

A MANAGEMENT OF REMOTE SENSING BIG DATA BASE ON STANDARD METADATA FILE AND DATABASE MANAGEMENT SYSTEM

Hui.Wu¹, Kun.Fu²

¹Guangxi zhuang autonomous region remote sensing institute of natural resources, Guangxi Nanning, China, wilkely5689@163.com

²Guangxi Huayao Space Information Co., Ltd., Guangxi Nanning, China, freek1987@163.com

KEYWORDS: remote sensing image, big data, standard metadata file, data management

ABSTRACT:

As a kind of information carrier which is high capacity, remarkable reliability, easy to obtain and the other features, remote sensing image data is widely used in the fields of natural resources survey, monitoring, planning, disaster prevention and the others (Huang, Jie, et al, 2008). Considering about the daily application scenario for the remote sensing image in professional departments, the demand of usage and management of remote sensing big data is about to be analysed in this paper.

In this paper, by combining professional department scenario, the application of remote sensing image analysis of remote sensing data in the use and management of professional department requirements, on the premise of respect the habits, is put forward to remote sensing image metadata standard for reference index, based on remote sensing image files and database management system, large data serialization of time management methods, the method to the realization of the design the metadata standard products, as well as to the standard of metadata content indexed storage of massive remote sensing image database management.

1. GENERAL INSTRUCTIONS

Gartner said: Big data is data that contains greater variety arriving in increasing volumes and with ever-higher velocity. This is known as the three Vs (Oracle, 2019). Because of the large amount of data and a large number of categories, especially its professionalism on processing technology, makes remote sensing images are rarely appears in the field of big data. In fact, remote sensing images have the characteristics of large storage, large amount of information and a variety of data sources, which is fully conform to the characteristics of big data. On the other hand, remote sensing images have the information of geographical spatial location, which is accepted and loved by the vast number of professional technical and management department. So they are being used in Land and resources, meteorology, environmental protection and other fields. With the development of technology and the abundance of the data source, remote sensing image data is increasing with geometric series speed (Zhang, Fen, Yan Gao, 2005). More and more industries have star to use the remote sensing images to carry

out the unique data investigation, analysis, monitoring, and so on, the demand of remote sensing image in the storage and management is more and more prominent.

At present most of the professional departments who are in use of remote sensing images are still adopts the traditional file management mode on data management, which in the way of copy the image data and store them in hard disks or optical disc, and then manual inventory management. But how to realize the visualization query and a quick call of the massive remote sensing image in order to satisfy the requirements of data storage and application management in a cozy high speed without mass migration to existing data, which is a problem that badly in need of research to solve in actual application.

2. CURRENT SITUATION OF DATA STORAGE MANAGEMENT

At present, there are basically 3 kinds of methods in remote sensing image storage management: the storage management based on file, the storage management based on geographical database, and the storage management based on file and database management.

file-based storage management: The remote sensing images are stored in a file server in the form of files, being organized and managed through the way of file. In this way, it is more flexible and easy to maintenance management and do not need to establish the characteristics of the image database.

Geographical database-based storage management: Remote sensing image data management based on relational database storage way, is a way that can give full play to the advantages of a relational database. The way of geographical database-based has advantages in data security, multi-user concurrent control, storage expansion and so on, but its disadvantages are also very obvious: firstly it is a way that needs the support of large relational database platform, secondly it will be a big costs on time and energy when in storing image data, thirdly image data will take up larger storage space after building library, at last, it will be a many and miscellaneous job when updating and backup of the local data.

file and database mixed storage management: A way combined with the characteristics of relational database and file system storage management. Though this way, the remote sensing image data is stored in the file system, and the metadata information of the image is managed by the relational database. Since the data files are not stored in the database, the database only manages the index information of the image data, so that the storage and retrieval efficiency of the image data is greatly improved.

However, due to the different metadata contents of different data sources, the storage paths are different, and there is a phenomenon that data needs to be moved. In combination with the requirements of image management of the current professional department, a management method of mixing documents and databases is adopted to propose image elements. The data is managed in the warehouse, and the metadata is linked to the original data file for management.

3. STANDARDIZED PRODUCTION OF REMOTE SENSING IMAGE METADATA

Multi-source and multi-temporal remote sensing images are the targets of large-data management of remote sensing images.

Some image data and their overall characteristics are shown in table1 (Baidu Wenku, 2015).

