

海岸带海洋遥感与地理信息系统系列丛书

Hai'andai Haiyang Yaogan yu Dili Xinxi Xitong Xilie Congshu

海岸带及近海 科学数据集成与共享研究

杜云艳 周成虎 苏奋振 王敬贵 李四海 李加洪 等著



海洋出版社

海岸带海洋遥感与地理信息系统系列丛书

海岸带及近海科学 数据集成与共享研究

杜云艳 周成虎 苏奋振 王敬贵 李四海 李加洪 等著

海洋出版社

2005年·北京

内容简介

本书旨在围绕海岸带及近海多源数据的集成与共享,从底层数据实体的设计、建设,数据技术系统的研发,多源、多格式数据的集成与发布等多方面探讨数据集成的难点和欲解决的新思路及关键技术。全书共分三个部分,首先,全面系统地阐述了海岸带及近海数据自身特点及基于此的数据平台实体设计;其次,进行地理信息系统空间数据的集成方法以及基于本体的空间数据显示和查询方法讨论;最后,探讨在地理本体理论指导下、采用空间数据引擎技术(SDE)、多用户视图动态生成技术、多源空间数据动态发布技术等研发的海岸带及近海数据集成平台系统。该技术系统按照集成等级的不同又分为基于 C/S 结构的数据库平台前端浏览和查询系统、基于 B/S 结构的网上信息共享系统和基于地理本体的用户级数据集成与共享系统。

本书可供遥感、地理信息系统、测绘、地理、海洋、大气等相关学科的科研人员、教师及研究生等阅读参考。

图书在版编目(CIP)数据

海岸带及近海科学数据集成与共享研究/杜云艳等著. —北京:
海洋出版社,2005. 11

ISBN 7-5027-6499-2

I. 海… II. 杜… III. ①海岸带—科学技术—数据—研究
②近海—科学技术—数据—研究 IV. ①P737.11 ②P72

中国版本图书馆 CIP 数据核字(2005)第 130811 号

HAI'ANDAI JI JINHAI KEXUE SHUJU JICHENG YU GONGXIANG YANJIU

责任编辑: 阎 安

责任印制: 刘志恒

海洋出版社 出版发行

<http://www.oceanpress.com.cn>

(100081 北京市海淀区大慧寺路 8 号)

北京顺诚彩色印刷有限公司印刷 各地新华书店经销

2005 年 11 月第 1 版 2006 年 6 月北京第 2 次印刷

开本: 787mm × 1092mm 1/16 印张: 16.75

字数: 400 千字 印数: 301~1200 册

定价: 68.00 元

海洋版图书印、装错误可随时退换

海岸带海洋遥感与地理信息系统系列丛书

编著委员会

主 任:周成虎

副主任:蒋兴伟 王 宏 罗续业 黄韦艮 张 杰

委 员:陈西庆 杜云艳 樊 伟 韩雪培 黄海军

李四海 林明森 刘宝银 刘高焕 刘建强

刘庆生 骆剑承 毛志华 裴相斌 苏奋振

唐军武 王 华 王钦敏 杨晓梅 杨燕明

恽才兴 赵冬至

秘 书:苏奋振 杨晓梅

丛 书 前 言

应对海洋战略,迎接海洋世纪,高新技术如初升旭日,勃勃生机!

我国 18 000 多千米的大陆海岸线和 300 多万平方千米的管辖海区,是中华民族持续发展的重要生存空间;海岸带是我国利用程度最高的国土,人口最密集、经济最发达;海岸带海洋也是海洋灾害频发和生态极其脆弱的区域,生存空间和可持续发展受严重挑战。

当前,我国海岸带在自然和人工作用下,变化剧烈。特别是经济的发展,海岸工程的建设,致使海岸冲淤变迁,直接影响国土开发与工程建设。与此相伴,生态环境持续恶化,部分水域出现荒漠化。由此,急需快速、及时、同步地获取海岸带资源环境及其演变信息,为海岸海洋的规划、开发和环境保护提供决策支持。

20 世纪 90 年代以来,国际高新技术与信息技术的迅速发展为海洋海岸带监测与信息服务提供了技术可能。当前,欧美纷纷开展区域海域遥感综合监测系统的研究,但岸线动态监测、滩涂信息提取、动力参数反演等仍然是难点。目前对地观测系统提供了海岸带近海动态监测的数据源,但现成的地理信息系统理论、方法和技术,严重缺乏对空间时序过程的处理分析能力。同时,海岸带诸多问题,需要集成多技术才能完成综合监测。为此需要完成对海量时空数据的综合集成、管理、自动处理和信息分发,从而及时、全面、实时、持续地提供资源环境信息。

