

Research on Management and Sharing of Large-Scale Image Based on ArcGIS

Liang Wang, Jun Tang*, Jinli Miao

¹College of Earth Sciences, Yangtze University, Wuhan

²China Geological Survey Development Research Centre, Beijing

Email: wangliang1623@163.com, tang0262@sina.com

Received: Jan. 2nd, 2014; revised: Feb. 1st, 2014; accepted: Feb. 8th, 2014

Copyright © 2014 Liang Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. In accordance of the Creative Commons Attribution License all Copyrights © 2014 are reserved for Hans and the owner of the intellectual property Liang Wang et al. All Copyright © 2014 are guarded by law and by Hans as a guardian.

Abstract: This paper introduces the model of traditional image data management and its existing problems, and the release and sharing of ArcGIS image management and image data are studied, including ArcGIS image management in different periods of the model and their respective characteristics, slicing and dynamic imaging services. According to the experimental verification, mosaic dataset can manage massive image data very well and make the sharing more convenient.

Keywords: Mosaic Dataset; ArcGIS; Image Data; Imaging Services

基于 ArcGIS 的大规模影像数据管理和共享的研究

王亮¹, 汤军^{1*}, 缪谨励²

¹长江大学地球科学学院, 武汉

²中国地质调查局发展研究中心, 北京

Email: wangliang1623@163.com, tang0262@sina.com

收稿日期: 2014 年 1 月 2 日; 修回日期: 2014 年 2 月 1 日; 录用日期: 2014 年 2 月 8 日

摘要: 本文介绍了传统影像数据管理的模式及存在的问题, 并对 ArcGIS 的影像管理模式和影像数据的发布与共享进行研究, 包括 ArcGIS 不同时期影像管理的模型和各自特点、切片和动态影像服务。实验验证, 镶嵌数据集能够很好的管理大规模影像数据, 让共享更加便捷。

关键词: 镶嵌数据集; ArcGIS; 影像数据; 影像服务

1. 引言

随着空间信息技术和我国卫星遥感技术的快速发展, 采集高分辨率的空间影像数据越来越容易。但是影像数据的分辨率越高意味着需要存储、管理的影像数据量越大, 它们之间是非线性增长的量化关系^[1]。对如此庞大、复杂的影像数据集, 进行快速处理、高效的存储、管理和共享, 缩短影像从获取到使用的

时间, 为公众、政府和企业提供满足需要的数据服务, 就成为一个需要迫切解决的问题。大规模影像数据存储管理研究最早可以追溯到 20 世纪 90 年代, 微软开发的 Microsoft TerraServer 系统, 将美国和俄罗斯数十年来拍摄的卫星数据放到 Internet; 在海量影像数据管理和处理方面, 具有里程碑意义的是 Google Earth, 它可以整合各类影像数据, 向用户提供高分辨率遥感影像公众服务。国内在该方面的软件也是异军突起,

*通讯作者。

如国遥新天地的 EV-Globe, 灵图的 VRMap, 武大吉奥的 GeoGlobe 等等。相比较之下, 国内的这些软件则主要侧重于海量影像数据的发布, 对于管理和处理方面的工作还不成熟, 不能够满足多用户的影像需求^[1]。本文基于 ESRI 公司的 ArcGIS 软件对大规模影像数据管理和共享的研究, 采用镶嵌数据集模式管理影像数据, 再以影像服务的形式发布共享, 让影像数据管理更加高效、共享更加便捷。

2. 传统影像数据管理的模式及问题

遥感影像数据的组成包括非结构化的像素数据和结构化的元数据信息。影像数据的管理模式主要可以分为以下三类。

2.1. 基于文件系统管理的方式

该影像管理模式将影像数据文件和元数据文件全部以文件形式存储在文件服务器上, 然后按照建立的影像编目规则, 以文件的形式来组织管理。数据量较小的情况, 这种模式操作简单, 效率较高。但影像数据的并发操作、快速检索以及空间数据的分布式存储和管理难以实现, 一旦文件服务器工作量特别大而崩溃时将导致整个系统的瘫痪, 从而造成不可弥补的损失^[1]。美国宇航局(NASA) EOS 和 World Wind, 谷歌公司的 Google Maps 和 Google Earth, 美国商业卫星遥感公司 Geoeye, 日本的 Earth Simulator 都属于基于文件的文件系统管理方式^[2]。