Satellite name	sensor	Number of bands	Resolution
Landsat	TM	7	Bands 1-5 and 7 are 30m and band 6 is 120m
EO-1	ALI	242	Full color band 10m , multi-spectral band 30m,
GF1	PMS	4	Full 2m, Multi 8m
GF2	PMS	4	Full 0.8m, Multi 3.2m
ZY3	PMS	4	Full 2.1m, Multi 6m
BJ2	PMS	4	Full 0.81m, Multi 3.2m
WorldView3		29	Full 0.31m, Multi 1.24m
IKONOS		5	Full 1m, Multi 4m
QuickBird		5	Full 0.61m, Multi 2.44m
SPOT6		5	Full 1.5m, Multi 6m

Table 1. Characteristics of some remote sensing satellite image

The first-level imaging products are equipped with metadata. After the image is refined, the key information can be extracted based on the existing metadata, and the product-level image metadata is automatically generated. For the convenience of management, the standard form of the product metadata is designed. The content includes product basic information, data source status and production process information. The basic product information is an important reference for data storage, including metadata name, satellite identification information, acquisition time, ground resolution and four corners of effective pixels of the image. Point coordinate values and other content, (see the table below) to carry out the production of the automated standard library, the key is to analyze the multi-source level 1 image product information, taking the domestic satellite image data as an example, the product record information is as following table2:

Data Source	GF2	ZY3-1
Feature ID	SatelliteID	SatelliteID
Time	ReceiveTime	AcquisitionTime
Orbit ID	OrbitID	OrbitID
Resolution	0.8	2.1
Four Ends	Left top Longitude	TopLeftLongitude LeftTopPoint-Longitude

	Left top Latitude	TopLeftLatitude	LeftTopPoint-Latitude
	Right top Longitude	TopRightLongitude	RightTopPoint-Longitude
.....			

Table2. Examples of product record information

4. SYSTEM IMPLEMENTATION

4.1 System architecture design

Remote sensing image database management is mainly for the management of metadata file library and the recording of file system path. The overall system architecture adopts four layers of architectures: data layer, service layer, service layer and application layer.

Among them, the data layer is the basis, using the file management method for the storage of massive remote sensing images of different formats and different sources, and the business layer management database as a dynamic access interface, providing access interfaces such as type, time and data format.

The business layer developed the data management system of the C/S architecture to realize the addition, deletion, modification and investigation of the image database.

Considering about data characteristics and application requirements, the service layer builds a variety of data services such as image visualization, data management, and spatial analysis, improving data transmission efficiency and application quality.

The application layer has developed a multi-source remote sensing image sharing application system based on B/S architecture, which mainly includes functions such as visual browsing, multi-temporal analysis, data association query, statistical analysis, data distribution and uploading.

4.2 Remote Sensing Image Management Database Design

The remote sensing image management database mainly serves the business layer, and the business layer realizes the purpose of managing massive remote sensing image data by using the dynamic database.

Specifically, by setting a standard entry space, the database management system automatically searches for standard image product metadata (xml format) in the space, and extracts image

name, resolution, projection, effective pixel coverage, and image acquisition time in the recorded metadata information. The absolute storage path is automatically recorded, and the contents of the database record and its corresponding metadata label are as figure 1 shown. Massive images are organized and stored in the disk system by means of files, and the disk is supervised by the database management system to realize dynamic retrieval and update of the remote sensing image storage path.

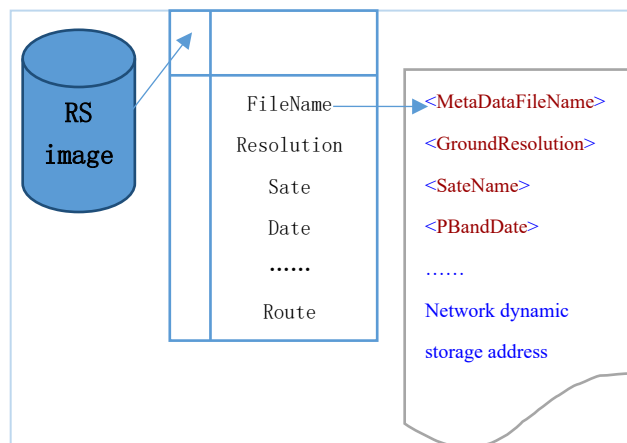


Figure 1. Remote sensing image management database

4.3 Image visualization management

The system framework adopts Microsoft Visual Studio 2015 integrated development platform, based on ArcGIS Engine 10.1 for secondary development to realize the functions required by the system, using Windows Form development mode, DevExpress 15.2 control library, C# language.

