

文章编号: 1671-8860(2003)03-0294-07

文献标识码: A

大型无缝影像数据库管理系统的设计与实现^{*}

王 密¹ 龚健雅¹ 李德仁¹

(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: 分析了大型无缝影像数据库的数据管理需求和系统设计原则, 提出了多比例尺、多数据源、分布式异构的海量影像数据管理模型, 并详细地讨论了大型无缝影像数据库管理系统的设计与具体实现技术。

关键词: 影像数据库; 异构数据库; COM/DCOM

中图法分类号: P208; TP392

影像数据作为空间数据的一种载体, 在地理信息领域中的地位变得越来越重要。正射影像本身具有多分辨率、多数据源和数据量巨大等特点, 是国家空间数据基础设施和数字地球的重要组成部分。当前, 国家空间数据基础设施和数字地球建设迫切需要高效地管理多比例尺、多数据源的海量影像数据, 因此如何高效地管理这些正射影像数据, 使之能够为我国的空间数据基础设施和数字地球建设发挥重要的作用, 是一个值得深入研究的问题。关于海量影像数据库的研究从 20 世纪 80 年代末开始初见端倪, 到 20 世纪 90 年代末, 在海量影像数据库系统的研究和实现方面国内外都已作了大量工作, 也推出了一些海量影像数据库管理系统, 如国外 ERDAS 公司的 Image Catalog、ERSI 公司的 ARCSDE 栅格数据管理模块, 国内武汉吉奥信息工程有限公司推出的 GeoImageDB3.0 海量影像数据库管理系统和武汉中地公司推出的 MSIDB 多源影像数据库等, 但是, 从管理模式和管理方式上讲, 这些系统目前还不能满足国家空间基础设施和数字地球建设中海量影像数据的无缝管理和应用需求(王密, 2001)。

随着应用的深入和计算机软、硬件的发展, 数据管理的方式也在由传统的文件方式向数据库方式转变。关系数据库是目前数据库发展的主流, 一些大型商用关系数据库, 如 Oracle、SQL Server、Informix 等已成功用于海量信息的管理领域。这些数据库都具有良好的安全措施、多用户访问、并发控制和数据恢复机制等功能, 并且支持存储复

杂数据类型的能力, 这使得影像数据能够用大型商用数据库来管理。

在基于文件的海量影像数据库管理系统 GeoImageDB3.0 版本的基础上, 本文重点讨论了基于组件技术和大型商用数据库的大型无缝影像数据库系统 GeoImageDB4.0 的管理模型、影像数据压缩以及系统设计和实现技术等。

1 大型无缝影像数据库的数据管理需求和系统设计原则

作为一个以影像数据管理为核心内容的大型海量无缝影像数据库管理系统, 应该满足以下一些需求: ①必须满足多分辨率的特性, 即支持多分辨率影像金字塔的结构; ②必须具有高效的索引和空间查询特性; ③能够集成化地管理多种数据源; ④同时高效地管理影像数据的元数据信息; ⑤必须考虑商用数据库的支持; ⑥所能管理的影像数据量必须是海量的, 应该达到 TB (1TB=10⁶M)级; ⑦采用有效的压缩算法对数据进行压缩; ⑧从数据可视化的角度上讲, 应该以零延迟作为系统追求的目标。

针对以上的数据管理要求, 系统设计应该采用以下的原则。

1) 数据的独立性和共享原则。数据的独立性是指要管理的数据与应用程序之间是分离的, 这样就使得同样的数据可以被不同的应用程序使

* 收稿日期: 2003-02-10。

项目来源: 国家自然科学基金资助项目(40201044); 测绘遥感信息工程国家重点实验室开放研究基金资助项目((01)0102)。

用。共享原则是指对数据的存储和管理不依赖于某个特定的软件系统,从而提高数据存储和管理的可操作性和延续性。

2)开放性原则。开放性原则主要包括数据的开放性和系统的开放性。数据的开放性原则是指系统可以读取多种影像数据格式。系统的开放性是考虑到系统的统一性和实施的可操作性,要使各级系统既相互联系又相对独立,保证系统具有良好的开放性。

