

# 面向网络的海量影像空间数据在线分发技术\*

朱欣焰<sup>1</sup> 陈能成<sup>1</sup> 王 密<sup>1</sup> 刘 琳<sup>1</sup>

(1 武汉大学测绘遥感信息工程国家重点实验室 武汉市珞喻路 129 号, 430079)

**摘 要:** 阐述了基于 J2EE 的海量空间数据在线分发系统原理和关键技术, 给出了海量影像空间数据在线浏览发布的实例。

**关键词:** 海量空间数据; J2EE; 影像数据发布; 无缝影像数据库

**中图法分类号:** P208

传统的影像空间数据分发方式主要靠纸质或磁盘、光盘介质, 不能快速、实时地提供最新的信息, 用户对信息的内容难以全面地了解和掌握。计算机网络的迅速发展和广泛应用, 为以计算机网络为基础的空间信息的分发提供了基础硬件平台。在系统构建方面, 分布式系统的构建为共享信息平台的建立提供了技术基础。目前, 流行的分布式技术体系主要包括 OSF 提出的 DCE, OMG 提出的 CORBA, Microsoft 提出的 DNA 和 Sun 提出的 J2EE。

空间信息的共享与分发包括元数据的管理、空间数据的下载、空间信息的在线浏览等方面。元数据提供共享信息的特征资料, 使数据用户可据此来确定该数据的名称、来源、组织结构、适用范围、质量状况等。通过对元数据的了解以及必要的管理流程和权限检查, 用户可以下载所需数据。在线浏览包括元数据的在线浏览和数据内容的在线浏览。由于空间信息(特别是卫星遥感信息)数据量大, 空间信息在线浏览与利用已经成为地理信息服务的重要研究方向。国内外许多学者致力于空间数据共享与分发方面的研究<sup>[1~3]</sup>, 并已建成了许多实际运行的系统, 但目前大部分系统采用数据在线下载方式<sup>[1]</sup>, 很少做到大数据量的空间数据在线浏览。在大数据量的在线浏览方面, 微软的 TerraServer 系统是目前世界上最大的系统<sup>[4]</sup>, TerraServer 可以在 Web 环境下供多用户实时在线浏览航空、卫星和地形数据, 并提供相应图像数据的元数据信息。大数据量的空间数据在

线浏览要求服务器有快速的响应速度, 解决的办法除在服务器端提高硬件的配置及组成服务器集群外, 还要处理好数据的组织、管理和调度, 这也是这方面的研究重点之一。文献[5, 6]对影像数据的组织、管理技术进行了较深入的研究。本文主要介绍基于 J2EE 体系的空间信息在线浏览与分发技术。

## 1 基于 J2EE 体系的空间数据分发系统的构成

J2EE 是一种利用 Java 2 平台来简化诸多与多级企业解决方案的开发、部署和管理相关的复杂问题的体系结构。J2EE 技术的基础就是核心 Java 平台或 Java 2 平台的标准版。J2EE 不仅具有“编写一次、到处运行”的特性, 具有便于存取数据库的 JDBC API、CORBA 技术以及能够在 Internet 应用中保护数据的安全模式等特点, 同时还提供了对 EJB、Java Servlets API、JSP 以及 XML 技术的全面支持<sup>[7, 8]</sup>。

### 1.1 空间数据在线分发服务系统构成

空间数据在线分发服务分为表示层、服务层和数据层三部分, 如图 1 所示。其中表示层为客户端; 服务层包括 Web 服务器、应用服务器和空间数据处理服务器; 数据层为空间数据库服务器。系统构建采用了 J2EE 分布式体系框架。