由此,攻克难点,综合集成,建立海岸带遥感综合应用系统,开展综合多要素、涉及多部门、覆盖全国海岸带的研究,完成我国海岸带资源环境宏观、动态、持续的监测,已成为我国海岸带经济社会持续发展以及海洋监测技术发展的急需。本丛书正是为此所著。

丛书在解决海岸带遥感分类标准、多源遥感信息融合等关键技术的基础上,建立起多尺度的海岸带及近海卫星遥感综合应用系统,主要围绕基础平台建设和多层次应用示范进行论述。基础平台分为基于本体的海岸带及近海遥感科学数据平台和基于海洋地理信息系统的遥感综合应用技术平台。其中,遥感综合应用平台集成了相关遥感监测模块,共同组成中国海岸带及近海卫星遥感综合应用系统。应用示范包括全国、区域、省、市多级应用,涉及海岸带滩涂空间资源的监测与变化评价、生态环境信息服务、河口海岸侵蚀、滨海湿地保护等方面,示范区分别为长江口、黄河三角洲、渤海等重点典型区。

丛书编撰出版由周成虎研究员任主编,苏奋振博士与杨晓梅博士任秘书,汇集了丛书众多作者多年心血,但鉴于时间、技术发展阶段以及著者的学科背景等诸多因素,纰漏定有,请方家不吝指正。

海岸带海洋遥感与地理信息系统系列丛书

已出版专著列表:

《空间投影理论及其在遥感技术中的应用》

作者:任留成著 出版者:科学出版社 出版年:2001

《黄河三角洲生态环境动态监测与数字模拟》

作者:刘高焕、叶庆华、刘庆生等著 出版者:科学出版社 出版年:2001

《环中国岛链——海洋地理、军事区位、信息系统》

作者:刘宝银、杨晓梅著 出版者:海洋出版社 出版年:2003

《海岸带及近海卫星遥感综合应用技术》

作者:恽才兴主编 出版者:海洋出版社 出版年:2004

《中国海岸带与海岛遥感调查——原则、方法、系统》

作者:刘宝银、苏奋振著 出版者:海洋出版社 出版年:2005

《海岸带及近海科学数据集成与共享研究》

作者:杜云艳、周成虎、苏奋振、王敬贵、李四海、李加洪等著 出版者:海洋出版社
出版年:2005

《海洋地理信息系统——原理、技术与应用》

作者:苏奋振、周成虎、杨晓梅、张杰、骆剑承等著 出版者:海洋出版社 出版年:2005

《海岸带遥感综合技术与实例研究》

作者:杨晓梅、周成虎、杜云艳、王华等著 出版者:海洋出版社 出版年:2005

前 言

中国海岸带及近海科学数据集成共享平台是“中国海岸带及近海卫星遥感综合应用系统技术”的主要研究内容,它的建立有三方面的迫切需求。首先,“中国海岸带及近海卫星遥感综合应用技术系统”这一海洋监测高新技术项目的任务所需。其次,当前海洋研究领域多学科交叉研究现状也为海洋的多源数据集成提出迫切需求。随着海洋数据获取手段的扩展和计算水平的提高,综合利用多源数据、多种分析方法和技术手段成为研究海洋新的思路。第三,当前科学数据共享的大好形势也为海洋科学数据的共享提出了迫切要求,同时也提供了有利的发展环境。海岸带作为海洋开发的前沿阵地和依托场所,以海岸带为基线的科学数据集成平台的建立也是国内外的大势所趋。

因此,该科学数据平台的主要任务是:围绕海岸带这一特殊的海陆相互作用和人类活动频繁区域,向海、向陆不同程度地扩展,在海岸带区域以不同分辨率的遥感影像为主,在海域以多专题的海洋观测资料和场数据资料为主集成以往的研究成果,实现我国海岸带及近海资源环境信息的集成框架和松散的科学数据共享。

