

RDBMS 和 COM 的海量遥感影像数据的管理和 Web 发布

吴信才¹⁾ 郭玲玲²⁾ 李 军¹⁾

¹⁾(中国地质大学(武汉)信息工程学院, 武汉 430074) ²⁾(北京大学数字地球工作室, 北京 100871)

摘要 介绍了基于关系型数据库(RDBMS)和微软组件技术(COM)的 Web 海量影像数据库(IMIDB)的研究背景、系统组织和实现;具体论述了 IMIDB 的数据库逻辑设计,包括数据库的选择、遥感数据的组织和存储;在此基础上,论述了存储在数据库中的遥感影像的 Web 发布体系结构,以及组件对象模型技术在该系统中的运用;最后给出了该系统的测试情况,测试结果表明,系统达到了设计目标。

关键词 遥感影像 海量影像数据库 关系数据库管理系统 组件对象模型 网络发布

中图分类号: TP751.1 TP311.133 TP393.4 文献标识码: A 文章编号: 1006-8961(2002)04-0384-04

Mass Image Database Management and Web Publishing Based on RDBMS and COM

WU Xin-cai¹⁾, GUO Ling-ling²⁾, LI jun¹⁾

¹⁾(Information Engineering Faculty, China University of Geosciences, Wuhan 430074)

²⁾(Digital Earth Studio, Peking University, Beijing 100871)

Abstract As a system for collecting, storing, analyzing, processing and managing geography spatial data, GIS's spatial data are mainly composed of vector data and grid data. The existing GIS systems are mostly based on vector data and have a great ability to establish and manage vector database. But they have poor ability to process grid data, especially to organize, schedule, store and manage mass grid data, without mentioning the ability to manage and integrate the multi-source, multi-scale and multi-phase image data. This paper does a deep research and a thorough discussion of mass image database and its realization and application. The paper mainly discusses the research background, system organization and implementation of mass image database(IMIDB) based on relational database system(RDBMS) and component object model(COM). It focuses on the logical design(including the selection of Database, the organization and the storage of remote sensing data); Web publishing architecture of remote sensing images in IMIDB and the application of its Component Object Model. Finally, a test of this system is done and the result of the test is given. From the result, we give a conclusion that the system has reached the aim of the design.

Keywords Remote sensing image, Mass image database, Relational database management system, Component object model, Web publishing

0 引言

随着遥感(RS)、全球定位系统(GPS)和地理信

息系统(GIS)日益紧密的结合,遥感影像数据逐渐成为 GIS 中的一种重要数据源,其数据量正呈几何级数增长.然而现有的 GIS 软件却难以组织、调度、存储和管理这样的海量数据,更没有考虑到多数据

源、多尺度、多时相影像数据的统一管理和集成等问题。此外,国家基础地理信息产业建设也迫切需要对海量影像数据进行高效、快捷地存储与管理。另一方面,Internet 技术的飞速发展使地理信息资源的共享在技术上成为可能,利用 Internet 实现地理信息系统(GIS)数据和信息的获取、发布和共享已经成为人们的共识。

基于以上两点,开发了一个基于全关系型数据库管理系统和组件对象模型的网络海量影像数据库系统(Internet Mass Image Database, IMIDB),这是国产地理信息系统基础平台 MapGIS 的一个组成部分。

1 IMIDB 的 RDBMS 逻辑设计

1.1 商用关系数据库的选择

图象数据库技术一直致力于解决海量数字图象的有效存储和管理问题。它是数据库技术的继承和发展,一方面,图象数据和文本数据存在着本质的区别,在文本数据领域得以成功应用的传统数据库技术,如果一成不变地照搬到图象数据领域,结果往往是低效,甚至无效;另一方面,传统数据库的许多成果,如 SQL 语言、索引技术等都值得图象数据库借鉴。上述两方面的结合成为目前图象数据库技术发展的主流。

大型后端数据库 SQL Server 7.0 是一个性能优越的数据库系统^[1]。它所构建的服务器大型数据库不仅可以支持字符型、整型、浮点型等常规数据类型,还能很好地支持影像(image)等二进制大对象(BLOB),其最大数据长度可达 2GB,同时,它具有良好的可收缩性、安全性、易用性,以及大容量的管理能力(支持 TB 级数据量)。因此,将它选为 IMIDB 的存储管理平台。

