Distribution of Large Volume Image Data Using the Mixed Mode of Java Servlet and COM on Web

WANG Mi  ZHU Xinyan  PAN Jun  CHEN Nengcheng

ABSTRACT  This paper reports some researches on distribution of large volume image data using techniques of the Mixed Mode of Java Servlet and COM on Web. The architecture and key technologies are discussed in detail. The web distribution system of image is implemented and the system is tested by the application instances. At last, the advantages and disadvantages for this web image distribution mode are analyzed.

KEY WORDS  seamless image database; WEB GIS; connection pool; COM & DCOM

Introduction

Since the 1990s, the researches and applications of the geography database based on Internet have become hot points in current computer and GIS fields. Furthermore, the concept of “Digital Earth” has received great approval by a lot of insight people. The core of building “Digital Earth” lies in the effective organization, management and distribution of all kinds of geographic information on Wide Area Network. Remote sensing image data is the important information source of all kinds of geographic information. With the development and application of satellites and remote sensing techniques, all sorts of remote sensing platforms working and hovering in the space continuously observe the earth and constantly provide all kinds of ground information of the earth. This makes the view of mankind reach the utmost extension and makes the remote sensing data find a very significant position in GIS. Therefore, Gore pointed out that the “Digital Earth” would be on the basis of 1-meter resolution satellite image. It goes without doubt that under the conceiving of “Digital Earth”, the remote sensing data will play an important role during the construction of “Digital Earth” in the future. So far, remote sensing has become an important science and technology field throughout last century. And how to process the remote sensing image data and distribute it on Internet has become a very important and urgent issue.

Referring to using an DBMS to manage large volume image data, the authors have already performed a deep research on the system structure and key technique and developed a large seamless image database management system, GeoImaeDB4.0, which is based on large commercial database (such as Oracle, SQL Server) and COM in a local area network environment. Nowadays, as for the distribution of image data on Web, many systems apply techniques such as ASP, ISAPI, Java and so on. The distribution system of image data on Web in this paper, however, is based on large seamless image database management system, GeoImageDB4.0, and...
the large volume image data management component, GeoImageDB. This paper emphasizes on the system structure and the key technique of the distribution of the large volume image data, which is based on the mixed mode of the Java Servlet and COM.

1 COM and Java Servlet Technique

1.1 Component object model (COM)

From the development course of software industry, the program design has involved in several innovations from the initial function decompounding to the structured programming, and then to the object-oriented programming, which is widely used today. The method for program design chosen by people is determined by its adaptability to the application software. The early development of software is major in the computing, so the function decompounding can meet the application requirements. As the complex degree of software increased constantly, the structured programming is always the important tool for software designer for a long time. But the object-oriented programming is entirely a leap in conception, which makes the program design fitter for the essential structure of application and plays a leading role on the whole life cycle of software production.

COM develops the object-oriented programming. It uses the object technique for the system design, and does further abstract on the course of the object-oriented programming. We can apply COM to the method for constructing the structure levels of system, and use the object-oriented programming more conveniently. The program design of COM emphasizes the true reusability of software and high interoperability. COM has the following characters.

1) The independence of programming language and developing environment. In the ideal software system, the components used to construct application system may come from different ISV (independent software vendor), and different ISVs may choose different programming language, that means the difference in calling manner and memory management. What’s more, though using the same programming language, the different developing environment (for example the compiler) will also bring inconvenience for integration of components. The criterion of COM establishes a criterion for components in binary level, and different ISVs only have to obey it and choose the appropriate programming language.

2) The location transparency of components. The location transparency is one aspect of the independence. As to the distributed application system, the location transparency can cross the computers. That is to say, the application programs in a computer can transparently visit the components in other computers. The location transparency is ensured from COM to DCOM.

3) Expandability. Every component is independent and has its own particular function. It only can communicate with the outside through interface. When a component needs to provide a new service, the new service can be accomplished through adding interfaces. And that has no influence on the users of the original interfaces. Furthermore, the new user can obtain service by choosing new interface over again.