要实现这样的任务目标,必须解决两个方面的技术难题,即海岸带及近海多源数据组织和管理技术和地理空间数据的集成技术。针对前一个技术难题,在分析对比GIS对多源数据的无缝集成、多尺度海量遥感数据快速查询和检索及无缝拼接等技术基础上,采用目前最先进的大型空间数据库管理搜索引擎技术——基于 Oracle9i 的 ArcSDE,进行全国海岸带及近海底层数据平台的详细设计,给出多级海量遥感数据、多专题数据的逻辑模型和 ArcSDE 存储模型和检索机制,并利用 ArcGIS 桌面系统完成了海岸带及近海科学数据平台实体的装载和组织。为了便于使用和浏览底层数据平台实体,采用 ArcObject 组件和 ArcIMS 软件,用 VB 和 JSP 作为开发语言,开发了基于 C/S 结构的前端浏览和查询管理系统和基于 B/S 结构的网上信息共享系统,提供了良好的环境。

面对前面提到的数据集成难题,通过对 GIS 数据集成的深入研究,提出了一种自上而下与自下而上相结合的方法,即基于地理本体的空间数据集成思路。该集成方法放弃原有的单纯从底层数据入手的集成思路,从用户概念模型入手,探索不同用户的概念模型之间集成的可能性。采用本体的思想定量表达数据生产过程中的概念模型,并通过概念模型到数据模型的映射关系,最终实现从概念视图到数据的有效集成。

本书主要从三部分讨论该平台所涉及的科学问题及其研究内容。首先,从海岸带及近海科学数据的特点入手进行该科学数据平台实体的设计,见本书第1~4章;其次,进行地理信息系统空间数据的集成方法以及基于本体的空间数据显示和查询方法讨论,见第5和第6章;最后,给出了我们在多年研究基础上所开发的海岸带科学数据集成平台技术系统,该技术系统包含了前面所提到的基于 ArcSDE 底层数据实体的前端浏览和查询系统、网上信息共享系统和基于地理本体的数据集成系统以及基于海湾本体的福建省海岸带海湾数据集成平台的实例研究,具体讨论见第7章和第8章。

本书是资源与环境信息系统国家重点实验室海岸带与海洋研究小组多年来精诚团结、刻苦攻关的研究成果。本研究的出发点是通过不懈的努力,力求在满足攻关项目目标的基础上,更广泛的收集海岸带及近海多源科学数据,并能在海岸带及近海科学数据集成上开辟一条新思路,探索一套更为实用的技术手段。本书把已成雏形的中国海岸带及近海数据集成平台展示在诸位的面前,以期达到抛砖引玉、投石击浪的目的。

目 次

第1章 概述	(1)
1.1 中国海岸带及近海科学数据集成平台研究的意义	(1)
1.1.1 我国海洋高新技术监测发展的需求	(1)
1.1.2 海岸带遥感技术应用的需求	(1)
1.1.3 科学数据共享的国家需求	(3)
1.2 地理本体的概念及其研究现状	(4)
1.2.1 地理本体的概念	(4)
1.2.2 地理本体的国内外研究现状	(6)
1.3 基于地理本体的海岸带及近海数据集成平台	(12)
1.3.1 地理空间数据集成研究概况	(12)
1.3.2 海岸带及近海多源数据的特点	(13)
1.3.3 海岸带及近海多源数据集成的技术难点	(14)
1.3.4 基于地理本体的海岸带及近海集成平台的提出	(16)
参考文献	(17)
第2章 海岸带及近海科学多源数据分析	(21)
2.1 海岸带及近海数据集成平台需求分析	(21)
2.1.1 数据需求分析	(22)
2.1.2 主要数据类型	(23)
2.1.3 数据平台实体构建的技术能力分析	(23)
2.2 海岸带及近海多源数据及其特点分析	(25)
2.2.1 多源数据总体特征	(25)
2.2.2 影像数据及其特征	(26)
2.2.3 基础背景数据及其特征	(27)
2.2.4 海岸带专题数据及其特征	(27)
2.2.5 海岸带文献数据与分析	(28)
2.3 海岸带及海洋数据标准与规范	(28)
2.3.1 海洋科学数据现有标准简介	(28)
2.3.2 海洋科学数据元数据标准和规范	(31)
2.3.3 海洋科学数据编码标准	(33)
2.3.4 命名规范设计	(35)
参考文献	(39)