3)数据库的可扩展性原则。由于影像数据生产的周期一般来说比较长,随着新数据源的出现和用户需求的增加,数据库中所包含的数据种类可能还会增加,因此就要求系统具有一定的可扩充性。当增加新种类的数据时,系统只要在原来的基础上增加一个对新数据的管理模块和若干数据库表后,就可实现对新数据的存储和管理。

4)数据检索和管理的高效性原则。数据管理和检索是影像数据管理的核心,因此如何高效地实现“知道有什么数据、在什么地方、迅速提取数据并加工成用户所要求的数据”是设计系统时考虑的一个主要因素。

2 多比例尺、多数据源分布式异构海量影像数据管理模型

海量影像数据管理模型对海量影像数据库管理系统所能管理的数据量、管理效率和可伸缩性来讲是至关重要的。在 GeoImageDB 以前版本中,为了解决海量影像数据的存储和管理问题,方涛等提出了分布式影像数据库文件系统 DIDBFS 的管理模式(方涛,1999)。这种管理模式在一定程度上满足了海量影像数据的存储和管理需求,但是随着应用的深入也显示出了明显的不足。其最明显的弱点是它不支持异质网络,只能工作在 Windows 系统平台的局域网范围内,另外系统的数据安全性、完整性、多用户访问和多数据源支持等方面也满足不了当前用户的需求。

作者在总结分布式影像数据库文件系统 DIDBFS 管理模式不足的基础上,结合当前海量影像数据库的管理需求,提出了多比例尺、多数据源影像数据库模型。该模型仍然以金字塔的数据结构为基础,但是对影像金字塔的含义进行了扩充。影像金字塔的建立可以采用以下两种方式。

1)多分辨率的数据源自动构建金字塔。从目前我国空间数据基础设施的建设来看,由于生产

的数据本身就是多分辨率的,因此可以直接建立不同的图像工程将数据入库,通过超图像工程自动构建金字塔。

2)影像数据抽取构建金字塔。对于只生产了基础层的影像数据,为了提高调度效率,其上层可以根据相应的比例尺从下层抽取数据来构建金字塔,这是目前多数系统所采用的方式。

对于多尺度的金字塔影像数据库,在大型商用数据库中一个比例尺可以作为数据库的一张或多张表,一块数据存贮为一条记录,这样建立起多比例尺空间数据库。应用时根据不同的显示比例尺或制图比例尺调用与之最相近的那一级比例尺数据库的数据,这也比较符合我国目前多分辨率空间数据基础设施的建设需求。多比例尺空间数据库可以根据需要和数据源的情况构建。图1表示了本文所研究的多比例尺、多数据源影像数据管理模型。

由图1的影像数据管理模型可知,由于每一个图像工程具有单独的空间索引,因此可以作为一个单独的子库存储在不同的数据库管理系统中,然后通过超图像工程将各个子库逻辑地联系起来,超图像工程就相当于整个影像数据库的全局逻辑视图。这样用户在访问超图像工程时,不必关心每个图像工程的具体存放位置和所采用的数据库管理系统,从而使各个子库对用户来说是透明的,并且每个子库可以存储在网络的不同节点上,形成分布式的异构存储模式。用户可以单独地访问每一个子库,也可以通过超图像工程透明地访问各个子库,保证整个影像数据库系统具有很好的开放性和可扩充性。图2表示了这种分布式的异构存储模式。

这种异构的影像数据库管理模式可以带来以下一些优点。

1)由于影像数据管理的基础平台是大型商用数据库,因而它具有许多文件系统无法具备的优点,如良好的安全性、多用户访问特性、数据的一致性。

2)对于大数据量的建库工作可以划分成不同的图像工程即子库,每个子库可以分别建立在不同的数据库服务器上,数据入库工作可以并行进行,提高了建库的效率。

3)对于生产单位,每个部门生产的数据的精度和数据源的类型不同,他们需要对自己生产的数据进行管理。由于每个部门的具体条件不同,因此所采用的数据库管理系统也可能不同,但是主管部门又需要将一个单位的数据进行统一管理,