The system realizes visual query of images. In order to improve query efficiency, the system design implements multiple query methods including spatial query, attribute query, and spatial attribute comprehensive query.

Spatial query: Spatial query is a query mode based on GIS geospatial range. The realization principle of spatial query is to use the geographical range selected by the user or the geographical range of the custom drawing to spatially superimpose with the actual range of the image. The superimposed form includes the relationship of intersection and cutting, and the spatial query function makes full use of the metadata information and reads the corner coordinates of the image recorded in the metabase. After the spatial query passes the geographic scope coordinates, the system constructs a query window according to the range, and generates corner polygons one by one according to the coordinates of all metadata corner points recorded in the database. The spatial space query is

judged by the query window and the corner range polygon, and a query result set is generated.

The client performs spatial query, including selecting administrative divisions and constructing arbitrary geometric shapes. By selecting the administrative division and drawing the outer rectangle of any geometric shape to obtain the geographical range, the spatial condition of the query is formed, and the image is superimposed and inquired.

Attribute query: The attribute query is to query the stored information in the database by setting the attribute condition information, return the image data that meets the conditions, and display it in the system, and the attribute query object is the data information in the metadata database. The system design user interface of attribute query, the user can set the query condition according to the window prompt, the system constructs the SQL query statement according to the user's selection condition and performs the query operation, returns the metadata number result set, and the metadata data result set is The base further queries the image table to generate a final image result set for display in the system. Includes queries for image type, resolution, imaging time, and spectral range.

Spatial attribute comprehensive query: The spatial attribute comprehensive query combines the two query modes of space and attribute, fully considers the query efficiency, and returns the result set satisfying the condition according to the attribute condition; on the basis of the result set, according to the geometric range query window transmitted by the constructed polygon and the client The spatial intersecting judgment is made, and finally the result of the query is obtained and displayed in the system. The results obtained by the above query method include preview of the original image, preview of the image fast view and preview of the metadata. In order to visually display the query result, the preview function is fully integrated into the main interface of the system.

5. REAL SAMPLES REVEAL

Depends on the remote sensing image big data management system, the medium and high resolution multispectral remote sensing images stored in the department's server in recent years were extracted and stored. According to the statistics, the quantities of remote sensing image is recorded as Table 3, and the storage situation is shown in Figure 2-4. Among them,

according to the query definition, the data of 2018 image results are grouped and analyzed by half of the year.

Time	Quantities
Year 2017	2121
Year 2018, month1-6	2299
Year 2018, month7-12	1246