第 3 章 海岸带及近海多源数据集成平台实体设计	(40)
3.1 中国海岸带及近海多源数据平台实体概念模型	(40)
3.1.1 多源数据平台的概念框架	(40)
3.1.2 中国海岸带及近海数据平台实体的概念模型	(41)
3.2 多源数据平台实体逻辑模型设计	(43)
3.2.1 海岸带及近海科学数据平台数据流程	(43)
3.2.2 海岸带及近海科学数据平台多层次数据组织及逻辑模型	(44)
3.3 基于 ArcSDE 的多源数据存储模型	(46)
3.3.1 遥感影像数据的 ArcSDE 存储模型	(47)
3.3.2 非影像数据的 ArcSDE 存储模型	(49)
3.4 海岸带及近海多源数据整编及质量控制	(50)
3.4.1 数据平台的数据准备和预处理	(50)
3.4.2 数据质量控制	(54)
3.5 数据平台实体数据组织及装载	(54)
3.5.1 数据组织	(54)
3.5.2 数据装载	(55)
参考文献	(59)
第 4 章 海岸带及近海科学数据集成平台质量控制	(60)
4.1 海岸带及近海科学数据集成平台质量控制概述	(60)
4.1.1 数据质量控制概述	(60)
4.1.2 海岸带及近海科学数据质量控制的必要性	(60)
4.2 海岸带及近海遥感数据质量控制	(61)
4.2.1 MODIS 数据质量控制及评价	(61)
4.2.2 TM/ETM+ 数据质量控制及评价	(66)
4.2.3 资源 2 号卫星影像质量控制及评价	(74)
4.3 海岸带及近海矢量数据质量控制	(78)
4.3.1 海岸带调查图集的概要介绍	(78)
4.3.2 海岸带调查矢量化前的质量控制	(78)
4.3.3 海岸带调查矢量化过程的质量控制	(82)
4.4 海岸带及近海观测资料的质量控制	(91)
4.4.1 海洋浮标(Argo)观测资料介绍	(91)
4.4.2 Argo 数据读取过程(转换)质量控制	(94)
4.4.3 Argo 数据总体的质量评价	(100)
参考文献	(100)
第 5 章 基于地理本体的空间数据集成方法研究	(101)
5.1 地理空间数据集成研究概况	(101)

5.1.1	地理空间数据集成概念及其类型	(101)
5.1.2	地理空间数据集成研究现状	(103)
5.1.3	当前地理空间数据集成研究面临的问题及其解决方案	(105)
5.2	基于地理本体的空间数据集成框架	(107)
5.2.1	不同的集成架构及其比较	(107)
5.2.2	混合架构下的集成方案	(108)
5.2.3	集成流程	(110)
5.2.4	地理本体集成方法	(110)
5.2.5	集成过程中数据的抽取与转换方法	(116)
5.3	土地利用本体和数据集成示例	(120)
5.3.1	土地利用本体的创建	(120)
5.3.2	土地利用本体的集成	(121)
5.3.3	土地利用数据的转换与集成	(124)
	参考文献	(125)

第6章 基于本体的海岸带及近海概念模型及显示查询方法

6.1	基于本体的中国海岸带及近海的概念模型	(127)
6.1.1	海岸带的定义	(127)
6.1.2	基于本体的海岸带及近海概念模型	(128)
6.2	中国海岸带及近海本体模型构建	(131)
6.2.1	一般本体的建模方法与表示语言	(131)
6.2.2	地理本体的建模方法与表示研究	(141)
6.2.3	中国海岸带及近海地理本体的建模	(155)
6.3	基于海岸带及近海本体模型的空间数据查询和检索方法	(161)
6.3.1	基于地理本体的空间数据显示方法研究	(161)
6.3.2	基于地理本体的空间数据查询方法研究	(168)
	参考文献	(175)

第7章 中国海岸带及近海科学数据平台技术系统

7.1	数据平台技术系统框架	(177)
7.2	基于C/S结构的中国海岸带及近海科学数据组织和管理系统	(178)
7.2.1	数据平台管理系统主要功能设计	(179)
7.2.2	数据平台管理系统开发设计	(179)
7.3	基于B/S结构的中国海岸带及近海网上信息服务系统	(186)
7.3.1	中国海岸带及近海科学数据网上信息共享系统功能设计	(186)
7.3.2	中国海岸带及近海科学数据网上信息共享系统开发设计	(186)
7.4	基于地理本体的海岸带及近海数据集成平台原型系统	(190)
7.4.1	地理本体原型系统的功能设计	(190)