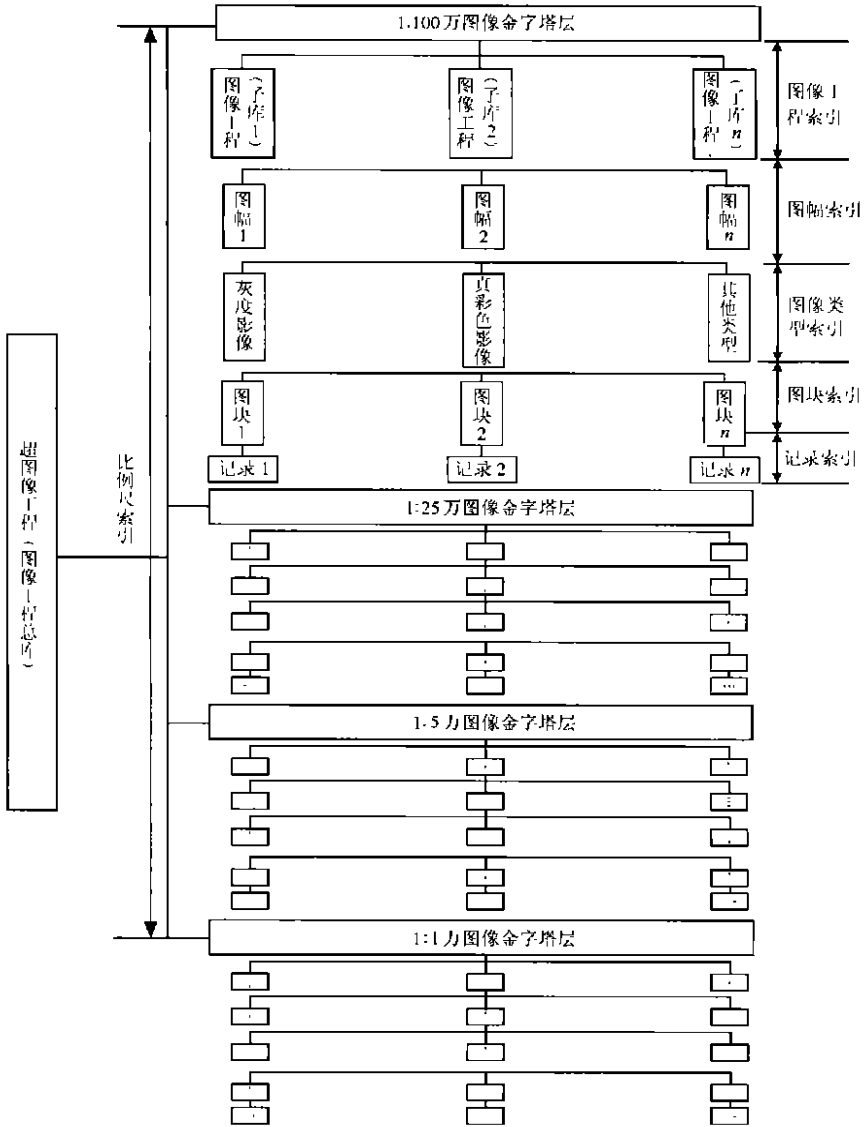


图 1 省级多比例尺、多数据源影像数据管理模型

Fig. 1 Multi-scale and Multi-resource Image Database Model in Province Level

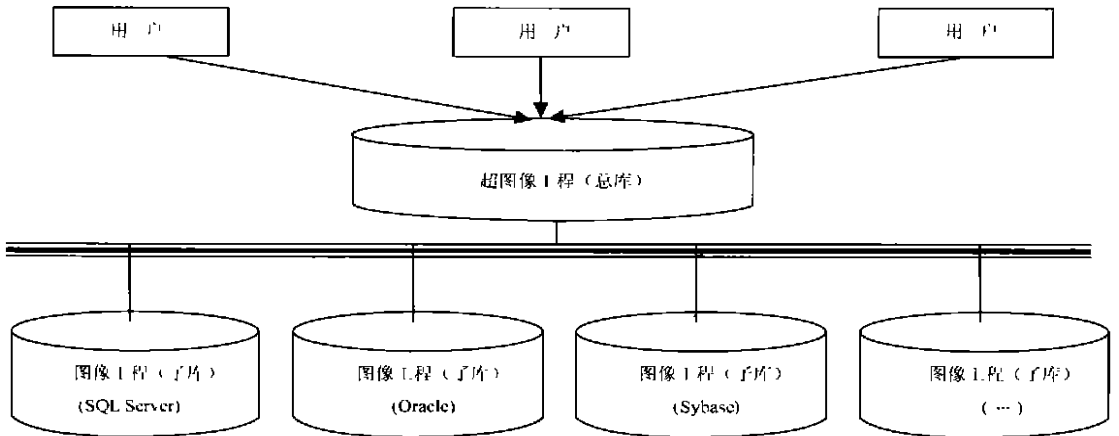


图 2 分布式的异构存储模式

Fig. 2 Distributed Heterogeneous Storage Mode

因此这种异构的管理方式在局域网范围内可以满足这种需求。

4)对异构操作系统的支持。目前有的数据库管理系统运行在 UNIX 操作系统上,有的运行在 Windows NT 操作系统上,这种异构的管理方式可以将不同操作系统的数据库统一起来。

5)系统具有良好的伸缩性,数据量可以无限地扩充,对于新生产的数据只需单独地建立子库,然后通过超图像工程进行连接即可。

6)数据分布式的存储有利于负载平衡,提高数据管理和访问的效率。

7)系统的需求环境比较宽松,可以利用已有的系统资源(如各种数据库软件资源)。

8)提高数据库访问的灵活性,用户既可以访问单独的图像工程子库,又可以访问整个金字塔影像数据库。

3 影像数据库的数据压缩