1.2 遥感数据组织^[2]

1.2.1 遥感影像的金字塔结构

为了实现影像数据库的实时缩放显示、快速获得多尺度的影像信息,需要对各种尺度的影像信息进行有效的组织管理。影像数据的多尺度组织和分块组织是解决上述问题的两个关键技术。

影像数据的多尺度组织是用影像金字塔结构来实现的。简单地讲,影像的金字塔结构就是由原始影像数据开始,建立起一系列影像,而各幅影像反映不同尺度下遥感影像的详尽程度。通常原始影像数据反映最详细、最真实的遥感影像,由它逐步生成其他

较低分辨率的影像,以反映原始影像在各尺度下的概貌。这样,在浏览影像数据时,就可以根据当前用户所需的分辨率,提取相应尺度金字塔层的数据。从而实现影像数据的快速浏览。

图 1 所示为单波段图象的金字塔结构。

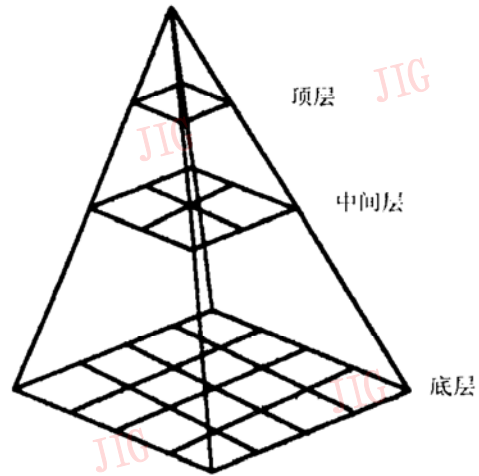


图 1 单波段图象的金字塔结构示意图

通常采用 2 的幂次来创建金字塔结构。具体创建过程为:

- (1) 存储原图象作为底层金字塔,记为第 0 层,其行数和列数分别为 N_{row}^0 和 N_{column}^0 。
- (2) 令层数加 1,将前一层图象按照位置对应关系进行重采样,得到当前层图象。经处理后的图象行列数分别变为 $N_{row}^{i+1} = N_{row}^i / 2$, $N_{column}^{i+1} = N_{column}^i / 2$ 。
- (3) 重复第 2 步,直至达到最顶层。

在不考虑压缩的情况下,建立金字塔结构后的影像数据量与初始影像数据量相比,约增加 1/3。

重采样的方法因具体数据的不同而异。对一般灰度图象,可采用最近邻、双线性或双立方插值算法;而对于编码影像数据,则仅能采用最近邻插值算法。

1.2.2 遥感影像的分块管理

由于遥感影像、航空影像等图象尺寸巨大,一般而言,其行列数远远超出了计算机屏幕显示范围。因此,在影像浏览时,屏幕可见区域显示的只是影像中的一个小矩形区域。这一问题可以用影像数据的分块组织技术来解决,这同时也利于图象的压缩处理,从而高效地完成影像存储,并减少图象读取时间。图象的块可以与关系数据库的记录进行很好的对应,当数据库记录与影像块一一对应时,就可以利用商用数据库管理海量影像数据。

图象块的大小,即图象块的行数和列数,通常取 2 的幂次方。具体分块时需要考虑以下因素:图象的局部相关程度、压缩算法、图象数据类型、图象块缓

冲区的管理算法、用户感兴趣区域的大小、网络的传输单元等.图2为多波段影像数据分块示意图.

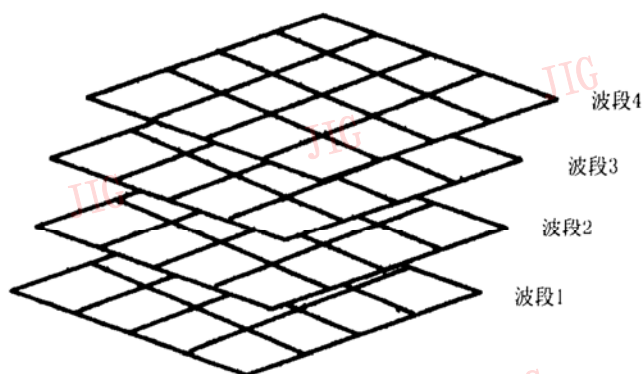


图2 多波段影像数据分块示意

1.2.3 遥感影像的数据库存储

MSI(Multi-Source Image)影像是MapGIS中自定义的一种图象格式,可以支持多种数据源的图象数据,如各种遥感数据、航测数据、航空雷达数据、各种摄影的图象数据,以及通过数值化和网格化的各种地质图、地形图、各种地球物理、地球化学数据和其他专业图象数据.