1) 客户端。客户端主要向服务器端提出空间

\* 收稿日期: 2003-03-10。

项目来源: 国家 863 计划资助项目(2002AA131040)。

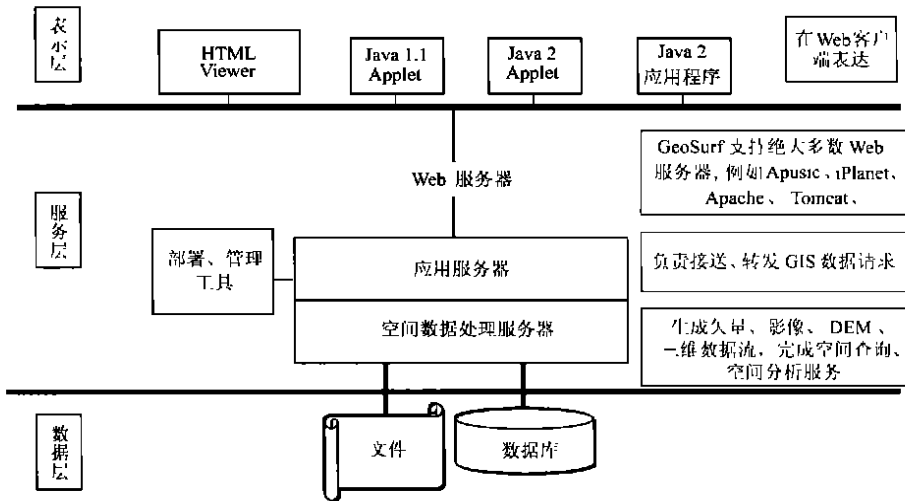


图 1 空间数据在线分发服务体系结构

Fig. 1 Architecture of Online Distribution Service of Spatial Data

数据服务请求, 并根据服务器端返回的结果进行实时处理、分析与显示。通常, 客户端的实现方法主要有 GIS 插件、GIS ActiveX 控件和 GIS Java Applets 几种。在本文基于 J2EE 体系的系统中, 客户端的主要实现技术采用 GIS Java Applets, 也可以采用基于浏览器的 HTML Viewer 或基于 JDK1.2 的应用程序。GIS Java Applets 支持栅格数据(JPEG、GIF)、影像数据流、矢量数据流、DEM 数据和三维数据流, 并且所有的表现均是基于 JavaBean 组件的方式提供, 用户可以根据需要定制自己的客户端。

2) Web 服务器。Web 服务器负责接受客户端的 GIS 服务请求, 传递给应用服务器, 并把结果送回给客户。Web 服务器选择目前支持 J2EE 的商业产品, 如 IBM Websphere、SUN iPlanet、Apache、Tomcat 等。

3) 应用服务器。应用服务器是一个基于 Java2 企业级版本的应用, 作为 Windows NT 的服务或者 UNIX 的后台进程存在, 应用服务器处理请求的负载平衡, 并且追踪空间数据服务运行在哪个空间数据处理服务器上。应用服务器可以将请求传递到一个适当的空间数据处理服务器, 处理后可以将结果通过 Web 服务器传给 Web GIS 客户机。

4) 空间数据处理服务器。完成空间应用服务的主要处理逻辑, 包括栅格地图、影像数据流、矢量数据流、DEM 数据流、三维数据流、空间查询和空间分析服务等, 生成的结果通过应用服务器经 Web 服务器传送给客户端。

5) 空间数据库服务器。存储和管理空间数据。空间数据遵循 OGC 的互操作协议和采用

Oracle9i 或 Oracle8i 的关系数据库存储。

### 1. 1. 1 客户端属性浏览技术

属性查询、动态页面浏览服务是动态网页生成的过程。动态网页生成的模式主要有 ASP、PHP 和 JSP 等。本文利用 Oracle 海量数据库系统管理数据, 用 Servlet 等高性能服务器端程序作为后台总控程序, Servlet 接受从客户端传来的用户请求, 分别调用不同的 JSP 程序向客户端反馈信息, JSP Servlet 通过 HTTP 协议在服务器端和客户端传递数据, 把参数传递给事先编好的 JavaBeans 和 EJB 组件, 对数据库进行操作, 如图 2 所示。元数据作为属性数据来存储管理, 客户端的浏览技术与属性信息类似。

### 1. 1. 2 客户端矢量、影像数据浏览技术

矢量、影像数据浏览(查询)服务模型如图 3 所示。由于现有的 Internet Explorer 或 Netscape Navigator 等不直接支持矢量地图及大数据量的影像和栅格地图的显示、漫游、查询, 因此必须扩充浏览器的功能。客户端信息表现包括矢量地图、影像、文本、声音、多媒体等, 系统主要是在表现和交互功能上。本文采用基于 Java Applet 的方式编写基于浏览器的用户界面, 并提供基于 JavaBean 的二次开发 GIS 组件。