为了减少影像数据在存储和处理时所占用的存储空间,提高数据的传输效率,对于海量的影像数据进行数据压缩是必要的。影像压缩是指在基

本不影响影像视觉和使用效果的情况下,采用一定的技术手段和压缩算法,将影像文件的尺寸压缩到原文件大小的几分之一或几十分之一,以节省存储空间和方便传输。压缩技术按是否有信息损失分为无损压缩技术和有损压缩技术。针对海量影像数据压缩的需求,分析现有数据压缩技术,我们认为传统的压缩算法已经成了定式,发展潜力不大;而基于小波变换的编码、分形编码,基于内容的压缩编码和基于视觉熵理论的编码等则是目前的研究方向。目前,小波技术在影像数据压缩领域应用得非常成功,相继出现了一些商业化的小波压缩软件,如美国 LizardTech 公司的 MrSID 和澳大利亚 ERMapper 公司的 ECW 等。

在本文研究的系统中,影像数据压缩采用了 ER Mapper 中的增强压缩小波(ECW)技术。它利用了离散小波变换(DWT)技术对图像进行压缩,通过局部转换,使得图像内部任何一部分都具有一致的分辨率和很好的图像质量。ERMapper 压缩对于灰度图像采用 10 :1 ~ 15 :1 的压缩比,彩色图像采用 25 :1 ~ 50 :1 的压缩比,可获得质量非常高的压缩图像。用户可以调整压缩比以获得理想的结果。