7.4.2 基于地理本体的数据集成原型系统实现	(193)
第8章 基于海湾本体的福建省海岸带数据集成平台示例	(199)
8.1 福建省海岸带地理空间数据库	(199)
8.2 基于本体的福建省海岸带海湾空间数据库	(199)
第9章 总结与展望	(209)
9.1 总结	(209)
9.1.1 实用化的海岸带及近海科学数据平台的实体设计方案	(209)
9.1.2 基于地理本体的空间数据集成方法	(209)
9.1.3 基于地理本体的空间数据分级和分类显示及智能化查询方法 ..	(210)
9.1.4 海岸带及近海科学数据平台技术系统研发	(210)
9.2 展望	(211)
9.2.1 基于地理本体的海岸带/海湾应用的进一步研究	(211)
9.2.2 地理本体中空间关系的形式化方法的深入研究	(211)
9.2.3 地理本体在地理信息网络发布和服务中的应用研究	(211)
9.2.4 各技术系统的进一步完善	(212)
附录 A 中国海岸带及近海多源遥感数据数据源清单	(213)
附录 B 土地利用类型编码说明	(222)
附录 C 海岸带资源调查图集几大专题的编码设计表:以福建省为例	(225)
附录 D 遥感数据的预处理流程	(229)
附录 E 各种类型 TM/ETM+影像几何校正精度评价结果	(250)
附录 F 五类 TM/ETM+影像的控制点分布图及各控制点误差	(252)

第1章 概述

1.1 中国海岸带及近海科学数据集成平台研究的意义

1.1.1 我国海洋高新技术监测发展的需求

我国从1996年开始将海洋高新技术作为第八个领域列入了国家高技术研究发展计划(“863”计划)，“九五”期间，该领域设立了3个主题，3个重大项目，25个专题，102个研究，46个青年基金，23个探索性研究，其中卫星遥感海洋应用技术作为海洋监测技术主题(818)中主要的高新技术之一，设立了“海洋渔业遥感信息服务示范系统”、“珠江口海域环境污染综合监测系统”和“海洋遥感应应用关键技术”等应用专题。经过四年努力，在众多青年学者和老一辈海洋专家的共同努力下，已经在各个主题取得了丰硕的成果。其中“海洋渔业信息服务技术”研制开发了具有自主产权的海洋渔业遥感信息处理、分析与服务软件系统，在海洋环境遥感信息反演、海洋渔业数据综合处理分析及中心渔场预报专家系统等方面取得重大进展，可以为渔业生产指挥和管理提供即时信息服务。而“海洋渔业遥感信息服务示范系统”完成了卫星高度计遥感信息应用技术，合成孔径雷达(SAR)遥感应应用技术，海洋光学遥感信息应用技术，卫星遥感海冰监测技术，海面风场和海流场多源遥感信息复合分析技术。所有这些项目的研究积累了大量的科学数据及相应的信息提取算法。一些主要成果已应用于海洋产业，获得了良好的社会效益。如何把这些宝贵的科研成果进行有效的集成并按照一定的机制进行领域内的信息共享，已成为巩固海洋高新技术研究成果及更有效地服务于海洋产业的关键。

此外，随着海洋监测技术主题在“863”的“十五”计划的延续，海洋卫星遥感综合应用技术依然作为一大专题列入重点发展计划。而“中国海岸带及近海卫星遥感综合应用技术系统”作为该专题的第一研究课题，担负着研究全国海岸带和近海的资源环境监测以及集成该主题中相应的遥感信息技术成果、数据成果的重要任务。为此，从数据集成和科学研究的持续性角度考虑，拟建立国家级的中国海岸带及近海科学数据共享平台。基本出发点是围绕海岸带这一特殊的海陆相互作用和人类活动频繁区域，向海、向陆不同程度地扩展，在海岸带区域以不同分辨率的遥感影像为主、在海域以多专题的海洋观测资料和场数据资料为主集成以往的研究成果，实现我国海岸带及近海资源环境信息的集成和数据共享。

1.1.2 海岸带遥感技术应用的需求

海岸带科学数据集成共享平台的建立也是当前海洋信息应用的迫切需求。总体上

看,海洋现象的数据获取手段主要有现场观测、数值模拟和遥感等(侍茂崇等 2000)。海洋领域的科学研究基本上是分别围绕这三种数据源进行各自不同的方法和应用实验研究。

