

RESEARCH ON MULTI-SOURCE SATELLITE IMAGE DATABASE MANAGEMENT SYSTEM

FU Ning¹, SUN Lin², YANG Hongzhen³, MA Jun⁴, LIAO Baoquan⁵

^{1,4,5} STATE GRID SHENWANG LBS(Beijing)Co.Ltd, BEIJING, CHINA

² Beijing HeadSpring Technology Co.Ltd, BEIJING, CHINA

³ STATE GRID ZHEJIANG ELECTRIC POWER Co.Ltd, HANGZHOU, CHINA

KEY WORDS: Multi-satellite image, Normalized processing, Relational Database, Electric Feature points

ABSTRACT:

For the exploration and analysis of electricity, it is necessary to continuously acquire multi-star source, multi-temporal, multi-level remote sensing images for analysis and interpretation. Since the overall data has a variety of features, a data structure for multi-sensor data storage is proposed. On the basis of solving key technologies such as real-time image processing and analysis and remote sensing image normalization processing, the .xml file and remote sensing data geographic information file are used to realize effective organization between remote sensing data and remote sensing data. Based on GDAL design relational database, The formation of a relatively complete management system of data management, shared publishing and application services will maximize the potential value of remote sensing images in electricity remote sensing.

1. INTRODUCTION

This document is a guide to using the GDAL database. With the continuous expansion of the scale of the power grid and the continuous upgrading and expansion of the information system, the real-time data is massive and has obvious characteristics of big data.^[1]

In order to observe the characteristic target points of the power system, this paper obtains multi-band high-resolution remote sensing images to observe the ground, and uses the method of visual interpretation and supervised classification to extract the power feature points on the ground. The overall operation process involves experimental testing of multiple satellites such as GF1, GF2, GF3, Terra SAR, and Cosmos. In order to view the image of the target area quickly and accurately, the designed display platform combining GDAL library and PostgreSQL for multi-source satellite images storage. While ensuring stable storage of data, it is also possible to find the target image quickly. The big data service platform can be applied to the power grid system to better solve the problems of storage, calculation and analysis of massive quasi-real-time data, which has high practical value.^[2]

2. KEY TECHNOLOGY

The section headings are in boldface capital and lowercase letters. Second level headings are typed as part of the succeeding paragraph (like the subsection heading of this paragraph).

2.1 GDAL

GDAL is a free, independent and powerful open source library, and an open source raster space data conversion library under the X/MIT license agreement.^[3] It uses a single abstract data model to express the supported file formats, as well as data conversion and processing through a series of command line tools.^[4] GDAL also includes an ORG library for manipulating vector data, so it has the ability to manipulate nearly 10 kinds of raster data such as bmp, jpg, tif, and nearly ten kinds of vector data such as *.dxf, *.gml, and ESRI Shapefile.^[5] The characteristics of GDAL have: 1. Scalability and portability are good. If you want to support a new data format, you only need to add the format driver; 2. The API is comprehensive and can be embedded in other programs; 3. can run across platforms; 4. support common spatial data formats.^[6] Based on the above advantages, Esri ArcGIS 10.4, Google Earth, ERMapper, IRSA remote sensing image processing software system and cross-platform Grass GIS system.^[7] Therefore, GDAL is used in GIS product development center widely.

2.2 Postgre SQL

PostgreSQL open source database system developed by the University of California, Berkeley.^[8] Can run on Windows, Linux and other platforms. PostgreSQL is one of the world's top relational databases, widely used in medical, astronomy, business and other fields.^[9] The main technical features of PostgreSQL include massive storage of PostgreSQL combined with the operational capabilities of structured query language

Corresponding Author: SUN Lin - E-mail:sunlinch121@163.com

Foundation Item: State Grid Science and Technology Project "Research on the Third-generation Precise Positioning Technology of BeiDou and its Application in Smart Grid" (No.5211XT180047)

and data processing capabilities of procedural languages, which can effectively support large-scale data storage and access, such as tablespaces of terabytes of databases. GB-level BLOB binary large objects and CLOB text large objects, etc., and ensure data integrity and security through various constraints. Standard Universal PostgreSQL is based on the widely used relational data model.

It follows the SQL international standard PostgreSQL to provide standards compliant ODBC3.0, JDBC3.0, OLEDB2.7 and embedded SQL interface PostgreSQL for application development. Verification, has become a standard, universal, secure, stable, practical, efficient data storage and management platform.