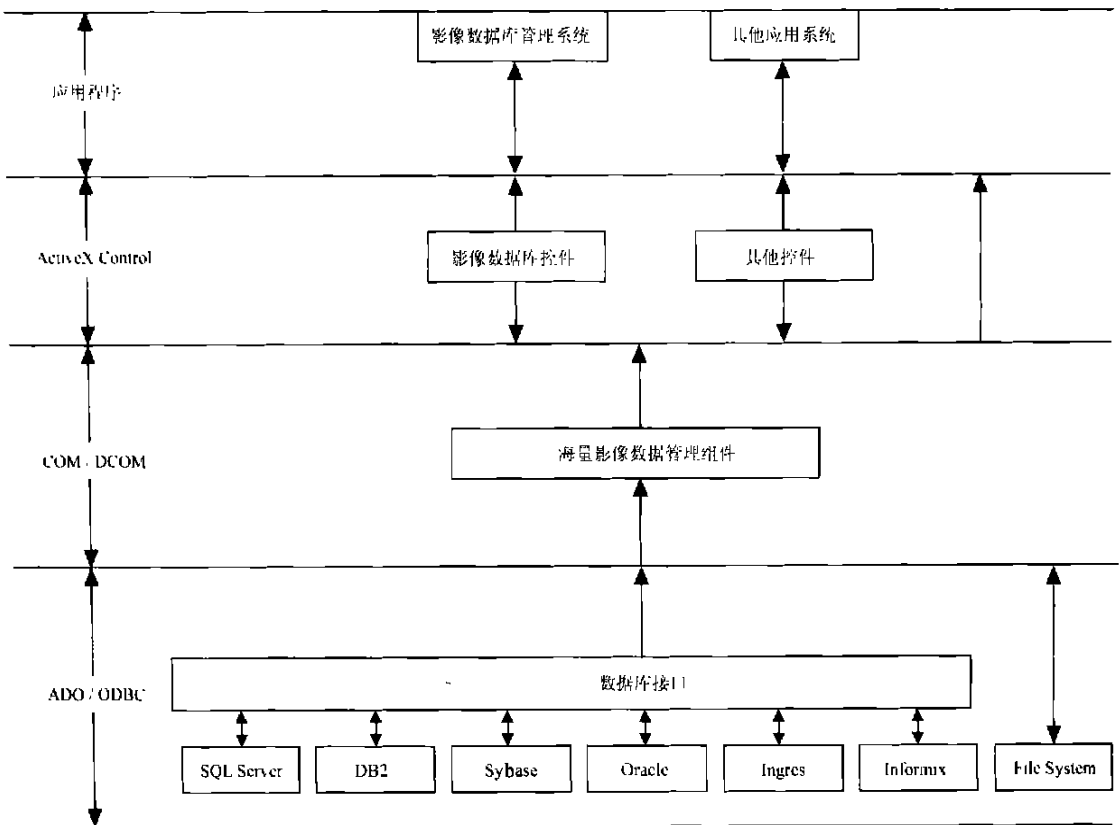


图3 GeoImageDB4.0 的软件体系结构

Fig. 3 Architecture of GeoImageDB4.0

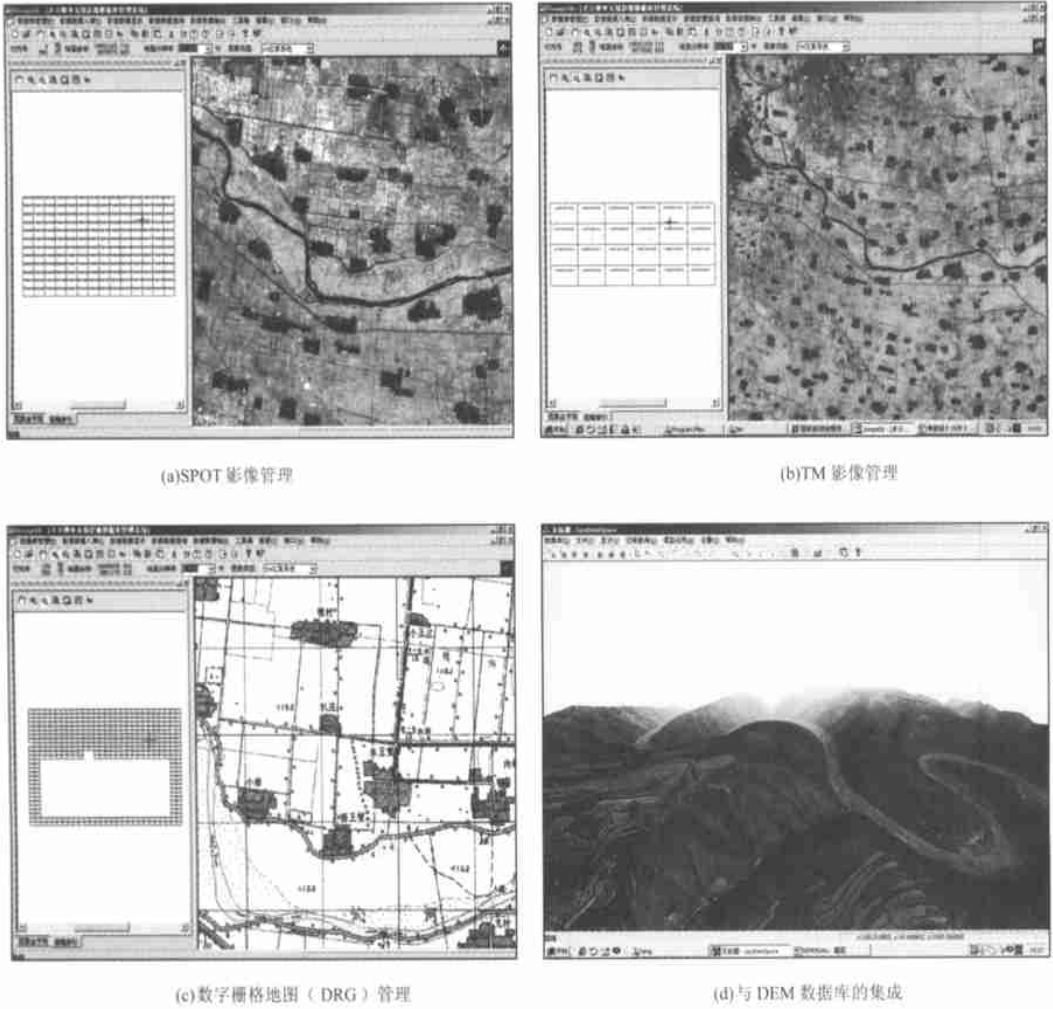


图4 GeoImageDB4.0应用实例
 Fig. 4 Instances of GeoImageDB4.0

ECW 压缩技术具有如下特点。

- 1) 开放的标准。ECW 免费提供各种开发环境的图像压缩/解压开发工具库。
- 2) 广泛的程序支持。ECW 为 Web 浏览器、GIS、CAD 和图像处理等软件提供有效的插件,更多的插件随时由第三方的开发者使用开放的标准 ECW 开发工具库加入。
- 3) 免费的压缩器。ECW 提供小于 500MB 图像的免费压缩和解压。

4 海量影像数据库管理系统的实现技术

海量影像数据库的无缝实现技术已经进行过详细讨论^[5,7], 本文重点讨论基于组件对象模型 (COM) 的 GeoImageDB4.0 软件实现技术。

GeoImageDB4.0 是在 Microsoft Windows 操作系统平台上开发的, 采用 COM/DCOM 标准, 将影像

数据分成小的数据块作为记录存放到关系表中。下面较详细地介绍 GeoImageDB 4.0 版本的实现方法。图 3 所示为 GeoImageDB4.0 的软件体系结构。

在图 3 中, 最底层是用 C++ 语言按照 COM 规范编写的海量影像数据管理组件, 它的主要功能是负责海量影像数据的调度、索引、压缩/解压和实时处理等, 这一组件通过 ADO 或 ODBC 与目前流行的各种商用数据库 (SQL Server、DB2、Sybase、Oracle、Informix、Ingres) 链接, 操作异构的数据库管理系统。另外, 也可以通过自己开发的文件接口来访问 3.0 版本的文件管理系统。