影像数据库中存储的是由MSI图象分层转化而来的JPEG图象.这主要基于以下考虑:JPEG和GIF是可以在网络上得到很好支持的两种图象格式;减少IMIDB的服务器端处理时间,提高IMIDB的网络响应性能;MapGIS的集成.

之所以采用JPEG压缩标准,主要基于以下考虑:JPEG是一种标准压缩算法,具有较高的压缩比;JPEG格式的图象在Internet上得到了广泛的支持,有利于遥感影像的最终Web发布;在IMIDB中,影像数据并不作进一步的分析、处理,而只是用作显示浏览和制图,因此允许对图象数据进行有损压缩.

2 组件对象模型(COM)技术在IMIDB中的应用

COM及相关技术是目前使用最广泛(Windows平台上)的组件开发模型和技术,是一种基于组件对象平台、具有某种标准通信接口、支持跨语言应用,由软件构件组成的新一代编程方法,具有很强的可配置性、可扩展性、开放性、使用灵活性和二次开发更方便等特点^[3].

海量影像数据库系统的COM组件包括影像数据库管理组件、影像数据库分析、处理组件和影像数据库显示、浏览交互组件3部分.它们都是基于MSI

图象组件构建的.IMIDB采用VC++中的ATL模板库实现了MSI数据管理层、数据处理层和数据显示层组件,共包括22个OLE自动化服务器和2个ActiveX控件.其中,OLE自动化服务器用于矢量数据的管理、影像数据的管理(包括数据库的连接、打开和关闭等操作)、图象文件和其他种类图象格式之间的相互转换、聚类分析、图象处理分析、图象坐标参照等方面;ActiveX控件用于影像数据显示、浏览和矢量数据的提取,以及对带数字高程模型的影像数据进行三维的显示和浏览.

这种COM结构具有面向对象一致性,语言独立性和透明交叉过程互操作性等优点,具有强大的生命力.对IMIDB的各个功能模块进行合理的划分和组织,使系统开发结构明确,便于IMIDB的二次开发,同时也能更好地适应于分布式和Internet应用.

3 IMIDB的原型系统

3.1 IMIDB的体系结构

IMIDB的体系结构如图3所示.其中,Image Database是服务器上存放遥感影像数据的SQL7.0数据库,MetaDatabase是服务器上存放关于遥感影像数据元数据的数据库.IMIDB Admin在后台对Image Database进行管理和维护,包括遥感影像的管理和影像块的数据库输入、预览、删除等,这个模块集中体现了COM技术在IMIDB中的应用.用户首先向Web Server请求一个HTML网页;然后根据返回的结果提出影像请求;ImgAgent在接受到远程客户端请求后,计算出请求的影像ID及显示比例,确认在服务器的影像缓冲区中是否已存在此影像,如果没有,则在Image Database中进行定位,把重采样和分块后的影像合成为一个临时JPEG格式文件,而后把路径传给WebImage Applet供其在客户端显示影像.客户端开辟有一块临时缓冲区,存储一次服务中曾浏览过的影像块.

在原型系统的设计和实现中,借鉴ASP环境的设计,制定了一套影像服务协议,较好地解决了多用户并发访问的问题.

客户端的WebImage Applet实现了IMIDB客户端的数据管理,实现了客户端状态在一定程度上持续性,从而优化了数据查询,减少了网络负载;同时这也是一种跨平台、可扩充的解决方案,容易进一步实现客户端的矢栅集成等功能.