3. OVERALL PLATFORM DESIGN

3.1 Platform design goal

The system design is based on the PostgreSQL database and the GDAL class library to realize the rapid identification of multi-source satellite remote sensing images and coverage areas and the statistics of related satellite-related parameters. Through the development of satellite image acquisition time, coverage area, data source and other classification systems, image segmentation of satellite imagery, image information extraction, screening by time, region and other conditions. And to ensure the standard storage of multi-source data, play a better practical price. Design should follow the principles of scalability, practicality, compatibility, reliability, stability and safety.

3.2 Platform design framework

The system is developed using the B/S structure system. The overall technical framework structure is shown in Figure 1.

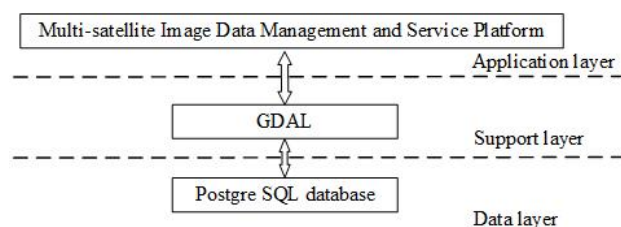


Figure 1. Overall technical framework

The system is logically divided into a data layer, a support layer and an application layer. The data layer specifically stores image data, the support layer provides various service components to access image data and respond to client requests, and the application layer is a client image data management system. In terms of data management, the system only involves image information such as satellite image and image capturing time, overall contour coverage spatial position, image resolution, etc., so the vector data is stored in a folder format in the Shapefile format and the image is in a grid format such as *.tif. Create a spatial database.

4. PLATFORM IMPLEMENTATION

In order to explore the distribution of the ground feature points of the power network, a large number of satellite images with different resolutions and different time items are obtained. According to the user's query requirements for the target image, this paper designs the system function from the perspective of

the user's daily business management requirements. The logical framework is shown in Figure 2.

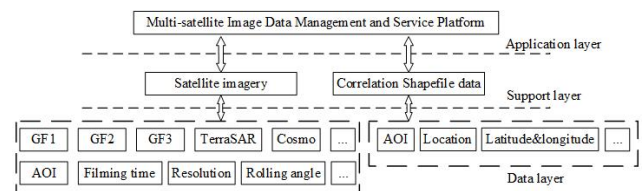


Figure 2. Overall logical framework

4.1 Data layer

The data is base of platform. The data layer is mainly the data resource content managed by the system, that is mainly composed of an image catalog database, a remote sensing image library, a vector information library. The data layer contains related data resources and is represented as a logical library. The specific data can be stored by Oracle or the file system depending on the type and usage. Spatio-temporal big data and storage characteristics in GIS environment (large data volume, heterogeneous data source, large data value, data uncertainty and information missing, real-time interactive visualization of data) determine the storage of large-scale spatial data: [10]

Distributed storage

The data layer management includes three types of remote sensing image, vector data and attribute data. The remote sensing image includes original image data, such as GF1, GF2, GF3, TerraSAR, Cosmo, etc. The vector data includes map tiles, national administrative division maps, power vector distribution maps, etc. The attribute data includes basic information of power, distribution information, market information, and user information. Based on the principles of standardization, standardization, security and practicability of database construction, the collected raw data is normalized for processing and storage management before database construction. At the same time, in order to better manage and access data, improve data computing and transmission efficiency, PostgreSQL is used to store data.

Data distribution strategies, metadata and their service models, there is the key technologies for distributed storage.

4.2 Support layer

The application service layer is composed of two subsystems: data management and data retrieval, according to different functional requirements.

(1) Data management subsystem

The data management subsystem satellite image catalog data production component completes various satellite image metadata information extraction and catalog data production; the original image storage component completes importing different types and different resolutions of original satellite image into the database mosaic data set; The library component completes the conversion of metadata for various products into database catalog data.

(2) Data retrieval subsystem

The data retrieval subsystem is composed of a data browsing component, a data query component, a data extraction component, and a data statistics component, and performs a data retrieval function together. The data browsing component completes the visualized directory data display, image data addition, layer control, and graphic basic operations; the data query component completes the spatial and attribute compound

query of various requirements, and completes the query result display; the data extraction component completes directly from the database Extract the corresponding data resources, or download the corresponding data files from the file database; the data statistics component completes the statistics, quantity, and capacity of various data resources by category, attribute.