2.2. 基于数据库系统管理方式

这种管理模式主要有两种形式: 一种是基于关系数据库的进行扩展, 它通过增加空间关系、空间的数据类型和操作来完成对影像数据的存储和管理; 另一种是基于空间数据引擎, 它是通过构建数据库管理中间件, 提供丰富而高效的空间数据访问接口, 来完成对影像数据的存储和管理。这两种形式都充分利用了数据库管理系统技术, 将遥感影像数据和属性数据同时存储在数据库中, 实现对元数据和影像数据的集中管理, 且具有多用户并发访问性和良好的数据安全性。不足之处在于都采用了 DBMS 来存储非结构化的影像数据, 使得存储的效率很低, 使数据库系统难以管理大规模的影像数据^[3]。微软的 TerraServer 和 Bing Maps 属于基于 BLOB 数据类型的关系数据库管理^[2]。

2.3. 基于文件与数据库结合的方式

该影像管理模式是将影像数据有关的属性数据和元数据存放到数据库系统中, 大量的非结构化的影像数据以文件的形式存放, 这种模式结合了文件系统和数据库系统的优势, 既能存储和管理大规模影像数据, 又便于数据的查询和使用。但相对于数据库系统存储管理模式, 该模式不具备多用户的并发访问特性^[3]。欧洲太空局(ESA)数据中心, 天地图, 中国资源卫星应用中心, 国家卫星气象中心, 国家海洋应用中心都属于文件和商业关系数据库混合方式^[2]。

3. 基于 ArcGIS 的影像管理

为了避免传统管理模式存在的问题同时又满足大规模影像数据更好的管理和共享, 本文研究了基于 ArcGIS 影像管理和共享解决方案。

ArcGIS 在不同阶段提出了不同的影像管理方案, ArcGIS 8X 利用栅格数据集(Raster Data Set)管理影像, ArcGIS 9X 利用栅格目录(Raster Catalog)进行存储, ArcGIS 10X 增加了专业的影像管理信息模型镶嵌数据集(Mosaic Dataset)^[4]。本文主要针对镶嵌数据集管理影像数据进行研究。

3.1. ArcGIS 影像数据管理

镶嵌数据集是管理和处理大规模影像数据的数据管理模型, 如图 1 所示, 采用文件系统加数据库系统的存储模式。通过数据库管理影像元数据文件及编目信息, 不会拷贝或者改变原始的影像数据, 它仍存放在文件系统中, 因此可快速构建镶嵌数据集管理大规模影像数据。镶嵌数据集的实时处理能力可实现多种类型影像产品的实时构建, 快速无冗余, 节省工作时间与磁盘空间。通过栅格类型(raster type)提供对多元传感器的支持^[5]。

3.2. 镶嵌数据集的特点

镶嵌数据集除了拥有完善的管理模型, 还具有以下特点^[5]:

1) 动态镶嵌

影像添加到镶嵌数据集之后, 通过动态镶嵌技术, 可以近乎实时的速度完成传统影像管理中影像镶嵌过程, 并可达到相同的效果。影像的空间参考也将动

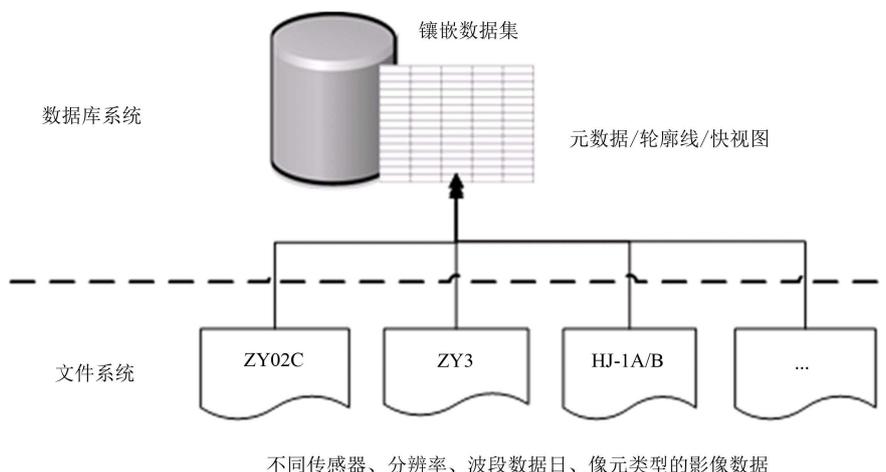


Figure 1. Mosaic dataset structure
图 1. 镶嵌数据集结构

态更改为与镶嵌数据集一致的参考，原始影像并没有发生更改，所有的过程都是动态虚拟的。

2) 实时处理

镶嵌数据集中执行的操作作为实时处理过程，它能够让用户以近乎实时的速度得到影像的处理结果，无需等待时间。通过实时处理技术，可以实现遥感影像的几何处理和辐射处理，不会产生中间数据，减少冗余数据的生成。