在海量影像数据管理组件的上一层是 ActiveX 控件, 负责影像的显示、漫游等可视化操作和一些查询等功能。控件和组件并不是应用系统, 只是一些主要功能的集合体。组件提供底层最基本的功能操作, 控件提供了一个可视化的窗口框架, 最上层是在组件和控件基础上构建的应用程序, 用来实现影像数

数据库的各种功能。用户也可以根据控件和组件定制自己的应用系统。另外,用户还可以很方便地通过组件与其他系统进行集成。

5 研究实例与结论

笔者在开发海量无缝影像数据库管理系统 GeoImageDB 4.0 时,采用并实现了前面所述的多比例尺、多数据源异构影像数据库管理模式,系统以目前流行的商用数据库 Oracle、SQL Server 数据库管理系统为数据库平台(目前实现了这两个数据库接口,其他数据库接口以后可以扩充),同时也支持文件管理系统,与 GeoImageDB 3.0 版本完全兼容。利用 Visual C++ 6.0 提供的面向对象 C++ 语言的集成开发环境,基于组件对象模型(COM)的全组件、控件式的开发方式,支持影像数据在局域网范围内透明、分布式、异构存储与管理。GeoImageDB4.0 实现了多比例尺、多数据源分布式异构海量影像数据管理,能够高效地管理多分辨率、多数据源海量影像数据,解决了目前我国的数字中国、数字化省、数字城市、数字小区等工程建设中海量影像数据库建设的技术问题,为这些工程建设奠定了一定的基础。作者在黑龙江测绘局、山西测绘局和国家基础地理信息中心等单位进行了 GeoImageDB4.0 系统应用测试,所用的影像类型有彩色航空影像、数字栅格地图、单波段遥感影像、多波段彩色合成影像,所用的数据库平台为 SQL Server 和 Oracle8i 商用数据库管理系统。测试结果表明,无论是在系统的工作效率还是在海量数据管理的可伸缩性方面,都明显优于 GeoImageDB 以前的版本。图 4 显示了 GeoImageDB4.0 影像数据管理的一个实例。

