GB 6390-86	地质图用色标准及用色原则(1:50000)
GB 9649-88	地质矿产术语分类代码
GB/T 13923-92	国土基础信息分类与代码
GB/T 13989-92	国家基本比例尺地形图分幅编号
DZ/T 0001-91	区域地质调查总则(1:50000)
DZ/T 0157-95	1:50000 地质图地理底图编制规范
DZ/T 0160-96	1:200000 地质图地理底图编绘规范及图式
GB 2260	中华人民共和国行政区划代码
DZ/T XXXX-1999	海区地貌图编图规范(送审稿)
GB 3100-3102	量和单位
DDB 9701	数字化地质图图层及属文件格式
DDZ 9701	资源评价工作中地理信息系统工作细则
DDB 9702	GIS 图层描述数据内容标准 山东省区域地质图计算机制图技术标准(试用) 空间数据库工作指南(征求意见稿)等一系列国家及部委的相关标准规范

3.4 海洋基础数据管理系统的术语

该术语是记录格式术语应采用固定定义:图元是图面上表示空间信息特征的基本单位,分为点、弧段、多边形 3 种类型。图素是图内信息中的各种实体类型,由代表各类实体的若干图元构成。图层是图内由一类图素或性质相近的一组图素的矢量数据,以及用于描述这些图素特征的属性数据为一个图层;一幅图由若干个图层组成。数据项是属性数据中不可再分割的最小的单元。数据项代码是一个或一组有序的易于计算机和人识别与处理的符号。它是鉴别编码对象的唯一标志,还能提供编码对象的有关分类、排序逻辑意义等信息。数据类型是定义数据项所表现的数据属性,如:字符型 C,数值型 N,日期型 D。图层属性文件图层属性是各个基本图层所对应的属性表,是描述各图层基本属性的数据集合。

4 海洋基础数据管理系统结构分析

根据各专业领域特点, 参照标准, 建立合理的属性数据结构, 见图 1。

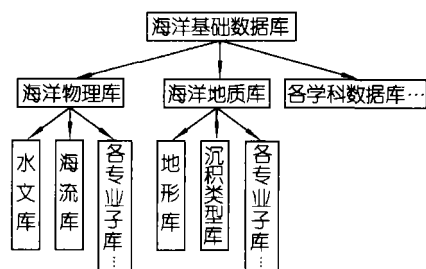


图 1 海洋基础数据管理系统结构

Fig. 1 Systematic configuration of marine basic data management system

4.1 资料预处理

由于历史的原因及各类资料采集设备的不同, 原始资料的数据结构与格式和应用标准间存在着很大差异, 需要编制相应的资料预处理软件对其进行抽取、整理、转换等处理; 另外还要将需要的图件及模拟资料数字化。

4.2 数据管理系统

数据管理系统主要包括数据的录入、维护、查询等功能。录入模块将预处理过的数据资料经筛选、质控检验入库; 维护模块负责对入库数据进行修正、剔除等操作; 查询模块可查询输出已入库数据内容。系统是面向数据管理与操作人员的。通过系统可保证数据准确无误地入库。

4.3 数据服务系统

系统包括信息发布、数据咨询服务等功能, 是面向最终用户的。

4.4 海洋基础数据管理系统的工作流程

工作流程见图 2。

5 海洋基础数据管理系统的技术方案

海洋基础数据管理系统基本分成两部分, 一是 GIS 的建立, 二是空间数据库与数据挖掘平台的开发。

建立海洋科学 GIS 系统面临两个主要问题, 第一是原始数据量大, 数以百 G 的数字资料, 大量的图件和模拟资料; 第二是原始数据类型复杂, 不同的学科特点, 不同的专业标准, 不同的存储设备, 不同的数据格式。这给历史资料入库以及将来数据扩充带来了很大的困难。为确保资料完整准确地入库, 首先要对历

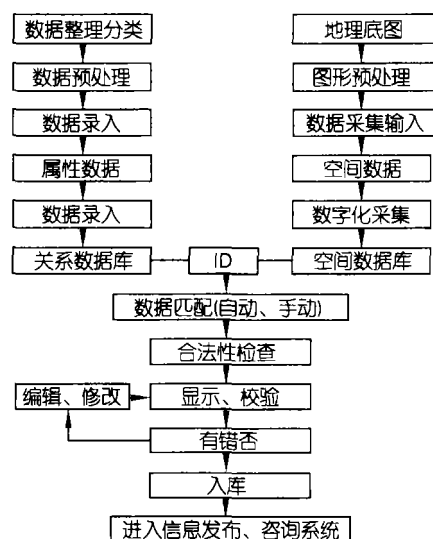


图 2 海洋基础数据管理系统的工作流程

Fig. 2 Flowchart of marine basic data management system

史资料、数据采集设备的输出数据的格式、精度、分辨率进行深入细致的剖析, 在此基础上建立有效的数据过滤转换通道, 包含存储介质转录、数据格式转换、数据点抽稀、插值、排重、数据质量评价与标志, 最终完成数据入库工作, 进而建立相应的信息发布与数据服务系统^[5]。数据库的建库过程中, 要保证结构合理性、数据库的开放性和可扩展性。数据库为一动态结构, 可以依今后调查研究的最新成果, 不断对数据库进行更新与扩充。

