

# 基于 COG 的遥感影像地图服务快速构建方法

陈昱行,盛辉,刘善伟,何亚文

(中国石油大学(华东)海洋与空间信息学院,山东 青岛 266580)

摘要:如今主流的地图服务软件中影像图层的发布需要手动指定参数,不利于影像图层自动化发布。 本文搭建私有对象存储服务管理遥感数据,利用 COG(cloud optimized GeoTIFF)支持局部访问的特性, 优化影像瓦片动态创建速度,实现了影像自动发布的瓦片地图服务,改进了瓦片唯一标识码生成算法, 提高了瓦片缓存的读写性能。在实例应用中,本文实现的遥感影像地图服务与 GeoServer 相比,在瓦片 动态创建中性能相近,在处理已缓存瓦片时更有优势,具有应用前景。

关键词: 遥感影像; 地图服务; 瓦片金字塔 中图分类号: TP751.1 文献标识码: A DOI: 10.11759/hykx20210208002

随着遥感对地观测技术的发展,遥感数据在国土 资源、海洋环境、气象等各行业中被广泛应用,成为 时空分析的重要数据来源。为更好地管理、共享这些 遥感数据,国内研究者基于分布式技术做了大量的基 础研究<sup>[1-4]</sup>,同时,如何能够在数据持续更新情况下, 为不同领域的用户提供及时的、便捷的遥感数据共享 服务<sup>[1]</sup>,成为遥感应用领域的研究热点。

遥感数据共享包括数据共享和数据预览。Web 端 的数据在线预览依赖于地图瓦片服务<sup>[5]</sup>,分为预先渲 染和动态创建两种方式:预先渲染需要在瓦片生成完 毕后才能提供服务,具有更快的响应速度,是构建地 图瓦片服务的主要方式<sup>[6-8]</sup>;动态创建方式在节约计算 资源方面更有优势<sup>[9]</sup>,适合在数据持续更新情况下使 用。目前,国际上知名的地图服务商业化软件 ArcGIS Server 支持使用镶嵌数据集(mosaic dataset)插件实现影 像瓦片的动态创建<sup>[9]</sup>;主流的开源地图服务软件 GeoServer 支持在不使用 GWC(geowebcache)的情况下 对遥感影像直接提供网络地图服务(web map service, WMS)服务。对于瓦片动态创建速度优化,可以在原始 数据内部增加 Tiling 和 Overview<sup>[10]</sup>,也可使用近年提出 的云优化 GeoTIFF 格式(cloud optimized GeoTIFF, COG), 它具有高效地提取 GeoTIFF 文件中的子区域的特性<sup>[11]</sup>。

主流地图服务软件支持瓦片的动态创建,也有 多种瓦片动态创建优化方法,如今主流的地图服务 软件中影像图层的发布需要手动指定参数,不利于 影像图层自动化发布;遥感影像数据具有多源性, 无法采用统一方法优化瓦片动态创建速度<sup>[12]</sup>。针对 文章编号: 1000-3096(2021)05-0047-07

上述问题,本文使用对象存储的 WebHooks 服务实 现遥感影像数据向 COG 格式的自动转化,利用 COG 格式的特性优化遥感影像瓦片动态创建的性能,设 计实现遥感影像自动发布的地图瓦片服务。

## 1 遥感影像瓦片快速动态创建基础

## 1.1 COG 内部结构

COG 是标准 GeoTIFF 格式,其对内部结构的改动都服从于 GeoTIFF 标准中允许但非强制要求的部分,具体结构<sup>[13-14]</sup>如图 1 所示。

头文件区中,标准 TIFF 签名后有一个隐藏区 (header ghost area),包含自身的大小、图像文件目录 (image file directory, IFD)的布局等描述信息<sup>[13]</sup>。头文 件区后面依次是全分辨率影像的 IFD、略缩图的 IFD 和影像数据存储区域。