通过使用动态镶嵌和实时处理技术，用镶嵌数据集建立影像库的速度会得到质的提升，也更有利于影像数据的发布与共享。

4. 影像数据的发布与共享

大规模影像数据通过传统的硬拷贝、FTP/HTTP 下载等方式进行共享，专业影像分析人员可以将影像数据下载到本地，利用专业软件进行分析。但是，这种方式并不能满足用户所有的需求。有些用户只需要将影像作为底图，传统的模式便不能达到这一要求。如何满足各种用户不同的需求是对大规模影像数据共享技术的挑战。

ArcGIS 影像解决方案以服务的形式，提供了灵活的数据共享方式。通过 ArcGIS Server 和其扩展 Image Extension，用户可以轻松实现海量影像的共享，如图 2 所示，基于 ArcGIS Server 的海量影像共享模式。通过 ArcGIS Server 发布的切片影像服务和动态影像服务，可以为用户提供企业级的影像数据访问能力^[6]。

1) 切片影像服务

ArcGIS Server 以影像切片服务共享的方式，为用户提供了即拿即用的服务，能够极快的显示影像，适合作为背景底图使用^[7]。

切片影像服务具有如下优势：

- 访问速度快，可获得类似于 ArcGIS Online、Bing Maps 和 Google Maps 的极速体验；
- 占用服务器资源少，影像服务单台服务器即可支持大量用户，通过集群方式可应对日益增长的访问量；
- ArcGIS 提供的切片工具，可以使用标准方案或自定义方案和集群处理能力，轻松快速创建影像切片。

2) 动态影像服务

Image Extension 扩展了 ArcGIS Server 影像服务的能力，使其能够将镶嵌数据集发布为动态影像服务。影像动态服务因为需要进行服务器动态处理，速度略逊一筹，但是提供了强大的服务器处理能力、快速的查询检索能力和可控的原始数据下载能力。它不仅适合公众用户作为背景底图使用，而且专业用户可以使用影像动态服务进行分析和处理。通过 Image Extension 发布的影像服务，可以快速访问镶嵌数据集和其中管理的影像数据，从而提升了影像资源的价值^[8-10]。

动态影像服务具有如下优势：

- 通过单一服务即可快速地访问海量影像数据；
- 编目管理的影像能够像镶嵌影像一样进行可视化和分析；
- 通过实时处理技术，从单一数据源创建多种影像

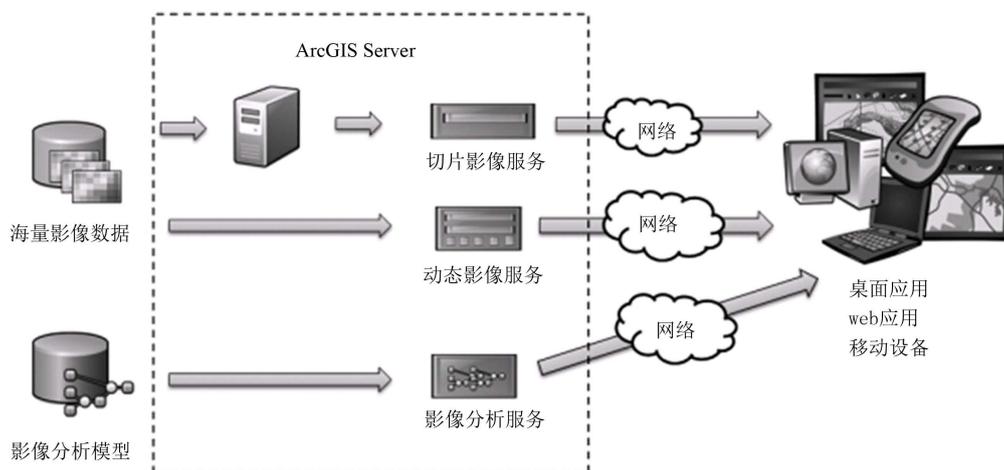


Figure 2. Manage massive image sharing model based on ArcGIS Server
图 2. 基于 ArcGIS Server 的海量影像管理共享模式