### 1. 1. 3 应用服务器的实现技术

图 4 为系统的应用服务器。它响应客户端的请求, 并且处理请求, 得到客户端能够解析的页面, 把它发送到客户端进行显示。在应用服务器中, 涉及到多用户和并发处理请求的通过 Servlet 和 JavaBean、EJB 来处理, 而其中的关键业务逻辑也可采用 DLL.COM 来处理。

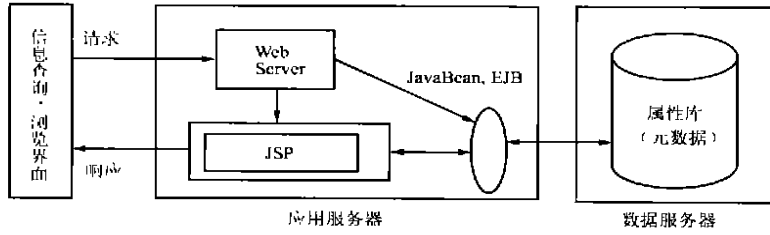


图 2 属性信息查询、浏览服务  
Fig. 2 Attribute Query and Browsing

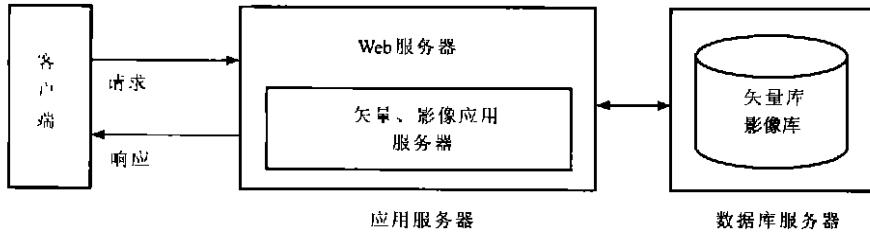


图 3 客户端浏览、查询服务  
Fig. 3 Browsing and Query on Client

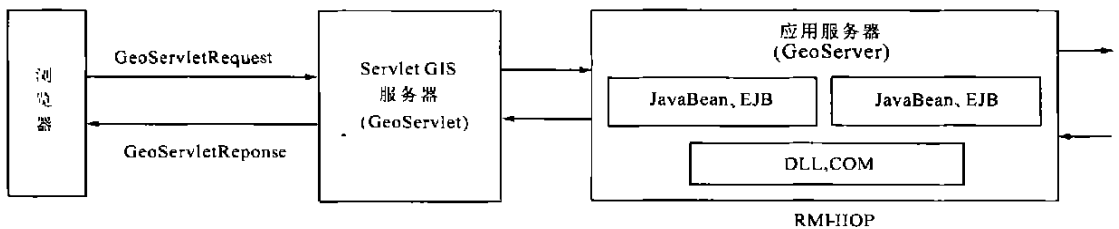


图 4 应用服务器  
Fig. 4 Application Server

其中, GeoServlet 是运行在 Web 服务器上的 Servlet 应用程序, 同时又是 RMI 客户机。系统通过它与客户机打交道, 接受请求和转送请求到 GIS 应用服务器中进行处理, 然后返回应用服务器的处理结果到浏览器。

GeoServlet Request 是 Servlet 服务的一个接口, 包括了客户端到服务器之间的联系 (包括客户端传送的阐述名称, 客户端正在使用的协议, 产生请求并且接受请求的服务器远端主机名, 请求的操作类型), 并且把这些请求通过远程方法调用 GIS 应用服务器上的 RMI 服务器, 启动相应的部件完成任务, 把任务传给 GeoServlet, 实例化 GeoServletResponse, 返回客户机。

应用服务器包括以下几个主要的组件: 用户权限验证组件, GIS 矢量数据请求, 应答服务组件, 动态页面生成, 页面管理, 影像数据请求。

### 1.2 空间数据在线分发服务流程

影像数据、矢量数据在线浏览服务流程如图 5 所示。用户对信息服务中心提出查询、浏览请求后, 首先由权限控制服务程序根据用户权限信

息进行检查, 通过了权限检查的请求, 被送到矢量、影像数据应用服务程序进行处理。应用服务程序根据请求的内容和范围, 利用空间索引调出相关的数据, 生成适合于屏幕显示的信息。这些信息经应用服务程序进行数据压缩、加密后, 通过网络传输到客户终端, 客户应用程序经过解密和解压后进行浏览、显示。如果用户的请求是要对指定范围的影像进行下载, 则生成下载文件, 并通知用户进行下载。数据加密是在数据传输过程中保证数据安全的一项重要手段, 可以采用信源加密和信道加密的方法。图 5 描述的流程为信源加密方法。数据的压缩是为了减少数据的传输量, 本文的数据压缩采用小波压缩技术。