IFD是 TIFF 文件标准中一种文件头的定义框架, 实质上是一组灵活的标签,包含影像数据在文件中 的位置信息,为文件读取提供捷径<sup>[8]</sup>。COG 中,前 16 kb 的数据中包含了所有 IFD<sup>[14]</sup>,读取这部分数据 就可以利用 IFD 访问数据局部,提取目标信息。

收稿日期: 2021-02-08; 修回日期: 2021-02-08

基金项目:国家重点研发计划项目(2017YFC1405600);国家自然科学基金(41976184)

<sup>[</sup>Foundation: National Key R & D Plan, No. 2017YFC1405600; National Natural Science Fundation, No. 41976184]

作者简介: 陈昱行(1995—), 男, 山东菏泽人, 硕士研究生, 主要从 事 WebGIS 相关研究, E-mail: chyuhang@qq.com; 刘善伟(1982—), 通信作者, 副教授, E-mail: shanweiliu@163.com





图 1 云优化 GeoTIFF 文件结构(据文献[13-14]绘制) Fig. 1 File Structure of Cloud Optimized GeoTIFF

## 1.2 COG 特性

使用 HTTP 协议请求影像数据时,可以在请求 头中使用 accept-ranges 字段指定所需数据的范围, 这种方式称为范围请求(range requests)。利用范围请 求,只需访问 COG 部分数据,就能从中提取头文件、 空间子集等数据,COG 中还为略缩图瓦片金字塔设 计了专门结构,可以在不读取完整数据的情况下预 览影像<sup>[7]</sup>。

为了分析 COG 部分读取与标准 GeoTIFF 总体读 取的效率差异,相关人员选用以下指标:(1)单点 的读取:选取稍微偏离中心点的点,读取该点的值。 (2)区域平均:取一个不规则的范围做空间子集的模 拟,获取该子集的平均值。(3)多波段影像读取:取一 个点,输出该点在所有波段的值<sup>[12]</sup>。通过对 COG 格 式影像执行以上3种操作得到对比结果(如图 2),在区 域平均和单点读取中,读取的字节数比率在 5%~15% 左右;多波段影像访问时的读取比率较高,但也只有 25%左右,相比总体访问模式效率提升明显。

地理空间数据抽象库(geospatial data abstraction library, GDAL)的虚拟文件系统(virtual file systems)支持使用/vsicurl 前缀指向在线数据,这种访 问同样支持范围请求,在此基础上,地图瓦片服务 的数据源将不再局限于自身文件系统中的数据,还 可以是局域网或互联网中的任意支持 HTTP 协议访 问的数据。



图 2 历问数据的日分比(据文献[12]修改] Fig. 2 Percentage of Read File Size<sup>[12]</sup>

## 2 遥感影像地图服务快速构建方法

## 2.1 影像数据存储

目前,遥感影像存储方式大致分为三种:关系数 据库(空间数据中间件)、网络存储和分布式存储<sup>[3]</sup>。 其中对象存储是网络存储的一种,使用 REST 和 SOAP 协议访问对象<sup>[15]</sup>,架构简单、部署容易,可以 更好地发挥 COG 的特性,本文设计的遥感影像地图 服务中使用对象存储服务管理遥感影像数据。

对象存储中,数据存储在桶(bucket)中,桶中对 象状态的变化可以通过事件广播监听,利用事件广播 的 WebHooks 接口,设计遥感影像上传、更新时的触 发器,构建 WebHooks 监听服务,流程如图 3 所示。





WebHooks 监听的触发事件采用 S3 中事件定义, S3 是 AWS(亚马逊云服务)的对象存储服务<sup>[15]</sup>,其接 口设计是目前行业事实标准<sup>[16-17]</sup>,监听事件与触发 事件的对应关系如表 1 所示。