产品的服务，在提供增值服务的同减少庞大的数据冗余；

- 保留了影像级别的元数据信息，客户端并能够快速进行查询和检索；
- 能够下载原始影像或裁切下载某一区域的影像，以便于影像分发。

通过 ArcGIS 提供的切片影像服务和动态影像服务，影像数据可以通过 Web 服务的方式进行共享。用户不需要安装客户端组件和程序，即可以通过网络快速访问到共享的影像数据，轻松和现有应用进行集成，解决传统影像共享模式中的问题^[11-14]。

5. 实验与结论

为了验证 ArcGIS 对大规模影像数据的管理和共享模式，本文从美国地质调查局网站获取到南极洲整个大陆 Landsat-7 的影像拼接数据 654 幅，数据量达 200 GB。按照研究思路，首先用 oracle 数据库和 ArcGIS 镶嵌数据集分别管理这些数据，建立影像库，然后利用 ArcGIS desktop 和 Server 分别对镶嵌数据集建快视图和缓存切片，并发布成影像服务，比较两种影像服务的显示速度。实验结果表明：

1) 对比两种方式入库效率发现，使用镶嵌数据集管理的影像数据能够快速建库，没有产生冗余数据，而且当原始数据变化时，可以自动同步，管理起来更加方便。而使用 oracle 数据库管理的影像数据入库时间过长，当数据量非常大的时候，会产生中断现象，不宜实时更新(表 1)^[15]。

Table 1. Image storage time comparison
表 1. 影像入库时间对比

管理方式	入库时间
Oracle 数据库	11.5 小时
Arcgis 镶嵌数据集	200 秒

Table 2. Comparison of image release
表 2. 影像发布方式对比

发布方式	生成时间	生成数据量	加载时间
影像切片	12 小时	40 (GB)	0.1 秒
影像快视图	2 小时	10 (GB)	0.5 秒

2) 利用 ArcGIS Server10.1 的集群功能，运行多个切片工具实例对原始影像数据生缓存切片，可以极大的提升切片生成的效率，最大化的利用服务器的资源。生成的多级影像缓存切片，能够在桌面、Web 和移动应用程序中快速浏览，但是生成时间过长。而 ArcGIS desktop 生成的镶嵌数据集的快视图可以极大的提高效率，不足之处是显示相对影像切片稍慢，可以满足实时性要求不高的影像服务(表 2)。

