THE MANAGEMENT OF MULTI-SCALE SPATIAL DATA BASED ON GRID AND CATALOG

SONG Guo-min, JIA Fen-li, MAO Yu-zhu
(Institute of Surveying and Mapping , Zhengzhou 450052, P.R.China)

ABSTRACT

With the deep application of GIS, multi-scale spatial data management becomes a basic and very important research spot. The difficulty of it is how to build a flexible mode of multi-scale spatial data management so fast query of multi-scale spatial data could be gotten. This paper uses the idea of grid for reference, constructs a metadata database of multi-scale and multi-source spatial data based on catalog. Then how to use this database to dispatch data is explained.

Keywords: Multi-scale and multi-source spatial data, Grid, Catalog, Metadata database

1. INTRODUCTION

With the deep application of GIS, one-scale spatial data can't meet our demands. As usual, multi-scale and multi-source spatial data is required to store in one system, so that they can switched naturally and fluently when used. Here, multi-source and multi-scale spatial data management becomes a basic and very important research spot. Multi-source and multi-scale spatial data management, it means to build a flexible and efficient mode of data catalog and data query so that it could be fast and efficiently gotten. This paper uses the idea of grid for reference, constructs a management of multi-scale spatial data based on catalog.

2. ACTUALITY OF THE MULTI-SOURCE AND MULTI-SCALE SPATIAL DATA

After more than twenty years efforts, relational departments of our country have achieved huge success in the construction of the national fundamental Geo-spatial database[6]. So far, about this construction, our country has constructed the nation's 1:4000000, 1:1000000, 1:500000, 1:250000 database, 1:50000 database of some territory, and 1:10000 database of the key anti-water areas of the seven big rivers; 1:1000000, 1:50000 DEM; 1:1000000 place-name database and gravity database of experiment areas. Also we are starting up for the construction of the nation's 1:50000 database, each province's 1:10000 database, and 1:5000, 1:2000, 1:1000, 1:500 database of important cities and areas. After several years construction of the national ground fundamental database, part of the national and provincial goal has been gotten, such as 1:500000 database of national land using, database of national land planning, database of provincial plan of land using, database of key cities' land using plan, as well as land using remote sensing monitoring database, database of national development areas, land's standard price database of national cities, and 1:10000-1:50000 land using database of national towns. During the ninth "five-year", national basic resource and environment spatial database

*Songgm@tom.com

is constructed, which consists of remote sensing photograph database, land using & covering thematic map database, digital images base of field surveying, multi-sampling framework database, and dynamic monitoring database. 7 national resource environment and area economy integrated databases are done with the use of the technique of multi-source geo-spatial data's spatial-time matching and integrating. Research is paid to the construction of national database for national resource environment and district information system, which served for national high level and synthesis programming management. Based on this, 6 national synthesis database including land, water, forests, oceans, mine resources, and district economy, as well as a provincial NREDIS(HaiNan province), national basic geo-database, national remote sensing resource environment synthesis database and NREDIS synthesis, totally 10 spatial information systems that can be shared with distributing nets, also spatial metadata base and net workstations, are founded by the use of spatial metadata base criteria, spatial meta-data criteria platform of our own and tools of meta-data soft.

In fact, these spatial data can be generally divided into two types: vector spatial data, such as series of scales of vector map data, other scales of local map data collected for special using. It is very difficult for this type in the algorithm of clipping and segmentation, more complex in merging technique, which needs complex operation like topology reconstruction, and element merging. raster spatial data, such as DTM, aerial images, satellite images, which are easy to clip or segment, also with much more simple merging technique. Data of this type is stored in the form of patch and scale (& resolution) in current spatial database in addition , the form of patch is mostly divided into rectangles. This paper wants to reach the goal of building a mechanism to use current data efficiently. So, obeying data's characters is the basic thought and basic gist on designing multi-source and multi-scale spatial data arrangement mechanism.

3. THE IDEA OF GRID & MULTI-SOURCE AND MULTI-SCALE DATA MANAGEMENT

Grid is a new research field rising in recent years. The concept is very open, and different field has different definition and understanding. In the field of geo-science, grid means the method of geo-space partition. The concept of Grid has been attempted early in geo-science. Binary decision tree, quad tree and octree, all these methods can divide certain spatial space according to different grid levels, thereby, different levels of grid are made[1, 2]. In fact, map sheet with different scale also makes a multi-level grid. For example, in DEM, covered area is divided into small grid. Each one is the smallest unit of information. It stores the height, origin of position's coordinate and grid distance. Also it can display the earth's surface using classification with LOD. The earth grid "Global Grid", researched by Goodchild, Dutton, Sahr and White[3, 4, 5], divides surface of earth into pieces according to different spatial dividing methods (such as division based on latitude and longitude, surface mapping division based on earth positive polyhedron). So it can depict spatial position relationship of the earth surface effectively.