## 2 海量遥感图像数据的在线浏览

海量遥感图像的在线浏览主要依赖于无缝影像库技术和数据缓存技术。无缝影像库技术针对海量影像数据的数据组织与管理, 主要包括影像分块技术、空间索引技术、数据的无缝组织技术和

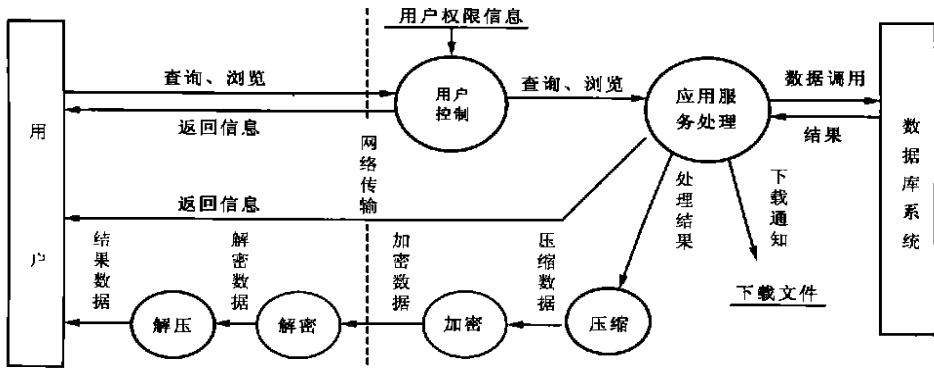


图 5 空间数据在线分发服务流程

Fig. 5 Flow of Online Distribution Service of Spatial Data

多分辨率金字塔的数据组织技术等。数据缓存技术主要解决在调度过程中减少数据 I/O 的次数, 从而提高效率。

### 2.1 无缝影像库技术

#### 2.1.1 影像分块技术

由于在海量影像数据库中, 每次调度和使用的图像数据只是数据库中的一小部分, 要高效地组织和管理数据, 数据分块是关键。通过数据分块可以减少数据的网络传输数据量, 方便数据压缩和有利于在计算机的内存中对图像数据进行运算处理, 因此对存储的影像数据采用影像分块和重构方法来组织数据。在本文中, 影像数据以关系数据库来存储与管理, 每个小图像块作为一个二进制字段存放, 形成一条记录。如图 5, 在图中所示的范围内, 按格网将数据分成 9 块, 对应表 1 所示的表结构。在数据分块时, 数据块的大小对图像调度效率有重要影响<sup>[9]</sup>。

#### 2.1.2 空间索引技术

常见的空间索引一般是自顶向下、逐级划分空间的各种数据结构空间索引, 比较有代表性的包括 BSP 树、KDB 树、R 树、R<sup>+</sup> 树、CELL 树、四叉树和格网型空间索引等。在上述空间索引中, 格网索引对于影像数据库来说非常合适。由于在影像数据库中数据按照分块的方式存储, 而且数据块的划分是非常规则的并且彼此之间没有重叠, 格网索引的建立和维护相当方便。按照格网对图像块进行编号, 然后通过编号和数据块的对应关系来索引图像块, 见图 6 和表 1。

在图 6 中, 设  $(X_0, Y_0)$  为格网的起始坐标原点,  $\Delta_x, \Delta_y$  分别是影像块的宽和高,  $X$  和  $Y$  方向的索引窗口的范围是  $(X_1, Y_1, X_2, Y_2)$ , 则