海洋科学 / 2021 年 / 第 45 卷 / 第 5 期

#### 表1 监听事件与触发事件

lab. I	Listening and trigger	ing events
	监听事件	触发事件
S3:	ObjectCreated: Put	数据存储, 缓存清理
S3: ObjectCreated: Post		数据存储
S3: ObjectRemoved: Delete		缓存清理

通过监听表 1 中的事件,完成遥感影像数据的 格式转化和管理,数据存储服务提供传统地图服务 中图层管理的功能,依据当前数据存储情况给客户 端提供当前可用的图层信息。

## 2.2 影像地图发布

依据OGC的WMTS服务标准,一次瓦片请求必须出现的参数如表2所示。

### 表 2 WMTS 必要请求参数

#### Tab. 2 WMTS required parameters

参数	信息
Layer	图层
TileMatrix	瓦片矩阵
TileCol	瓦片矩阵中的列号
TileRow	瓦片矩阵中的行号

在传统地图服务中, Layer 参数指向的是一组配 置,这组配置包括原始数据地址、合成波段、空间范 围等信息。除原始数据地址、合成波段外,其余信息 都可从原始数据头文件获取。本文利用 Layer 参数指 定原始数据地址,添加参数 Indexes 指定合成波段, 替代之前 Layer 参数的功能。

生成相同数据的不同瓦片时,一些信息被重复 计算,如数据的值域范围等,本文设计影像数据统 计信息服务,通过数据地址返回一个包含统计信息 的 JSON 文本,在正式请求瓦片之前,先通过该服务 获取影像数据的统计信息,流程如图 4 所示。



图 4 客户端请求流程 Fig. 4 Workflow of client's requests

由图4,客户端从数据存储服务中获取可用的影 像数据(图层)列表,通过数据地址从统计信息服务 中获取该数据统计信息,之后将这些信息整合进请 求参数中完成瓦片请求。

## 2.3 影像瓦片动态创建

在接到瓦片的动态创建指令后,进入影像瓦片 动态创建流程,整个流程如图 5 所示。



图 5 瓦片生成流程图 Fig. 5 Workflow of the generation of tile

由图 5, 流程依次执行: 1) 解析参数, 计算虚拟 变换参数与空间子集: 解析波段索引(Indexes), 获取 合成波段; 解析瓦片矩阵(TileMatrix), 行号(TileRow) 和列号(TileCol), 得出投影坐标系下的瓦片的空间范 围,最终计算出虚拟变换参数、瓦片空间子集。2)读取瓦片数据:依据虚拟变换参数,从源数据创建虚拟变形数据集(warped virtual dataset),根据波段顺序和空间子集读取瓦片数据。3)线性拉伸:依据该波

段数据的值域范围对瓦片数据做线性拉伸,提升显示效果。4)无值区掩膜(mask)提取:若存在 alpha 波段,则将该波段作为无值区掩膜,否则计算得到无值区掩膜。5)图片渲染:将处理后的瓦片数据与掩膜数据叠加,渲染为图片。

## 2.4 影像瓦片缓存处理

为避免瓦片缓存对地图服务主进程的影响,使 用外置 NoSQL 数据库实现瓦片缓存,设计瓦片缓存 机制如图 6 所示。



图 6 瓦片缓存机制 Fig. 6 Workflow of tile caching

本文中瓦片唯一标识码的生成方法与 GeoServer 类似,不同的是,在 GeoServer 中,为了兼顾 WMS 协议,生成标识瓦片的唯一标识码时使用瓦片的外 包矩形作为参数,计算 Hash 值时为了保证外包矩形 各坐标中浮点型精度,增加计算量。本文生成方法使 用原始数据地址、瓦片所在缩放等级和行列号以及 其他参数计算出唯一标识,其中缩放等级和行列号 均为整数,减少计算量,过程如下:1)将请求参数 按照特定方法排序,并将请求参数统一转换为大写 字符。2)遍历排序后的请求参数序列,将其拼接成 以&字符连接的字符串。3)计算字符串 Hash 值得到 瓦片的唯一标识码。