海洋科学空间数据仓库开发工作的主要任务是首先对来源于海洋空间数据中类型不同、结构不同、存储格式不同的、内容与格式丰富多采的与海洋科学相关的原始数据进行标准化、过滤与匹配、净化、表明时间戳和确认数据质量的处理, 即求精过程; 然后再根据任务的需要, 对来源于分布式数据库的不同性质、不同格式的相关数据再进行集成与分割、概括与聚集、预测与推导、翻译与格式化、转换与再映像处理; 然后再进行数据库仓库的建模、概括、聚集、调整、确认及建立结构化查询等, 它能够定义和生成各种领域专用模型, 如海域规划模型、海洋地质灾害评价模型等; 运用这些模型对数据进行有效分析, 并把分析结果或是决策咨询建议以直观、清晰的形式输出。需要特别强调的是, 与用于工商业领域的数据库系统不同, 海洋科学空间数据仓库中 GIS 系统的灵活运用将起到至关重要的作用, 因为只有利用 GIS 系统, 才

能准确描述各种海洋数据属性,而数据分析和成果展示及输出中,也离不开 GIS 系统的支持。

由于海洋科学信息系统中的数据具有多来源性,

这就决定了网络技术和分布式数据库技术会在今后的海洋科学信息系统发挥日益重要的作用,加上空间数据仓库技术,可将海洋科学信息组织为如图 3 所示

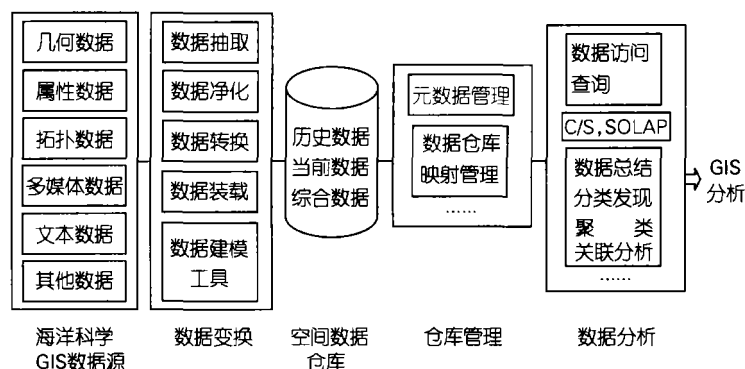


图 3 海洋科学信息组织的整体结构

Fig. 3 Integration of marine science organization

6 结束语

海洋基础数据管理系统不仅要服务于海洋的科学研究,同时服务于海洋预报、海洋工程环境的评估、军事海洋学的战场环境准备、生物多样性的研究、海域使用评价与论证以及海洋科普工作等国家需求的各个方面。

综上所述,可以看出,系统的建立与开发,不仅对海洋科学研究有强劲的推动作用,还有着广阔的市场前景。不仅具有重要的经济意义,还具有极大的社会效益。同时还会推进我国海洋产业部门及管理机构的信息化建设。作者提出的方法虽然只是一种尝试,但在实践中已表明它是可行的和有前途的。

参考文献:

- [1] 黄杏元,马劲松,汤勤. 地理信息系统概论[M]. 北京: 高等教育出版社, 2001. 24- 25.
- [2] Mallach E G. Decision Support and Data Warehouse Systems[M]. 北京: 电子工业出版社, 2001. 3.
- [3] 陈常松,何建邦. 基于地理要素的资源与环境数据的组织方法[J]. 地理学报, 1999, 54(4): 373- 381.
- [4] 吴信才. 地理信息系统的基本技术与发展动态[J]. 地球科学, 1998, 23(4): 329- 333.
- [5] 李德仁,关泽群. 空间信息系统的集成与实现[M]. 武汉: 武汉测绘科技大学出版社, 2000.

Discussion on design of marine basic data management system

LUAN Zhen- dong, FAN Feng- xin, YAN Jun

(Institute of Oceanology, the Chinese Academy of Sciences, Qingdao 266071, China)

Received: Mar., 13, 2003

Key words: geography information systems; data warehouse; database configuration

Abstract: This paper studied the method of rapid information extraction from expanding data, and discussed the necessity of constructing a marine basic database. A basic idea is put forward to construct a basic data management system on the basis of database and GIS.

(本文编辑: 刘珊珊)