A Marine Remote Sensing Spatial Database Engine for Web Publishing

CHEN Zhirong  XU Caijiang

Abstract  To meet the requirements of efficient management and web publishing for marine remote sensing data, a spatial database engine, named MRSSDE, is designed independently. The logical model, physical model, and optimization method of MRSSDE are discussed in detail. Compared to the ArcSDE, which is the leading product of Spatial Database Engine, the MRSSDE proved to be more effective.

Keywords  marine remote sensing data; spatial database engine; geographic information system; web publishing

Introduction

How to organize, access and publish marine RS (remote sensing) data effectively on the internet has already been an urgent matter in marine RS technology[1]. The main difficulties lie in web-oriented storage, index and transmission of huge image data[2]. The goal is to establish effective storage and access strategies for the marine RS data in the Object Oriented relational database to solve problems[3]. Many researchers have put forward several ideas with regard to integrative storage of image data and metadata, and have obtained good results. For example, according to the characters of Internet and RS image, Yang Chaowei accomplished RS image web publishing by integrating spatial index of RS commercial relational database into RS data files[4]. Using advanced database technology for RS image data storage and management, Chen Huabin provided a simple web metadata system, which can easily achieve query, browsing and analysis of RS data[5]. He Zhengguo discussed some key technical studies of a raster data web publishing system, such as database connection pool, intercurrent access, and so on, but has given the realization framework as a whole[6]. These researches have ignored the vital feature of marine RS data—the short update cycle. Currently, marine RS data access still applies a common technical approach, and it is difficult to achieve marine RS data real-time update and dynamic publishing functions. Therefore, the challenge is to develop efficient data access and reading technology to meet the requirements of frequent interaction with database and high-performance data access.

Based on the analysis of existing spatial database engine technologies, a marine RS spatial database engine-MRSSDE is designed independently. Compared with other existing engines for performance testing, the correctness and feasibility of MRSSDE has been validated.

1  MRSSDE architecture

1.1  MRSSDE logical model

This paper is based on the expansion of the object-oriented relational database which was built for...
RS images. It imports the conception of an object into relational database to extend the data model in order to enable it to store image data.

According to the object-oriented perspective, marine RS data model can be reduced to cell(or pixel), block(or tile), band (or layer), metadata and other elements, as shown in Fig.1.

1) The core raster data

The core raster data is a multi-dimensional pixel matrix. In the matrix, the coordinates of a two-dimensional plane can be represented as \((X, Y)\), and the third dimension can be height, time or spectral information. Pixel depth represents the size of data.

2) Raster block

The core data need to be divided into blocks in order to optimize the storage and access. Block will be stored in the raster block table in the form of binary large object (BLOB). Along each dimension’s positive direction of the block border, MRSSDE will fill the block with filling-pixel (as Fig.2), if dissatisfied with the size of a block, to ensure that all blocks have the same size. The value of filling-pixel is zero. The block after filling has the same pixel depth as the others.

3) Band or layer

MRSSDE data model applies the logical hierarchical structure. The core data consist of one or more layers. For example, for multi-band RS data, it uses multiple layers to represent multiple bands. Layer here is a logical concept; usually one layer corresponds to one band.

4) Metadata

In MRSSDE data model, the core data are taken out, thus metadata is left. Metadata can be divided into object information, marine RS data, band information and spatial reference information.

1.2 MRSSDE physical model

The mainstream SDE products have their own characteristics in the areas of architecture, vector storage model, spatial query and index, etc., but are weak in the raster data storage and management, such as image data seamless storage, especially in the management of marine RS data, which is short-cycle update and has real-time characteristics. There are still many needs for sound.

In view of the above, taking full account of the temporal phase character of marine RS data, MRSSDE uses four tables to realize the seamless storage in Oracle9i.

1) User metadata table

This table is the systemic table in the system, mainly taking charge of maintenance and management of marine RS data.

2) Raster table

Related storage parameters and spatial coverage range are stored in the raster table. The table structure fields are defined as follows: RASTER_ID is the number of marine RS datasets, mainly responsible for distinguishing which image the data block belongs to when many images are stored in one table; BAND_NUM data means the image band; WIDTH and HEIGHT correspond to the images’ width and height; DATATYPE corresponds to image type; PYRAMID_MAXLEVEL means the highest level of the pyramid; BLOCK_WIDTH and BLOCK_HEIGHT correspond to the width and height of the block; XMIN, YMIN, XMAX and YMAX save the spatial coverage range of RS data, if there is no coordinating extension, the image size takes its place.