4) Reusability. COM provides two kinds of reusable models including and aggregation. In the process of realizing the component, we can extend the existing component to get a new one in such two ways.

So, Bill Gates pointed out that; “The technique of component has been the current development of software. It is a revolution that developing the software based on component. Not only It is a new way of distributed computing but also a extensive structure supporting the concept of whole lifecycle including designing, developing and deploying”. Considering the merit of component technique as well as concerning that it is the mainstream technique, GeoImage DB4.0 developed by the authors is a large volume seamless image database management system based on component and control technique.
1.2 Java Servlet technique

Servlet is the newly added function of Java 2 developed by Sun company. It can extend the Web server function just as the CGI script but stronger, and occupies much less resource on server. So it is beyond the CGI in performance, and can be more adaptive to the rapid development of Internet. Many Web servers can support the Java Servlet such as the Apache, IIS, Tomcat and so on.

Servlet is the Java program using Java Servlet API (Java Servlet application program interface) and other correlative classes and methods. Java Servlet API defines a standard interface between Servlet and the server, which makes the Servlet have the character of the platform. When starting up the Web Server or client asking for service to the Servlet for the first time, Web server load the Servlet automatically. After that, Servlet goes on running and responds the request from other clients until the Servlet Engine on the application server is shut down. A Servlet only has one case, which serves for the multi requests in an application. Servlet comes from a series of advantages of Java such as calling back memory automatically, object-oriented character, the exception processing mechanism. Compared with CGI, Servlet has many virtues:

① It is very simple and easy to write Servlet, and debugging and maintaining is much easier than CGI.

② Servlet concurrently responds all the requests from the clients and only is there an exclusive case in the memory to run until all the clients receive the response. So Servlet does not have to load, start up, shut down the program frequently like CGI. That not only raise the running speed, but also saves the valuable memory resource.

③ Servlet is written by Java Servlet API. So it has all the virtues of Java language and does not have to care the details like CGI. Therefore, Servlet can be combined with all kinds of servers conveniently.

2 Multi-tier system architecture of image data distribution

With the rapid development of computer and network techniques, the complexity of software increases, so the system structure plays a more and more important role in the design and development of software. The early software is almost all the stand-alone software system. Usually the developers and the designers pay little attention to the system architecture. As a whole, software does not have the conception of tires, and usually an application program includes all the user interfaces, operation rules and data processing. The applied data may be stored in the local computers or in the long-distance computers, but the logic of accessing the data is included in the application program.

Because the one-tier structure can not comply with the quite complex request of application to the software now, especially in the application of the network or distributed environment, the one-tier architecture can not satisfy the requirement obviously. So the application program is naturally divided into two parts: client and server. It is the C/S mode and sometimes is called Client/Server architecture. Along with the development and popularity of Internet, the structure is transited from dual-tier architecture to tri-tier or multi-tier architecture.

The distribution system of the seamless image database on Web adopts four-tier architecture. Its system architecture is shown in Fig. 1.

In the four-tier architecture, the first tier is client. It realizes GUI of the whole distribution system on Web, including many operations of image such as showing, zooming in, zooming out, roaming and so on. Its running code can be downloaded to execute from the Web server of the second layer to the local client.

The second tier is Web Server. It mainly answers for the responding to the request of HTTP sent out by the browser on the client, and sends
the request to the application server to process.

The third tier is Application Server. It mainly answers for the request of the Web server, communicates with the database, completes processes the operations of the seamless image database such as dispatching, merging and so on, and transmit the result to the Web Server.

The fourth tier is the Database Server. It is made up of the database systems, and mainly answers for the work such as storage and management of the image data, safety certification to the user, and so on.

In addition, what is directly connected with database is GeoImage DB4.0, a seamless image database management system, which is a developed system based on the C/S mode and is mainly responsible for accomplishing the establishment and maintenance work of seamless image database on server.