$$\text{起始格网的行号为: } \text{int}((X_1 - X_0) / \Delta_x)$$

$$\text{终止格网的行号为: } \text{int}((X_2 - X_0) / \Delta_x) + 1$$

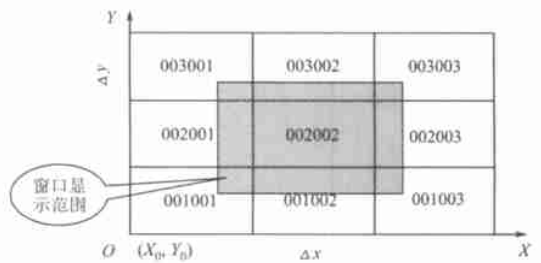


图 6 图像格网索引

Fig. 6 Image Grid Indexing

表 1 块号和数据块对照表

Tab. 1 Relation Between Blocks Number and Image Data

块号	数据块
001001	Image Block1
001002	Image Block2
001003	Image Block3
002001	Image Block4
...	...

$$\text{起始格网的列号为: } \text{int}((Y_1 - Y_0) / \Delta_y)$$

$$\text{终止格网的列号为: } \text{int}((Y_2 - Y_0) / \Delta_y) + 1$$

由起止格网的行号和起止格网的列号可以索引到窗口  $(X_1, Y_1, X_2, Y_2)$  所包含的数据块编号, 再在表中由数据块编号找到相应的数据块。

#### 2.1.3 数据的无缝组织技术

从理论上讲, 由于在地理坐标系下, 地球上的每一点都有相同的坐标基准, 因此如果按照地理坐标(经纬度)来进行数据组织是很容易实现无缝的, 目前大范围矢量数据建库的建立采用的就是将高斯坐标投影到经纬度, 以经纬度为基准来建库。由于矢量的数据量小, 可以利用严格的投影公式进行投影转换, 但是对于海量的栅格影像数据, 如果按照地理坐标来建库会给数据的使用和分发带来不便, 因此本文直接使用数据生产时的空间参考和原

始数据来建库,但是这样就会带来新的问题:由于每个投影带的空间基准不同,都有自己的坐标原点,使得整个数据库不能满足无缝的要求。但是由于相邻两带之间可以通过地理坐标建立起几何关系,在带边处,可以通过这种几何关系把相邻带内的数据实时投影到当前带内,与当前带内的数据无缝地镶嵌到一起同样可以满足无缝的要求。这样建库就保证了原始数据的完整性,同时又可以对整个数据库进行无缝漫游。

2.1.4 多分辨率金字塔的数据组织

为了快速实现由粗到精、由整体到局部的影像漫游、浏览,对影像采用多级金字塔的结构来组织数据,在进行放大、缩小时,不同金字塔层之间可以进行无缝切换。

2.2 数据缓存技术

对影像数据进行分块组织后,在显示某一区域的影像时,需要同时调入多个影像块进行拼接,显示的效率和调入影像块数以及每块的大小密切相关。在每块大小固定的情况下,调入的块数越多,且数据库的IO次数越多,花费的时间就越长。通常情况下,一个用户在影像的浏览过程中,相邻两屏先后有重叠的区域。如图7(a)所示,在先后相邻两屏影像浏览过程中,002001和002002这两个影像块是相同的。如果在前一屏调入后,在服务器端的内存中开辟一个缓冲区暂时保存,后一屏显示前就可以直接利用,而不需要对这些在缓冲区中存在的影像块重新从数据库调入,这就是缓存技术。图7(b)是4个用户同时调用空间上相邻的区域出现重叠影像块的情况。缓存可以在服务器端实现,也可在客户端实现。但在

客户端实现时,只能解决单个用户相邻几屏的数据重复调用问题,因此本文的缓存技术用在服务器端。调度时,缓存中影像块的淘汰方法可以用先进先出法、先进后出法或频率最小的影像块先淘汰的方法。通过缓存技术,可以很好地提高系统的效率。

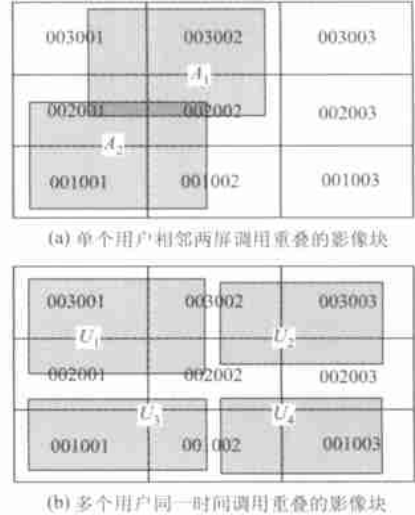


图7 基于缓存技术的影像调度管理

Fig. 7 Image Calling Management Based on Buffer Technology

2.3 实例

图8是航空影像在专网上在线浏览的实例。实例中按多比例尺金字塔结构建立无缝影像库,入库影像图1767幅,金字塔最底层图像分辨率为0.4m,总数据量约45.7GB,在IE浏览器下可以实现实时在线浏览。