参 考 文 献

1 Gore A. The Digital Earth; Understanding Our Planet in the

- 21st Century. <http://www.digitalearth.gov/speech.html>, 1998
- 2 Batory D J, Barnett J G. An Extensible Database Management System. IEEE Transactions on Software Engineering, 1988, 14(11): 358~363
- 3 John A T, Anthony J T, Mark A K. The Handling of Large Digital Orthophoto Raster Database in GIS. ISPRS, 1994, 30(4): 457~459.
- 4 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
- 5 方 涛, 李德仁, 龚健雅, 等. GeoImageDB 多分辨率无缝影像数据库系统的开发与实现. 武汉测绘科技大学学报, 1999, 24(3): 189~193
- 6 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
- 7 王 密, 龚健雅, 李德仁. 大型遥感影像数据库空间无缝的数据组织. 武汉大学学报·信息科学版, 2001, 26(5): 419~424
- 8 李德仁, 龚健雅, 朱欣焰, 等. 我国地球空间数据框架的设计思想与技术路线. 武汉测绘科技大学学报, 1998, 23(4): 297~303
- 9 龚健雅, 朱欣焰, 朱 庆, 等. 面向对象集成化空间数据库管理系统的设计与实现. 武汉测绘科技大学学报, 2000, 25(4): 289~293
- 10 李德仁. 信息高速公路、空间数据基础设施与数字地球. 测绘学报, 1999, 28(1): 1~5
- 11 龚健雅. 地理信息系统基础. 北京: 科学出版社, 2001

第一作者简介: 王密, 博士。主要研究方向是海量空间数据管理、无缝影像数据库系统的设计、研制及立体无缝影像数据库和可量测虚拟现实。主要代表成果为无缝影像数据库管理系统 GeoImageDB。目前已经发表论文 10 余篇。

E-mail: wangmi@x263.net

Design and Implementation of Large-Scale Image Database Management System

WANG Mi¹ GONG Jianya¹ LI Deren¹

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

Abstract: At present, the construction of National Spatial Data Infrastructure (NSDI) and Digital Earth imminently needs efficient management of multi-scale and multi-resource mass image data. This paper

gives systematic and deep researches on distributed heterogeneous management mode of super-image \1 project and image-project based on commercial database management system (such as ORACLE, SQL Server) and Component Object Modal (COM) after the characteristics of image data and its management requirement is comprehensively analyzed. The design and implementation method for large-scale seamless image database management system is also described.

Key words: image database; heterogeneous database; COM /DCOM

About the first author: WANG Mi Ph. D. His major research interests are mass spatial data management, design and development of seamless image database, stereo seamless image database and measurable virtual reality. His typical achievements are seamless image database management system—GeolImageDB.

E-mail: wangmi@x263.net

(责任编辑: 涓涓)

(上接第 271 页)

the main direction and sequentially lead to the initial orientation parameters values. Then, more precise and stable values are found by an optimizing algorithm, based on conditional least squares adjustment with unknown parameters. During the iterative process in this algorithm, orientation parameters of the image are recalculated. Last, these orientation parameters are used to reconstruct textures landscape of the buildings by combining the vector and texture data in image.

Key words: vanishing point; camera calibration; least squares

About the first author: ZHANG Zuxun, professor, ph. D supervisor, member of Euro-Asia International Academy of Sciences. His research field includes photogrammetry and remote sensing. His representative achievements include a fully digital mapping system—VirtuoZo and a literature—“Digital Photogrammetry”.

E-mail: zxzhang@supresoft.com.cn

(责任编辑: 涓涓)