3) Raster block table
The table stores raster block related parameters. RASTER_ID is the foreign key, according to the raster table; RASBAND_ID is the band number; PRD_ LEVEL means the current pyramid level; ROW_ NUM and COL_ NUM correspond to the block’s row number and column number; BLOCK_ DATA saves the actual data storage location; ENVELOPE is responsible for the block borders, and combined with spatial index mechanism provided by Oracle Spatial to achieve efficient marine RS data retrieval.

4) Raster aux table
The table mainly stores general statistical information and palette information of RS data. RASTER_ID and RASBAND_ID stand for foreign keys, according to the raster block table; BAND_MINVALUE and BAND_MAXVALUE correspond to the minimum and maximum pixel values of the band; PALETTE saves palette information, does the deed only for single-band IMG.

1.3 Storage optimization design
Dynamic real-time updating of marine RS data has directly caused the rapid expansion of spatial database size. Very frequent database interaction makes the performance problem of the database system become more and more prominent. Therefore, optimization of the performance of spatial database is particularly important. MRSSDE adopts two steps to improve operational efficiency:

1) Adjusting data block size
Data block is the basic unit of data manipulation in Oracle Database. The right block size setting is very important to the entire system’s performance. For example, first, to the question of disk reading performance and network transmission performance, too big or too small block will cause an increase in overall operations (disk reading and writing, the memory copying, network transmission); second, to the question of multi-resolution hierarchical model, arbitrary block size is not conducive for establishing image pyramid and Quad-tree index, and the best block size should be a two integer power. As shown in Fig.3, when MRSSDE stores marine RS data, the data are automatically separated into several pieces into 128 × 128 left-to-right and up to down, and one data block is saved as one record. Once smaller than one block size, which is defined by the user, it should be stuffed by filling the pixel and saved as one block.

2) Building image pyramid
When the program needs to browse the data stored in the database, the higher the resolution is, the more obvious the details of the ground are, and the greater the volume of data is, the slower the speed is. So, on the basis of data block, MRSSDE builds a multi-level resolution image pyramid in the database (Fig.4). When the client requests data, the corresponding level pyramid image, which is in accordance with the current scale and extent, will be the feedback. In this way, seamless browsing among different resolutions is realized, and the transmission efficiency of the network is improved by less data transmission but the same visual effect.

2 Performance test and analysis
In order to verify the feasibility of the design of the engine, the requirement of storage space, data up-loading efficiency and response time, three aspects of MRSSDE performance have been tested, and the results show that the engine is effective.

The test used Oracle9i (Oracle9.0.1 version) as the
object-relational database and Windows2000 as the operating system. The server-side software are Oracle9i, MRSSDE and Zhejiang Marine Water Quality Real-time Inspecting System, and the client browser is IE6.0. Comparison target is the mainstream spatial database engine—the United States ESRI Corp’s ArcSDE9.0.

2.1 Requirement of storage space test

The main purpose of this test is to verify the relationship between data block size and the storage space required in the database. 15.8G floating-point data is used as test data. Table 1 shows the requirement of storage space test.

Through the results, we can see that the record number in database lies on the data block size in lossless compression condition. The smaller the block size, the more the records in the raster table are. Conversely, the bigger the block size is, the less the record number is. From the effective rate of storage space perspective, the data block with 32 × 32 size looks the same as that with 128 × 128 size. But from the transfer efficiency perspective, the data block with 128 × 128 size has fewer table indexes because of its small number of records, so it will be more effective. Block 256 × 256 looks most effective from the figures. But for the system as a whole, a very big data block will cause problems in display and management. Therefore, the right size of data block needs to be chosen to save storage space.

| Block size | Row average length(Byte) | Record number | Database storage space/G | Database storage space after build pyramid/G | Efficiency rate of storage space/% |
|------------|--------------------------|---------------|--------------------------|---------------------------------------------|----------------------------------|
| 32×32      | 2 251                    | 5 208 150     | 10.96                    | 13.37                                       | 82.0                             |
| 64×64      | 9 130                    | 1 361 160     | 10.90                    | 17.66                                       | 61.7                             |
| 128×128    | 33 716                   | 329 030       | 10.07                    | 12.95                                       | 77.8                             |
| 256×256    | 134 064                  | 82 960        | 11.56                    | 12.76                                       | 90.6                             |

2.2 Data upload efficiency test

The test of data uploading efficiency can reflect the storage efficiency of the database engine objectively. In the experiment, the selected data size is 60.2 M, the image width is 3 600 pixels and height is 5 840 pixels, the image file is in BSQ format, and the band number is 3. The size settings of the tested block are 32, 64, 128, 256, 512, 1 024 and 2 048. Table 2 shows the result of data uploading efficiency of ArcSDE and MRSSDE in the same environment.