对于每一次瓦片请求,首先通过中间件根据计 算出的唯一标识码确认是否有对应的缓存,若已有 缓存,则直接返回缓存数据,否则进入下个流程:将 该标识码与处理池已有的标识码比对,若未找到相 同的标识码,进入生产流程;若有相同值,则线程进 入等待状态,直到该瓦片生产流程结束,在一张瓦 片生产流程结束后,持有相同标识码的等待线程同 步做出响应,删除处理池中的该标识码,通过异步 的方式将瓦片缓存至瓦片缓存数据库。

## 3 应用实例与分析

## 3.1 应用环境与服务构建

实验选取主流开源地图服务器软件 GeoServer 作为对照,应用环境配置如表 3 所示。

#### 表 3 应用环境配置表

Tab. 3	Experimental	environment	configurati	ion
			0	

内容	参数
处理器	AMD EPYC 7551(2.0 GHz) 4 核
内存	8 G
操作系统	CentOS 8.0
GeoServer 版本	2.16

GeoServer 使用跨平台二进制 2.16 版本,使用 startup 启动脚本运行。影像发布步骤为:1) 手动执行 格式转化命令,将 GeoTIFF 影像转化为 COG 格式。 2) 建立 GeoTIFF 格式的数据存储(datastore):为数 据存储命名,指定源数据地址。3) 通过数据存储发 布图层:为图层命名,选择坐标参考系统(CRS),指 定空间范围和合成波段顺序。

本文设计的遥感影像地图服务实现过程如下: 数据存储端,考虑到易部署性和易拓展性,选用 MinIO(一款基于云原生技术的对象存储服务软件)构 建服务,编写 Python 脚本结合 MinIO 的 WebHooks 接口实现监听服务,包括文件格式转化和缓存清理 功能;地图瓦片服务端,使用 rasterio(GDAL 的 Python 绑定库)读取 COG 数据,完成瓦片的提取,使 用 aiohttp 和 asynio 优化并发访问性能。瓦片缓存端, 选用高性能、支持数据持久化的 Redis 数据库配合本 文设计的瓦片唯一标识码生成方法缓存瓦片数据。 影像发布流程为使用 MinIO 管理工具将数据上传至 对象存储服务。

#### 3.2 数据与实验设计

数据使用无人机多波段 DOM 影像,图 7 为影像局部,影像详细信息如表 4 所示。

实验分两组对照,一组使用 GeoServer 对同个数 据两种格式(GeoTIFF 和 COG)分别提供地图瓦片服 务,通过响应速度验证 COG 在瓦片动态创建中的速 度优势。另一组使用相同的 COG 数据,使用 GeoServer和本文设计实现的遥感影像地图服务分别 提供地图瓦片服务,通过响应时间测试两者在瓦片 动态创建和处理缓存瓦片两个方面的性能差异。





图 7 无人机 DOM 影像瓦片 Fig. 7 UAV DOM image tile

- 表 4 数据详细信息
- Tab. 4 Imaging Data Information

参数	内容
原始数据大小	1.5 G
COG 数据大小	1.7 G
图幅尺寸	21 410 × 28 940
数据类型	uint 8
缩放范围	15~22
波段数	3

实验过程中,向地图服务发送若干并发请求, 记录各请求的响应时间,为避免各瓦片包含的数据 量不同影响结果,测试用的瓦片(由缩放等级和行 列号描述)集,通过如下流程生成:1)读取影像的 空间范围、有效缩放等级范围,转换到投影坐标系。 2)对每个缩放等级,计算空间范围左上和右下两个 坐标所在的瓦片的行列号,得到该缩放等级下所有 瓦片,形成瓦片集。3)将瓦片集中的各瓦片的范围 与空间范围做叠加分析,移除未完全包含于空间范 围的瓦片。4)将瓦片集打乱顺序,按照10、50和100 的长度切片,得到瓦片测试集。