3) 通过动态影像服务的发布，可以轻松实现影像数据的查询浏览应用，而且实现了传统模式的影像数据分发下载，使影像数据共享更加便利。

发布后的影像效果图如图 3 所示。

6. 结语

本文在传统影像管理模式不足的基础上，研究使用 ArcGIS 的镶嵌数据集模型管理大规模影像数据，

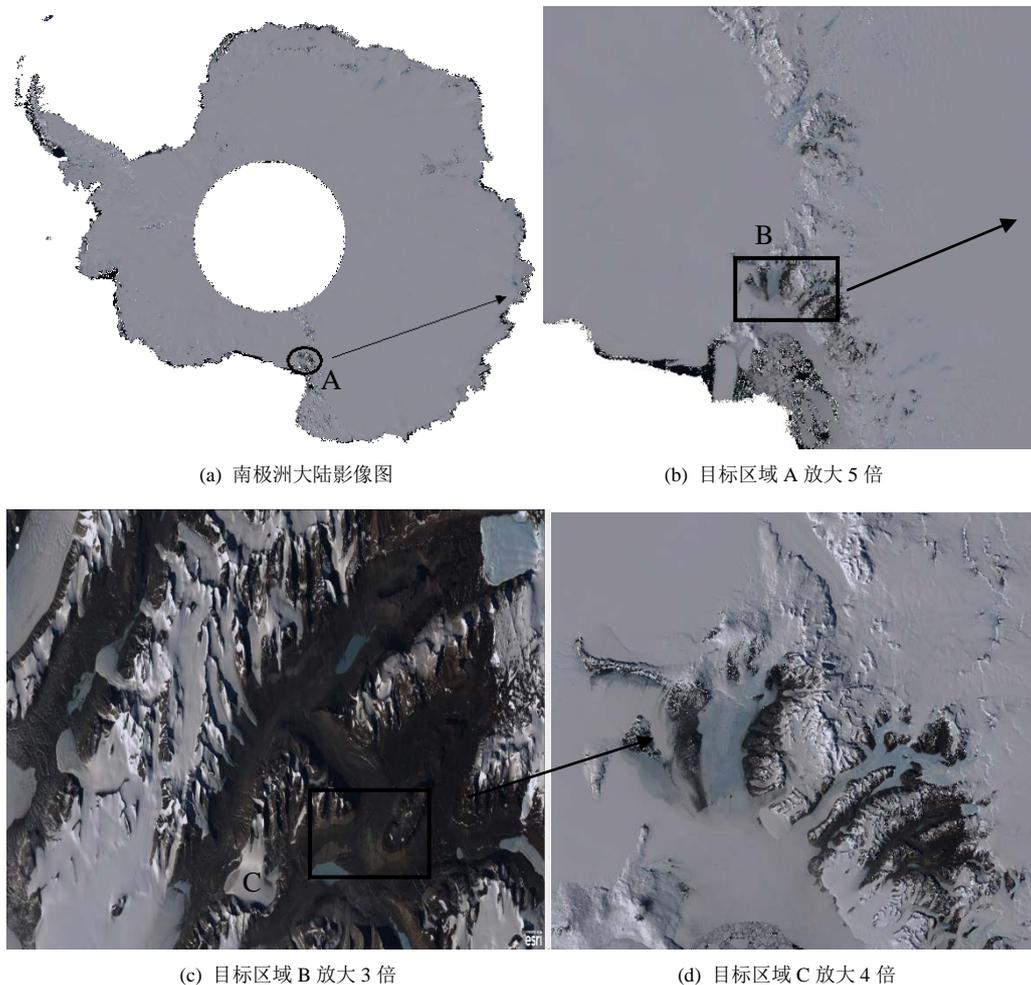


Figure 3. Image renderings
图 3. 影像效果图

充分体现了文件系统和数据库系统各自优点。利用 ArcGIS Server 以影像服务形式发布,使影像更好的共享,发挥其应有的作用。

但是随着影像数据量的增大,达到 PB 级,镶嵌数据集是否能继续发挥优势;随着用户并发数达到一定规模,文件系统的管理模式是否会出现瓶颈;随着计算机技术的发展,影像数据的切片处理时间和在线分析处理能力是否能得到更大的提升,这些仍需要研究和改进。

参考文献 (References)

[1] 邹雪冰 (2012) 基于 Oracle 的海量空间影像数据管理. *科技创新导报*, **10**, 85-89.
[2] 史少维, 刘云广 (2013) 基于 Oracle 遥感影像库的构建与实现. *城市勘察*, **2**, 17-22.

[3] 穆增光 (2013) 基于 ArcGIS 的综合影像数据库系统研究. *测绘技术装备*, **1**, 7-10.
[4] 霍亮, 刘桂红 (2010) T 级影像数据管理平台的设计与实现. *测绘科学*, **35**, 87-89.
[5] 董杰, 黄炎, 董平 (2012) 基于镶嵌数据集的大规模遥感影像管理研究. *红外*, **10**, 30-32.
[6] Esri 中国信息技术有限公司 (2011) ESRI 海量影像管理与共享解决方案白皮书. Esri 中国信息技术有限公司, 北京.
[7] Esri 中国信息技术有限公司 (2012) 解密 ArcGIS 影像管理和共享技术. Esri 中国信息技术有限公司, 北京.
[8] Esri 中国信息技术有限公司 (2012) 开启镶嵌数据集管理影像的新航程. Esri 中国信息技术有限公司, 北京.
[9] ArcGIS 影像技术之镶嵌数据集概览. (2012) http://blog.sina.com.cn/s/blog_790697610100zycp.html
[10] 吕雪锋, 程承旗, 龚健雅, 关丽 (2011) 海量遥感数据存储管理技术综述. *中国科学*, **41**, 1561-1573.
[11] 周成虎, 欧阳, 李增元 (2008) 我国遥感数据的集成与共享研究. *中国工程科学*, **6**, 51-55.
[12] 戴芹, 刘建波, 刘士彬 (2008) 海量卫星遥感数据共享的关键技术. *计算机工程*, **6**, 283-285.
[13] Sample, J.T. and Loup, E. (2010) *Tile-base geospatial information system: Principle and practices*. Springer, New York, 23-200.

- [14] Gong, J.Y., Xiang, L.G., Chen, J., et al. (2010) Multi-source geospatial information integration and sharing in Virtual Globes. *Science China Technological Sciences*, **1**, 1-6.
- [15] Wu, X., Guo, J., Wallace, J., et al. (2009) Evaluation of CBERS image data: Geometric and radiometric aspects. *Innovations in Remote Sensing and Photogrammetry*, **2**, 91-103.