直接观测资料以其“真实性和离散性”的特点对模式检验、海洋卫星遥感资料的校验应用方面起着不可替代的作用;而卫星海洋遥感资料的出现,以其“大面积、准同步、多时相、廉价”的基本特点为海洋现象的研究提供了崭新的数据集。在海表温度,叶绿素浓度,悬浮质浓度,DOM 浓度,海洋初级生产力,海洋光学参数等方面源源不断的提供着海洋环境参数和信息;同时由于数值技术和计算机技术的发展,复杂、精确、高效的海洋数值模式不断的诞生(如美国普林斯顿大学的 POM 模式)。这些功能强大的数值模式,可以模拟现实状况下的海洋运动,经过实测数据的验证,为海洋现象的研究提供大量的场数据(冯士筭 1999)。

这三种主要的数据源为海洋科学的研究提供了有力的数据基础,但由于各自的局限性,虽然有以上多种数据来源,但是很难说有哪一种单一的数据源可以提供一幅完整的海洋现象特征视图。因此单纯的利用某一种数据源并不能非常准确、全面的研究各种海洋现象。如利用数值模拟进行的研究,可以在 x, y, z, t 四维上给出海洋的定量描述,但时空分辨率太粗,误差也比较大。给出的模拟结果往往是“概况”视图,不能反映具体环境现象,对于无法获取其他数据的地方,可利用该类数据;而利用观测数据进行的数据分析和处理,虽然在局部区域可以比较真实和详细的给出海洋状态,是海洋复杂运动的真实反映,但只限于特定调查区域范围,并不能从宏观上给出海洋的整体状况;随着卫星遥感资料的广泛获取,科学家们应用卫星数据广泛开展了海洋的应用研究,使用最多的是卫星反演的海表温度,利用海表温度研究了 El-Nino 现象、黑潮和湾流的特征、探测了诸多中尺度涡旋(李凤歧等 2000),并结合其他数据研究全球气候变化,研究得出了太平洋暖池的平均温度和位置,赤道海域 Kelvin 波、Rossby 波的传播过程,中尺度涡旋、上升流、锋面的变化状况,计算海洋热收支,二氧化碳气体交换系数,并在小尺度海洋动力特征方面,研究了涡动的精细结构及海气相互作用。而微波传感器主要用于探测海洋动力参数,其中较为成熟的是(星载)微波高度计(Altimeter, ALT)。通过对海平面高度、有效波高、后向散射的测量,可同时获取浪、流、潮、海面风速等重要动力参数。卫星资料虽然是同步覆盖,能提供大量的多时相、多平台信息,能够反演各种海洋环境参数,但只提供表层信息。因此,在进行卫星海洋应用研究的时候,单纯的利用遥感反演的表层数据并不能全方位给出海洋的真实状况,充其量也只是反映了部分海表现象的真实状况,其次由于受地形和海陆相互作用的影响,卫星反演的数据在海岸带和近海区域的精度并不能得到有效的保证,同样给海岸带及近海区域的海洋现象的研究带来了缺陷,还必须考虑现场观测和卫星遥感同步发展,相互验证和补充。

鉴于上述原因,在当前海洋观测的手段和计算机技术飞速发展的现实状况下,如何有效综合使用这些数据是全面地进行海洋研究的关键。为此,不仅需要在技术上研究数据相互之间的融合,出现了“数据同化”(数据模式中加入观测数据)和“数据融合”技术(多源信息进行多层次处理以提高信息质量的过程),同时还需要从数据的角度保证多源海洋数据的有效集成和存储。本研究拟建立的“中国海岸带及近海科学数据平台”一定

程度上就是希望在研究海洋多源数据集成技术的基础上有效地组织和管理海洋观测数据、卫星遥感数据及其反演的海表要素场数据和各种模式计算的导出数据,一定限度内为海洋科学研究的人员提供所需要的特定区域特定时间段的数据,同时为加强海洋领域内各个应用学科之间的合作、交叉研究提供最基础的数据保障,从而为提高我国海洋监测技术水平奠定坚实的信息基础。

1.1.3 科学数据共享的国家需求

除了上述两个原因外,建立“中国海岸带及近海科学数据集成平台”的另一个重要原因就是当前国内外关于科学数据共享的大趋势。

近年来,科学数据的管理与共享工作已经成为我国基础数据库建设迫切的任务。科技部在 1999 年就实施了“国家科技基础性工作专项”,陆续开展一批急需的科学数据库建设。为了推进我国的科学数据共享工作,我国较早的加入了世界数据中心(简称 WDC)组织,并组建了 9 个学科中心(孙九林等 2002)。