为了避免客户端与服务端之间网络波动对结果 的影响,执行并发请求和结果统计的 Python 脚本在 服务端运行,结果中的响应时间不包含网络延迟。

### 3.3 结果分析

本文设计的遥感影像地图服务称为简化地图服务,标准 GeoTIFF 格式称为非 COG。

瓦片动态创建时, COG 与非 COG 之间、两种服务之间性能对比如图 8 所示。





由图 8 可知,使用非 COG 数据动态创建瓦片时, 响应时间显著增长,由图 8c,在处理 100次并发请求 中最后一个响应时,使用 COG 与非 COG 之间响应 时间相差超过 2 min, COG 有明显优势。

## 研究论文 · Linn → ARTICLE

在同样使用 COG 格式时,简化地图服务与 GeoServer 对比,由图 8a,在处理 10 次并发请求时, 平均响应速度更快;由图 8b 和图 8c,在处理 50 次和 100 次并发请求时,简化地图服务在处理前部少量请 求时有速度优势,GeoServer 总体响应时间更优。

处理已缓存瓦片时,两种服务对比情况如图 9 所示。



Fig. 9 Response time comparison with caching

图 9结果表明,简化地图服务在上述 3 组测试中, 响应速度保持在 0.15 s 之内,而 GeoServer 的响应速 度随着并发请求个数的增加显著增长,依据图 9c, 在 100 次并发请求的测试中,完成最后请求所用的 时间超过 0.4 s。

综上得出,简化地图服务在动态创建瓦片的场

景下, 性能与 GeoServer 各有优劣; 在瓦片已缓存的 场景下比 GeoServer 表现更好。

## 4 结语

本文设计实现一种自动发布的地图影像瓦片服 务,相比传统方法构建的地图服务,有以下优点:

1) 构建了自动处理工作流,将源数据转化为 COG 格式,加速瓦片动态创建。

 2)源数据在地图服务外部管理,影像数据上 传、更新时自动发布地图服务。

该方法构建的遥感影像地图服务依然存在不足, 源数据压缩后会影响瓦片动态创建的速度,而存储 未压缩的数据占用更多存储空间,如何寻找两者间 的平衡点还需进一步研究。

### 参考文献:

- YAN J, MA Y, WANG L, et al. A cloud-based remote sensing data production system[J]. Future Generation Computer Systems, 2018, 86: 1154-1166.
- [2] MA Y, WU H, WANG L, et al. Remote sensing big data computing: Challenges and opportunities[J]. Future Generation Computer Systems, 2015, 51: 47-60.
- [3] 季艳,鲁克文,张英慧.海量遥感数据分布式集群化存储技术研究[J]. 计算机科学与探索,2017,11(9): 1398-1404.

JI Yan, LU Kewen, ZHANG Yinghui. Research on distributed clustering storage technology for massive remote sensing data[J]. Journal of Frontiers of Computer Science and Technology, 2017, 11(9): 1398-1404.

[4] 戴芹,刘建波,刘士彬. 海量卫星遥感数据共享的关 键技术[J]. 计算机工程, 2008(6): 283-285.
DAI Qin, LIU Jianbo, LIU Shibin. Key techniques of massive remote sensing data sharing[J]. Computer Engineering, 2008(6): 283-285.

 [5] 廖芳芳, 裴春营, 李永峰. 基于最高层级的影像分布 式切片技术研究[J]. 计算机产品与流通, 2020(10): 38-39.

LIAO Fangfang, PEI Chunying, LI Yongfeng. Research on distributed slicing technology of image based on the highest Level[J]. Computer Products and Circulation, 2020(10): 38-39.

[6] 史玉龙,侯传燕. Hilbert 索引的栅格瓦片金字塔数据 存储方案[J]. 新疆师范大学学报(自然科学版), 2020, 39(2): 13-16.

SHI Yulong, HOU Chuanyan. Raster tile pyramid data storage solution using Hilbert index[J]. Journal of Xinjiang Normal University(Natural Sciences Edition), 2020, 39(2): 13-16.