Table 3. Image storage statistics

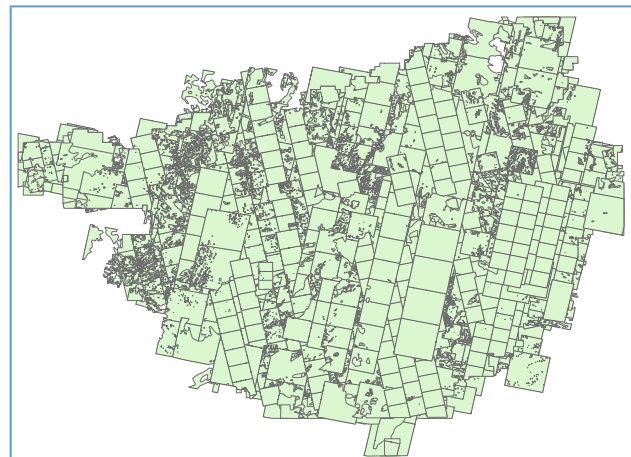


Figure 2. Store images of year 2017

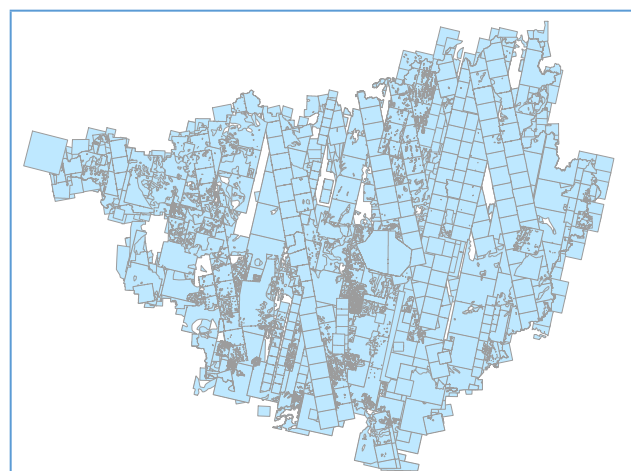


Figure 3. Store images of year 2018, month 1-6

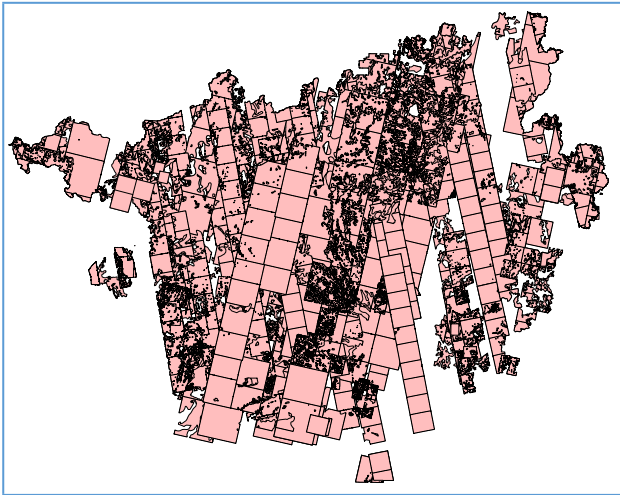


Figure 4. Store images of year 2018, month 7-12

After images are recorded in the management system, using the system's comprehensive query function, the 2018 annual image is divided into two periods from January to June and from July to December. Based on the remote sensing image results that have been stored in the library, the image coverage area are the results calculated by the system though the way match the location of the image and the image location stored in the database, and finally, the system will draw out the result due to the four-to-information record recorded by the image library. The calculation result of the query on a point (star) on the map is shown in Figure 5. In addition, the image display function can be adjusted to display the image data set that meets the conditions in the query window.

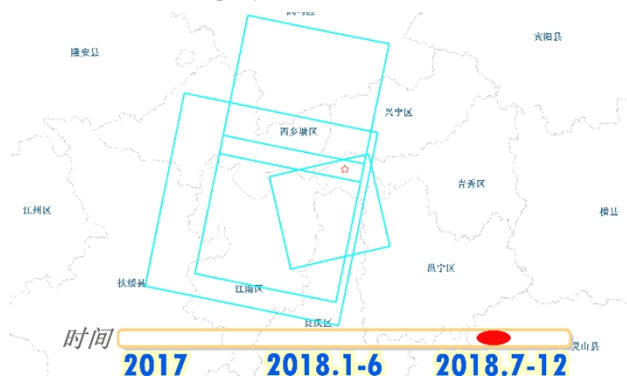


Figure 5. The query of a star point shown on the map

6. CONCLUSIONS

Guided by the needs of satellite remote sensing data in various industries, and using advanced technologies such as remote sensing, computer networks, databases, and geographic information systems, establish a database of remote sensing

image resources by fully integrate existing mass remote sensing image data, existing production technologies, equipment and other resources, and then develop a massive image data management system; build a unified, stable, scalable and compatible satellite remote sensing image big data management system with the core of providing efficient and fast remote sensing image comprehensive services. The remote sensing image big data management system based on file and database management system realizes visualization and time serialization management of massive remote sensing images, realizes centralized management, processing and service of massive images, and improves the utilization and range of remote sensing images. In the future, satellite remote sensing(Luo, Rui, et al, 2000).

REFERENCES

- Huang, Jie, Renyi Liu, Nan Liu, et al, 2008: Research and Implementation of Massive Remote Sensing Image Management and Visualization System. *Journal of Zhejiang University (science edition)*, 1008-9497, 35(6):701-706
- Oracle, 2019. What Is Big Data?.
<http://www.oracle.com/big-data/guide/what-is-big-data.html>
- Zhang, Fen, Yan Gao, 2005: The Application of MRSID Technology in Image Database System. *Bulletin of Surveying and Mapping*, 0494-0911(2005)04-0040-03: 40-42
- Baidu Wenku, 2015. Summary of common remote sensing satellites and sensors.
<https://wenku.baidu.com/view/2e11d0e048d7c1c708a145fe.htm>
- Luo, Rui, Yongsheng Zhang, Ying Li, et al, 2000: The establishment of remote sensing image data model and image database. *Journal of institute of Surveying and Mapping*, 1009-427X, 2000(04), 273-276