From Table 2 we can see that: (1) the blocks sizes have a marked impact on the efficiency, and too big or too small size is not conducive to the database storage and network transmission; (2) when the data block size ranges from 128~256, the data transmission is more efficient, and relatively stable; and (3) the network transmission speed of MRSSDE which is based on OCI technology is faster than that of ArcSDE.

| Block size | ArcSDE9.0 cost | MRSSDE cost |
|------------|----------------|-------------|
| 32×32      | 96.65          | 79.45       |
| 64×64      | 73.60          | 56.48       |
| 128×128    | 50.08          | 42.66       |
| 256×256    | 54.50          | 42.47       |
| 512×512    | 64.02          | 51.27       |
| 1 024×1 024| 53.13          | 47.64       |
| 2 048×2 048| 64.88          | 49.78       |

2.3 Test of response time

Response time refers to the time cost for response of user operation. Table 3 shows the specific test results.
From Table 3 we can see that: (1) the block size has a significant impact on client response time cost, and too big or too small size does not help to improve the speed of response; (2) when the data block size is 128, the average response time is the least, and that of MRSSDE is less than that of ArcSDE.

| Block size | ArcSDE average response time | MRSSDE average response time |
|------------|-----------------------------|------------------------------|
| 32×32      | 0.608                       | 0.653                        |
| 64×64      | 0.570                       | 0.510                        |
| 128×128    | 0.496                       | 0.398                        |
| 256×256    | 0.632                       | 0.618                        |
| 512×512    | 0.794                       | 0.840                        |
| 1 024×1 024| 1.487                       | 1.283                        |

3 Conclusion

Focusing on the characteristic of marine RS data, this research has designed MRSSDE, and describes the logical model, physical model and optimization methods of MRSSDE in detail. Compared to the mainstream database engine ArcSDE, MRSSDE has been proven to be more effective and feasible, for meeting the requirements of marine RS data storage, management and Web publishing.

References

[1] Zhao Yanling, He Xianqiang, Wang Difeng, et al. (2005) Web-based querying system of ocean satellite products[J]. *Donghai Marine Science*, 23(1): 32-39

[2] Zhu Tiewen, Chang Lei, Li Qi, et al. (2005) Research on database-based management technology for image data[J]. *Computer Science*, 32(1): 95-97

[3] Du Yunyan, Su Fenzhen, Yang Xiaomei, et al. (2004) A study and development on the scientific data platform of China coast and offshore[J]. *Acta Oceanologica Sinica*, 26(6): 29-36

[4] Yang Chaowei, Li Qi, Cheng Jicheng, et al. (2000) Research and implementation on web publication of remote sensing images[J]. *Journal of Remote Sensing*, 4(1): 71-75

[5] Chen Huabin, Bi Jiantao, Cao Yanrong, et al. (2005) Construction of remote sensing data warehouse and web services[J]. *Science of Surveying and Mapping*, 30(5): 84-85

[6] He Zhengguo, Du Juan (2003) An web issuing system of raster map[J]. *Bulletin of Surveying and Mapping*, (1): 56-58

[7] Xu Bo, Weng Huanxin, Dong Chengsong (2004) Application of the ocean environmental information database on the basis of GIS in ocean environmental information visualization analysis[J]. *Zhejiang University (Science Edition)*, 31(4): 471-475

[8] Li Zonghua, Peng Mingjun (2005) Remote Sensing Image database based on RDBMS[J]. *Geomatics and Information Science of Wuhan University*, 30(2): 166-169

[9] Hadid M S (2000) Representing and reasoning on conceptual queries over image databases[J]. *Journal of Intelligent Information Systems*, (14): 131-154.

[10] Wu Xincai, Guo Lingling, Li Jun (2002) Mass image database management and Web publishing based on RDBMS and COM[J]. *Image and Graphics*, 4: 384-387

[11] Liu Nan, Liu Renyi (2002) Principle & Appliance of Web GIS[M]. *Beijing: Science Press*