- [7] 李瑞清, 熊伟, 吴烨, 等. 一种基于 MBTiles 的地图 瓦片存储技术[J]. 地理空间信息, 2019, 17(12): 58-62.
  LI Ruiqing, XIONG Wei, WU Ye, et al. A map expansion storage technology based on MBTiles[J]. Journal of Geomatics, 2019, 17(12): 58-62.
- [8] 王冬至,林东铨. 一种适用于移动端国土调查应用的 瓦片地图存储方法[J]. 测绘地理信息, 2020, 45(5): 94-96.
   WANG Dongzhi, LIN Dongquan. A storing method of

tile data for land survey applications on mobile ends[J]. Journal of Geomatics, 2020, 45(5): 94-96.

- [9] 杨会元,冯钟葵,李山山. 基于 Web 的遥感影像在线 分类实现技术研究[J]. 遥感信息, 2015, 30(1): 101-106. YANG Huiyuan, FENG Zhongkui, LI Shanshan. Implementation of online remote sensing image classification[J]. Remote Sensing Information, 2015, 30(1): 101-106.
- [10] Data Considerations GeoServer User Manual: [EB/OL].
   [2020-10-16]. https://docs.geoserver.org/stable/en/user/ production/data.html.
- [11] LEHTO L, KÄHKÖNEN J, OKSANEN J, et al. Supporting wide user-base in raster analysis – GeoCubes Finland[J]. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2018, XLII-4: 329-334.

- [12] CHAN Y, CHRIS D, PATRICK Q, et al. Cloud-optimized format study[EB/OL]. [2020-01-01]. https://ntrs. nasa.gov/citations/20200001178.
- [13] EVEN R, ELI L A, CHEN Y. COG Cloud optimized GeoTIFF generator[EB/OL]. [2020-10-12]. https://gdal. org/drivers/raster/cog.html.
- [14] EVEN R, DAN B, MARKUS N. Cloud optimized GeoTIFF - GDAL[EB/OL]. [2018-08-01]. https://trac. osgeo.org/gdal/wiki/CloudOptimizedGeoTIFF.
- [15] KHAJA R, KOTA V. Hybrid cloud framework for object storage based geo spatial remote sensing data processing[J]. International Journal of Engineering Sciences & Research Technology, 2014, 3(12): 449-454.
- [16] 曹守欣,赵琉涛,金翊. 基于对象存储的云存储系统 研究[J]. Computer Science and Application, 2014, 4(12): 333-343.
  CAO Shouxin, ZHAO Liutao, JIN Yi. The study of cloud

storage system based on object-based storage[J]. Computer Science and Application, 2014, 4(12): 333-343.

[17] 陈阳, 王丹. Ceph RadosGW 对象存储集群的部署与 优化[J]. 现代计算机, 2020(14): 17-20.
CHEN Yang, WANG Dan. Ceph RadosGW object storage cluster deployment and optimization. Journal of Frontiers of Computer Science and Technology, 2020(14): 17-20.

# A rapid construction method of remote sensing image map services based on COG

## CHEN Yu-hang, SHENG Hui, LIU Shan-wei, HE Ya-wen

(College of Oceanography and Space Informatics, China University of Petroleum (East China), Qingdao 266580, China)

**Received:** Feb. 8, 2021 **Key words:** remote sensing imagery; map services; tile pyramid

**Abstract:** The existing image publishing process of a map tile server is complicated because it is not conducive to the construction of an automated process. This paper builds a private object storage service to manage the remote sensing of data. Cloud optimized GeoTIFF (COG) is used to support local access. The dynamic creation speed of image tiles is optimized, and a tile map service that automatically publishes images is realized. In the example application, the remote sensing image map services that are implemented in this paper are similar to those of the GeoServer in the dynamic creation of the COG data tile.

(本文编辑:杨 悦)