当前科学数据共享工作的大好形势,为海洋科学数据共享平台的建立提出了迫切的要求,同时也提供了有利的发展环境。科学家预言,21 世纪是海洋世纪,海洋势必在未来的经济发展、人民生活中占据重要的地位。而海岸带作为海洋开发的前沿阵地和依托场所,更是人类从内陆走向海洋首先要关心的领域。

多年来,我国政府一直重视海洋基础数据的管理与共享工作,国家海洋局海洋信息中心就是在这一思想的指导下应运而生的研究机构,从 20 世纪 80 年代开始,一直致力于海洋基础数据的收集和管理工作。目前该机构主要是面向海洋基础数据库的建设,其内容主要涵盖海岸带、海岛、大洋等普查和综合调查资料。

随着遥感技术的发展,时至今日,世界各国的遥感卫星相继成功发射,以不同的时空分辨率实时的监测着地球,其中专门应用于海洋监测与研究的卫星相继入轨运行,从海表温度、高度、风场、海浪、水色等方面大面积、周期性地监测着海洋,其中比较典型的有海岸带水色扫描仪(CZCS)、海洋宽视场传感器 SeaWiFS 和装载有高度计的 Topex/Poseidon,ERS-2,Geosat 卫星等;此外,根据特定目的还进行了一些航空遥感探测,这些都为海洋科学研究人员提供了崭新的数据集。由于这些数据集的多时相、大面积等特点,对于海洋各种现象的研究提供了前所未有的现场资料。针对海量遥感数据,当前各个海洋研究机构针对自己的研究问题,在相应的遥感数据上进行各种分析,无论是在数据还是在成果上都在一定程度上相互保密,致使资料不能有效沟通,一定程度上造成了资料的重复、国家投资的巨大浪费,制约了我国海洋科学研究水平的提高。

为此,从科学研究的角度建立以遥感数据为主的海洋科学数据共享平台已成为解决海洋研究领域数据共享问题的有效途径。该数据平台拟采用松散的共享机制,在探讨海量数据管理、组织、集成和共享等关键技术问题基础上,为实现各个研究机构数据共享提供整体技术支撑,从而为海洋科学的研究提供一个有利的数据支撑环境。鉴于海洋数据的多样性和时空复杂性等难点,本平台围绕中国海岸带首先建立以遥感数据为主的多源数据集成平台,一俟在方法和技术上进行攻关,进而沿着海岸带向近海和内陆延伸,最终形成比较系统的全国海岸带及近海科学数据集成共享平台。

1.2 地理本体的概念及其研究现状

1.2.1 地理本体的概念

本体论(Ontology)是一个哲学概念,最早出现于16世纪后期,由德国经院学者郭克兰纽(1547—1628年)在其著作中第一次提出和使用了本体一词(桑新民 1993),最早由人工智能界引入计算机或信息科学领域。1991年Neches等最早在AI领域使用“本体”这个术语,并将本体定义为“构成某个问题领域词汇的基本术语和关系以及组合这些术语和关系以规定词汇外延的各种规则”。Gruber是将本体方法用于信息科学的先锋之一,他将本体定义为“概念模型明确的规范说明”(Gruber 1993)。经过十多年的讨论,目前计算机学界对什么是本体还存在分歧,但大多数学者普遍认可Gruber的定义。其后Borst在其博士论文中强调概念模型必须是公共认可的和形式化的,这样有利于本体的共享和重用及计算机处理,并将本体定义为“共享概念模型的形式化规范说明”(Borst 1997)。1998年Studer等综合了两个定义的优点,并认为本体应该对领域知识中的各种常识、隐含假设进行显式说明,从而将本体定义为“共享概念模型的明确的形式化规范说明”。这个定义包含“概念模型”,“明确的规范说明”,“形式化”,和“共享”四层含义(Uschold, Gruninger 2004; 邓志鸿等 2002; Studer et al. 1998)。

由于不同学科或领域对本体的使用目的和方式不同,其定义存在偏差甚至可能完全不同。从20世纪90年代末开始,本体概念及其思想引起了地理信息科学研究者的注意,许多国际学术组织和研究机构开展了一系列相关研究计划,地理信息科学界的许多国际学术会议也设立了本体相关的讨论主题,这些研究计划和学术会议就地理本体的基础理论、方法及其应用问题进行了卓有成效的探索和研究,取得了一定进展。那么从本体发展到地理本体,是如何定义的呢?