Obviously, grid in geo-science area covers the meaning of multi-scale. So it is a good and effective way to catch the grid thought to design multi-source and multi-scale spatial data management mechanism. This paper uses the idea of spatial information multi-level grid for reference; its essential is to dismiss the concept of map sheet widely used in current spatial data management, to use grid and unique global ID to manage data. It can not only satisfy current management way based on map sheet, but also has a strong expansibility that provides a base for conversing to GIS based on grid later.

3.1 The Way of Grid Division

As using grid idea, the data scale can't be divided into seven levels according to series maps. It should meet the needs of depicting global geo-information and multi-source data characters. So division according to longitude difference and latitude difference is a good method. Based on the consideration of getting a better match about current map of basic scale, longitude difference 6 , latitude difference 4 are chosen to construct the grid, which is in accordance with the division of map sheet of current 1:1000000 scale. 24°×16° as the starting grid, curtailing in half of the current grid size or widening in two times of the current grid size, then specifying the starting longitude and latitude of each grid level, we can get a intact grid system after 24 times of curtailing.

The advantages of this division method are as follow:

- The grid division mechanism this paper mentioned is for the purpose of adapting both current using map with basic series scales and grid system in the future. It belongs to a transition grid mechanism. In this way, we can use grid number as basic index tools when coding. So this will greatly reduce the change when conversing to grid mechanism in the future.
- Though the grid mechanism of this paper is according to the longitude difference and latitude difference of current series maps, we can also put images and DEM and other multi-source data into grid to manage easily because of using grid. Most of these data is raster data, whose clipping and merging is easier than that of vector data. So it is convenient to manage both raster and vector data, also convenient for multi-source data fusion.

3.2 The Way of Grid Coding

In a system using map sheet numbers as its index tools, calculating is very slow. It mainly has two reasons:

- National map sheets of current maps have the difference between new and old, letters between capital and lower, and signs of both half the earth between added and not added, all the above, The calculation will be loaded down with trivial details.
- Numbers and letters mixed in map sheets slow the speed of calculation.

Quality of codes affects the efficiency of programs when data is distilled and called frequently. This paper uses grid coding as the index tool in attempt to get the grid row and column numbers quickly when searched. When used, firstly we ensure which level this data located, and then we ensure which area the grid covers. As bit operation is the most fast in the computer operation system, here, we use bit operation coding.

It is very simple to calculate a point (B, L)'s row and column number. The formula is as follow:

$$C = (L - L_0) / L_c; \qquad R = (B - B_0) / B_c$$

Here, L_0 B_0 are starting longitude and latitude; L_c B_c are grid size.

For the grid code containing all levels of grid, we use a 64 bit unsigned integrate data to store. So if C and R form a grid code, that means to move C left for 32 bits, which is depicted in C language as

```
N = C;
N <<= 32;
N += R;
```

following:

4. THE MULTI-SCALE DATA MANAGEMENT BASED ON CATALOG

This catalog mechanism means cataloguing multi-source and multi-scale data, getting names of relevant scopes and relevant levels through a well-designed meta-database.

4.1 The Multi-scale Data Management

Map data storage has two forms: one is file, the other is database. For the management of multi-scale data, this paper designs a multi-scale grid metadata base based on grid. It considers multi-source and distributed stored data, and designs the following data table: data format table, data path table, data level table, grid level table and data tables of different levels (scales) & formats. When used, each data source node has a metadata base to depict its data structure.

Data format table is used to store kinds of data and the prior level of different format data in this data source node.

Data path table is used to store paths of different format data in data source nodes.

Data level table is used to store the levels of data, including level names and descriptions.

Grid level table is used to store the corresponding relationship of grids and display scale, and division of this level grid.

Different levels (scales) and formats data tables take charge of storing grid numbers, names, longitude and latitude scopes, changing time and data sources of each data file (or database table), in order to select in data of different formats.