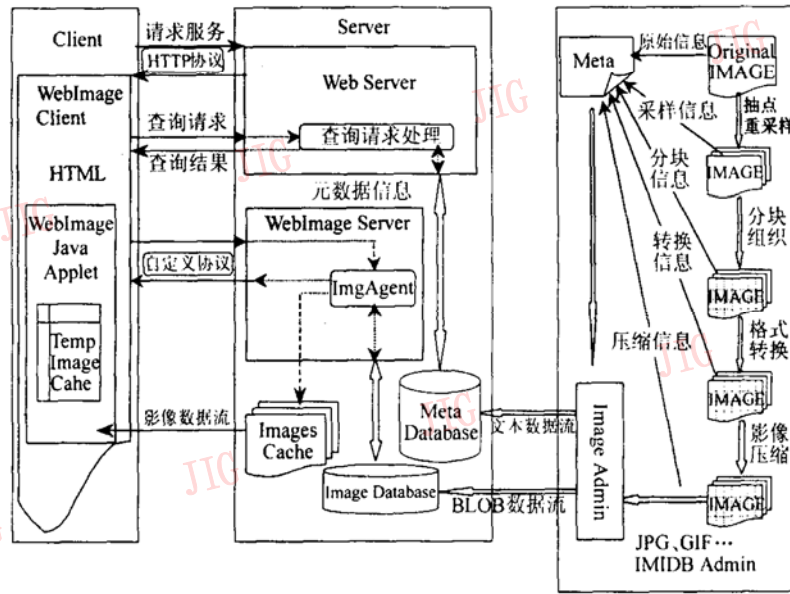


图 3 IMIDB 体系结构

3.2 原型系统运行环境

在一个局域网内构建了如下的 IMIDB 运行环境:

- (1) IMIDB Server 硬件环境: Pentium III 650MHz, 256M 内存, 硬盘 18G. 系统的虚拟内存设置为 1500M;
- (2) IMIDB Server 软件环境: 中文 Windows 2000 Advanced Server, IIS5.0, SQL Server7.0;
- (3) IMIDB Client 硬件环境: 主机 Pentium III 550MHz, 64M 内存, 硬盘 15G;
- (4) IMIDB Client 软件环境: 中文 Windows98, IE5.0;

3.3 原型系统数据组织

在 IMIDB Server 上建立了一个 55.8GB 的影像数据库, 共有 496 个影像文件. 每个影像文件大小为 5000 行 × 7500 列 × 3 个波段, 按 JPEG 方式压缩, 压缩后的数据量为 13.3GB, 压缩比为 4.195.

3.4 测试结果和分析

Web 发布时, 客户端程序 Applet 运行后, 显示服务器端一个 309K 影像块的平均时间为 3s, 对于一个 195K 的影像块, 平均时间为 1s. 由于 WebImg Applet 中实现了客户端的数据管理, 所以请求的影像块如果已经成功请求过, WebImg Applet 就直接调用本地的备份, 其响应时间为 0s.

在多用户并发的情况下, WebImg Server 端由于实现了 Cache 管理, 所以在不同客户对相同的影像块先后提出请求的情况下, 后续请求的响应时间得到了有效的减少.

由于实现了空间索引, ImgAgent 根据当前分

辨率和坐标范围可以直接定位到 ImgDataBase 中相应的图象块, 所以在影像库中的数据量增多的情况下, 服务器的性能不会因为需要遍历而降低.

由原型系统的运行效果和以上的分析, 我们认为 IMIDB 达到了系统设计目的, 可以满足海量遥感影像的 Web 发布的实际需求.

参考文献

- 1 <http://www.microsoft.com/china/sql/./1998-11>[EB/OL].
- 2 Tom Barclay, Robert Eberl, Jim Gray. Microsoft terraServer [EB/OL]. <http://www.terraserver.com/./1998-6>.
- 3 Dale Rogerson. Inside COM[M]. Microsoft Press, 1997.



吴信才 1956 年生, 中国地质大学信息工程学院教授, 博士生导师, 国产优选地理信息系统 MapGIS 主持开发人. 目前研究方向为地理制图学与地理信息工程.



郭玲玲 1976 年生, 中国地质大学信息工程学院硕士毕业, 现为北京大学数字地球室博士生. 研究方向为地理制图学与地理信息工程.



李军 1968 年生, 中国地质大学信息工程学院讲师, 博士. 研究方向为地理制图学与地理信息工程.