地理本体在有些文献中定义为“地理心智模型的显式形式化”(Fonseca 2001)或“研究地理信息科学领域内不同层次和不同应用方向上的地理空间信息概念的详细内涵和层次关系,并给出概念的语义标识”(孙敏等 2004)。显然,这两个定义都有可取之处,但都不够完全表达本体的含义。前一个将地理本体形式化说明对象限定为地理心智模型,没有明确说明这种心智模型的涵义,如果心智模型指的是整个学科感知世界而不是地理认知世界,则这个定义还是可以接受的;后一个将地理本体的内容局限在“地理空间信息概念的详细内涵和层次关系及其语义标识”上,而没有将概念实例及其之间的空间关系纳入到地理本体中,显然也是不完整的,而且还混淆了地理本体与地理本体理论之间的区别。因此,在一般本体定义的基础上,考虑到地理本体的特殊性及其内容的广泛性和丰富性,我们认为地理本体可以定义为:地理本体是指特定地理空间信息共同体概念化模型或学科感知世界明确的形式化规范说明。

所谓的信息共同体,指的是至少在一段时间内共享同一数字地理信息语言和空间要素定义的人群,例如某个政府机构或集团、某个行业、同一学科中的一群研究者、某个项

目上的合作伙伴,等等。相同的信息共同体共享相同的世界观和抽象模型、要素表达和元数据(OGC 2003)。对于一个小的信息共同体,他们共享粒度非常细且特别具体和详细的地理概念化模型,相应的地理本体就是详细程度很高、应用范围很小的低层本体;对于一个由多个小的信息共同体组成的更大的信息共同体,他们只共享了粒度更粗、范围相对较大而详细程度较低的地理概念化模型,相应的地理本体就是详细程度较低、应用范围较广的高层本体。因此,特定的地理本体总是与特定的地理空间信息共同体相对应的。

地理本体形式化说明的对象是某个地理空间信息共同体的概念化模型或学科感知世界。在这里,概念化模型不同于通常意义上的概念模型,指的是人们对地理现实世界的多个抽象层次,包括地理认知世界、地理概念世界、地理空间世界、地理尺度世界及地理投影世界,等同于所谓的学科感知世界(图 1-1)。因此,地理本体既可以是地理认知模型或概念模型的显式形式化,也可以是逻辑层次或表达层次上地理模型明确的形式化说明。

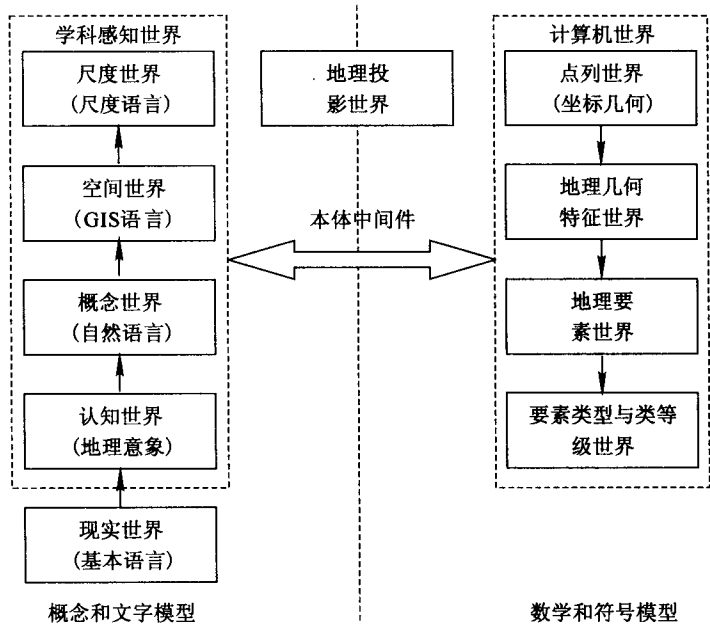


图 1-1 地理现实世界的抽象层次

地理本体通过明确说明地理空间信息共同体的概念化模型或学科感知世界,可以为地理数据使用者提供数据生产者的学科感知世界;通过概念或类的明确说明,可以避免术语使用上的混乱和歧义,因而可以解决认知和命名这两种语义异质性。地理本体是与地理数据相分离的,而不是像数据/知识包那样将数据与语义存储在同一个结构中,这样可以避免为在语义上同类的地理对象重复存储该语义描述符,从而解决了数据冗余问题;而地理本体通过建立与地理数据集之间的映射或关联关系,可以有效地表达地理数据集的语义。此外,地理本体通过采用标准的表示工具和语言来进行描述和表示,或通