We can manage data of different types and formats through meta-database, which is a conversion to grid GIS in the future. Structure of the table is as follow:

Table 1. Data format.

serial number	Data name	Data type	length	byte	instruction
1	type	int			Data type corresponding to system enum type
2	format	char	4	4	Such as“.shp” “.mid”

Table 2. Data path.

serial number	Data name	Data type	length	byte	instruction
1	type	char	4	4	Such as“.shp” “.mid”
2	path	char	1000	1000	

Table 3. Data level table means data collection’s levels, such as settlement place and transportation

serial number	Data name	Data type	length	byte	instruction
1	format	char	4	4	Such as“.shp” “.mid”
2	Level number	int			Only numbers, no other meaning
3	Level name	char	50	50	Data level names
4	description	char	255	255	Description of this data level

Table 4. Grid level table (level numbers and division methods of each data grid)

serial number	Data name	Data type	length	byte	instruction
1	type	int			Data type
2	Level number	int			Grid level
3	row	Long int			Grid division of this level
4	column	Long int			
5	Start longitude	double			Start longitude and latitude
6	Start latitude	double			
7	Lowest scale	double			Display scale of this level
8	Highest scale	double			

Table 5. ×××_××table (×××means data types   ××means levels)

serial number	Data name	Data type	length	byte	instruction
1	Piece number	char	64	64	
2	Smallest longitude	double			
3	Smallest latitude	double			
4	biggest longitude	double			
5	Biggest latitude	double			
6	Data collection name	char	255	255	Corresponding file and table

4.2 The Dispatch of Multi-scale Data

Dispatch of multi-scale data means using the metadatabase and transfering relevant data at relevant time when multi-scale metadatabase comes into being. Here, we call components of data storing, indexing, abstracting for “data engine”, which brings the concept of “SDE” in recent research of database, and modifies the multi-scale data management and dispatch . “SDE” is firstly put forward by ESRI corporation. It means SDE itself can be thought a succesive spatial data model in the perspective of spatial data management, by which we can put spatial data into connection data base. Generally speaking, SDE just provides storing, reading, index, managing data and basic operations, no spatial analysis and complex operations.

In this paper, there are differences between data engine and SDE. Firstly, data engine in this paper takes charge of data storing, including in file ways and database ways. So these difference between two ways is transparant to users. Secondly, data engine should take charge of data indexing, and resolve how to query data quikly. For file storing ways, spatial index should be constructed in data engine, while for database storing ways, we can use services of commercial artery database. But for users, using data engine needs the same interface ways by whatever storing ways. Thirdly, data engine should provide the same data model that ajust to different data format and database format.

In data engine, dispatch of multi-scale data is made up of two modules: module of data dispatch and module of data extraction. The module of data dispatch is used to extract data with suitable scale and suitable type in metadatabase according to the known data area; the module of data extraction takes charge of extract appointed-area data from data file or database. So, getting the data set name from meta database, it is the key point. Here, programming process is depicted as figure 1.