4.3 Application layer

The user layer mainly includes the administrators and business personnel of the database. The database administrator maintains the database, and the business personnel process the daily business through the application system.

5. PLATFORM APPLICATION

5.1 Development environment

The operating system is Windows 7/10, the operating system environment is .net framework 4.6.1, the locale is VC++ and JAVA. The database uses PostgreSQL 11, the development platform uses GDAL2.4.0.

5.2 Function realization

The overall implementation goes through 4 steps.

(1) Multi-source remote sensing image preprocessing. Convert the .tif image format to .png using GDAL-WAP and convert the coordinate system from WGS1984 to CGCS2000.

(2) Multi-source image building. Using the .xml file in the original data, the attribute information in the remote sensing image is read, and the necessary fields of the database are established, and the fields including the "satellite name", "product time", and "upper left latitude and longitude" are established. Specific byte is shown in Figure 3.

raster image	raster information	vector data
PK rasterID	PK rasterID	PK power feature point
createtime	rollangle	rasterID
sensor	(B, L)	createtime
datapath	gray	...
...

Figure 3. Overall specific byte

(3) GDAL platform image reading. The image is read using the pyramid file, and each band of the remote sensing image is quickly read into the memory by using the interface function provided by GDAL. Part of the image showing the size of the window is selected from the pyramid for reading, saving data reading time. At the first reading, the default is the upper-left portion of the input image. Then, the input image is stretch-converted to create a bitmap file, and the R, G, and B components are calculated according to $Gray=0.299 \times R+0.587 \times G+0.114 \times B$, and the gray scale is assigned to the pixel.

(4) Image browsing. In image reading, a new range is selected for each movement to read and display. Therefore, in the algorithm design, the selection of the image to be read range is mainly considered to ensure that it does not cross the boundary. In this platform, define a public class member variable Extent_No. to record the coordinates of the Four corners of the

area to be read. At the current resolution, the size of the area to be read defaults to the size of the display window, if it is to be read. The essence of image scaling is to resample the image, the amplification is upsampling, and the reduction is downsampling. This platform mainly uses the nearest neighbor pixel method for resampling. Platform data management display interface is shown in Figure 4.



Figure 4. Platform data management

6. CONCLUSION

Based on the actual work requirements, this paper designs a remote sensing image data management platform based on PostgreSQL and GDAL through the learning of WebGIS architecture, and realizes a large number of image data management and retrieval functions in the power system. It not only satisfies the management of the regional image by the power management department, but also satisfies the query requirements for the image coverage of the target observation area.

Remote sensing image data management is only a subsystem of the platform. The next work will be based on the existing ones, further enrich and improve the platform, increase the breadth and depth of the system, and make a larger number of data read queries in the future.

REFERENCES

- 宋亚奇,周国亮,朱永利.智能电网大数据处理技术现状与挑战.电网技术, 2013, 37(4): 927 - 935.
- 李 钢,茅海泉, 基于 Postgre SQL 的海量准实时数据服务平台访问方案.Computer Systems & Applications,2019,28(2):274 -279.
- 赵辉辉.基于 GDAL 的农田信息系统研究[D].哈尔滨:东北农业大学, 2011.
- 吴 萍. 基于 GDAL 的电力 GIS 系统设计与实现[D].厦门:厦门大学, 2015.
- 薛小虎,沈大川,袁关伟,等.基于 GDAL 的谷歌影像瓦片数据拼接技术研究[J].测绘与空间地理信息, 2016,39(2):126-128.
- Felzenszwalb P F, Huttenlocher D P. Efficient Graph-based Image Segmentation[J]. International Journal of Computer Vision,2004,59(2):167-181.
- Marius Appel,Florian Lahn,Wouter Buytaert,Edzer Pebesma. Open and scalable analytics of large Earth observation datasets: From scenes to multidimensional arrays using SciDB and GDAL[J]. ISPRS Journal of Photogrammetry and Remote Sensing.

Jiang ZM, Hassan AE. Examining the evolution of codecomments in Postgre SQL. Proceedings of 2006 InternationalWorkshop on Mining Software Repositories. Shanghai,China. 2006. 179 - 180.

孙乔,焦艳斌.实时数据库系统研发及其在智能电网中的应用. 电力信息化 , 2012, 10(11): 105 - 109. [doi: 10.3969/j.issn.1672-4844.2012.11.023].

栾峰, 宁方辉, 滕惠忠.面向海量遥感影像的数据模型与管理模型设计 [J] .海洋测绘, 2014(5) : 76-78.