实验系统基于J2EE分布式体系结构,在GeoSurf 4.0的基础上进行开发,数据库使用



图8 影像在线浏览实例

Fig. 8 An Instance of Image Online Browsing

Oracle9i。GeoSurf 是由武汉大学开发的跨平台、分布式、多数据源 Web GIS 平台, GeoSurfV 4.0 采用了 Java2 企业级(J2EE)的分布式体系框架, 具有高可扩展性、弹性配置、跨平台等特点。

### 3 结 语

本文讨论了面向网络的海量影像空间数据在线分发技术, 研究并实验验证了基于 J2EE 体系构建这类系统的有效性。应用服务器的构建采用无缝影像库技术和数据缓存技术, 较好地解决了海量空间数据的有效组织、管理、调度和快速响应问题。实验也表明, 在服务器端的关键算法实现上, 采用 C 语言比采用纯 Java 具有更高的效率。

#### 参 考 文 献

- 1 Su Y F, Slottow J, Mozes A. Distributing Proprietary Geographic Data on the World Wide Web — UCLA GIS Database and Map Server. *Computer and Geoscience*, 2000 (26): 741 ~ 749

- 2 刘荣高, 庄大方, 刘纪远. 分布式海量矢量地理数据共享研究. *中国图像图形学报*, 2001, 6(9): 865 ~ 872
- 3 龚健雅, 袁相儒. 跨平台分布式地理信息组织与处理. *武汉测绘科技大学学报*, 1998, 23(4): 364 ~ 369
- 4 Barclay T, Gray J, Slutz D. Microsoft TerraServer: A Spatial Data Warehouse. *The ACM SIGMOD*, Austin, 2000
- 5 方 涛, 龚健雅, 李德仁. 影像数据库建立中的若干关键技术. *武汉测绘科技大学学报*, 1997, 22(3): 266 ~ 269
- 6 王 密. 大型无缝影像数据库系统(GeoImageDB)的研制与可量测虚拟现实(MVR)的可行性研究:[ 博士论文]. 武汉: 武汉大学, 2001
- 7 屈晓晖. *Enterprise JavaBeans 编程实战*. 北京: 科学出版社, 2001
- 8 Karmoss. *Java Servlets 编程指南*. 北京: 科学出版社, 2000

第一作者简介: 朱欣焰, 教授。主要从事空间数据库、地理信息系统、Web GIS、空间信息移动用户应用服务等方面的研究。代表成果: 地理信息系统中多用户环境下数据共享一致性问题研究; 无缝空间数据库的概念、实现与问题研究。  
E-mail: Geozy@263.net

## Online Distribution Technology of Mass Image Spatial Data Based on Network

ZHU Xinyan<sup>1</sup> CHEN Nengcheng<sup>1</sup> WANG Mi<sup>1</sup> LIU Lin<sup>1</sup>

(1 State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing,  
Wuhan University, 129 Luoyu Road, Wuhan, China 430079)

**Abstract:** With the development of aerial and spatial remote sensing technology of China, the ground receiving station stores and manages mass spatial data at TB level. Thus, the traditional distribution mode cannot meet the application demand. So as to meet the demand of network infomazation, it is required that the mass spatial data be distributed, shared and processed on the special network. Functions as online browsing, distribution and processing of spatial data, especially the mass spatial data have become an important research orientation in geographic information service. The paper focuses on the principles and key technologies of mass spatial data online distribution system based on J2EE. An instance of mass image spatial data online browsing and distribution is given.

**Key words:** massive spatial data; J2EE; image data distribution; seamless image database

**About the first author:** ZHU Xinyan, professor. His main research orientation are GIS spatial database and spatial information services for mobile users. And his typical achievements are study on data consistency in spatial database system; concepts Implementation and problems of the seamless spatial database.

E-mail: Geozy@263.net

(责任编辑: 涓涓)