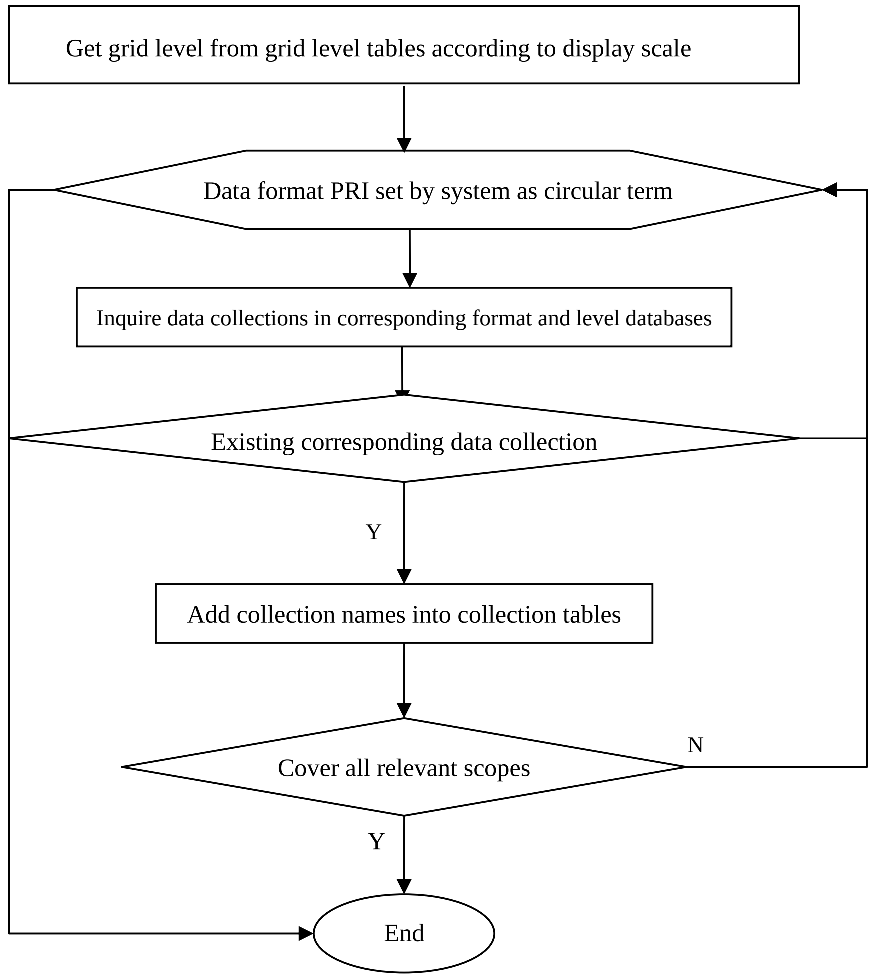

Fig.1. inquire corresponding data collection names by metadata bases

After getting the corresponding data set names according to metadatabase, we should extract data accoring to data path information stored in metadatabase. Here, programming process is depicted as figure 2:

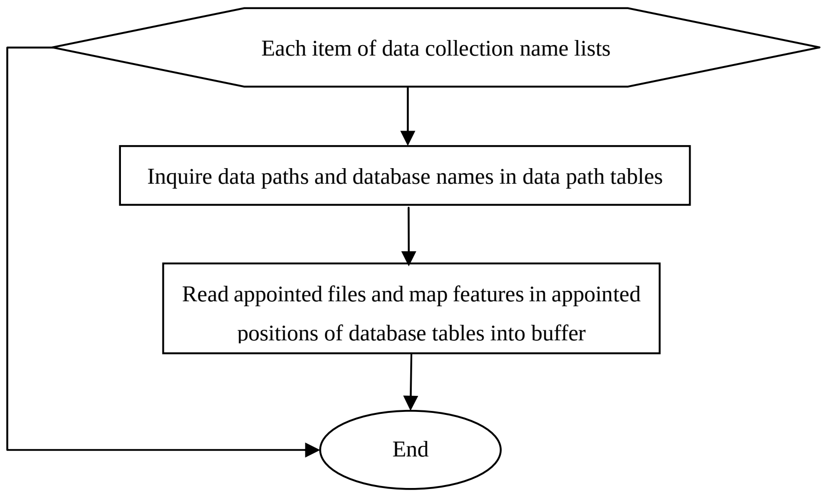

Fig.2. obtain features according to data collection names

Before reading appointed-area factor in data files and database tables, and putting them into the the internal storage, spatial index must be founded first. As it is not the main point of this paper, we would not narrate any more than needed.

5. CONCLUSION

This paper uses the idea of grid for reference to construct a meta database based on grid, which is used to manage multi-source and multi-scale data; also it lists a basic programming process of how to fulfill the visits on data by the use of metadata base. Its advantage is it can put data of different types into one grid, and at the same time, it uses potential operation of grid coding to calculate, whose speed is much higher and algorithm is much easier than that of general map sheet coding calculation consisting of letters and numbers. Considering data distributed storage, tables in metadata bases record data paths in order that it can be easily used. Based on experiments, this method can efficiently manage vector data and raster data such as images and digital elevation models, and can get a good effect.

REFERENCES

1. LJ Deren, SHAO Zhenfeng, ZHU Xinyan. "Spatial Information Multi-grid and Its Typical Application".*Geomatics and Information Science of W uhan University*. 2004, 11, p945.
2. LJ Deren, XIAO Zhifeng. Etc. "Research on Grid Division and Encoding of Spatial Information Multi-Grids". *ACTA Geodaetica et Cartographica S1nica*. 2006, 2, p52.
3. Goodchild M F, Yang S. "A Hierarchical Data Structure for Global Geographic Information Systems". *Computer Graphics*. Vision and Image Processing. 1992, 54(1), p31.

4. Dutton G. "Encoding and Handling Geospatial Data with Hierarchical Triangular Meshes". *Proceedings of 7th International Symposium On Spatial Data Handling*. 1996, p34.
5. Sahr K, White D. "Geodesic Discrete Global Grid Systems". *Cartography and Geographic Information Science.* 1 998, 30(2), p121.
6. Song Guomin. "Research and Implementation on Battlefield Geo-spatial Data Sharing". *Information Engineering University*. 2006, 4.
7. XIA Yu, ZHU Xinyan, and LI Deren. "Indexing Technology in Spatial Information Multi-Grid". *geospatial information*. 2006, 12, p4