3 Key techniques of image data distribution based on the mixed mode of Java Servlet and COM

This paper mainly researches the following two key techniques of the image data distribution based on the mixed mode of Java Servlet and COM.

3.1 Mixed call technique of Java and COM

The mixed call technique of Java and COM researched by this paper is implemented by using Java. The so-called native methods indicate that the methods do not use the Java language, but other languages of the native environment. Using some instructions provided by Java and the system library together can make Java call the native methods. The principle of calling the native methods of Java is implemented by the mixed programming of Java and the language used to realize the native methods. The mixed call technique of Java and COM researched by this paper will use the mixed programming of Java and C. The basic principle of the mixed programming of Java and C is to realize the connection between Java and C. The connection is realized by DLL. As we all know, DLL is an executable piece entity unit, it is available to all the application programs based on the same environment, and when the application programs run, DLL can run as a part of the application programs. Therefore, so long as insetting the DLL realized by C code into Java can reach the purpose of the mixed programming. The function of calling COM in Java can be realized by the way of packing a DLL which can be called by Java based on the components. Then COM can be called by Java through DLL. Due to Java is a network programming language, so the native methods are forbidden to be called by Java Applet, which can only be used in the application program of Java. Java 2 provides the technique of Servlet, which is a small program on server. Servlet can run on the application server, call the DLL executable code configured on server, and then call components on server through DLL. The implemen-
tation procedure of calling the native methods by Java is described as follows.

1) Writing the Java classes with the native methods and mapping the methods provided by the components as the native methods of Java.

```java
public class ImageDB {
    public native static boolean startDbServer();
    public native static void getInitParam(double[] rect);
    public native static void getInitResolution(double[] resolution);
    public native static void getImageData(double x0, double y0, double x1, double y1, int width, int height, int[] imagePixels);
    public native static void closeDbServer();
    static {
        System.loadLibrary("ImageDbServer");
    }
}
```

2) Use the javah, exe tool and make the native methods of Java be compiled as the header files which can be recognized by C. The result is

```java
JNIEXPORT jobject JNI CALL Java_ ImageDB_startDbServer (JNIEnv *, jclass);
JNIEXPORT void JNI CALL Java_ ImageDB_getInitParam (JNIEnv *, jclass, jdoubleArray);
JNIEXPORT void JNI CALL Java_ ImageDB_getInitResolution (JNIEnv *, jclass, jdoubleArray);
JNIEXPORT void JNI CALL Java_ ImageDB_getImageData (JNIEnv *, jclass, jdouble, jdouble, jdouble, jint, jint, jintArray);
JNIEXPORT void JNI CALL Java_ ImageDB_closeDbServer (JNIEnv *, jclass);
```

3) Use Visual C++ to build a DLL project and make the above functions be the output functions of DLL to build the ImageDbServer. DLL that is a dynamic link libraries. The concrete function of each function is realized in the way of calling components through DLL.

Through the above steps, calling the function provided by components in Java Servlet can be realized after compiling the native method classes of Java as the byte code of Java.

### 3.2 Connection pool technique

In the applications of Web, it is necessary to frequently access database because the visits come from many users. In order to visit database, the connection to the database should be established for every network customer. As for the database, the connection to database is an time-consuming operation. Establishing a connection will cost a few seconds at least, furthermore, each connection to the database will consume the memory and the connection to the network also needs to be preserved. Finally, closing the connection will cost the spending on the callback of garbage. In this way, both the time and the system resources waste a lot. Therefore, a better solution is to build a middle component named connection pool to manage all the connection between the network consumers and the database. When they need to visit the database, consumers apply a connection to the connection pool, the database operations are executed, and then the connection is released. But the connection pool does not close the idle connections. It provides the idle connection for other consumers. Because the database connection on Web researched by this paper is realized by the components, all the database connection must be managed by themselves. The system adopts the connection pool technique and realizes the multi-user concurrent visit to database.

Fig. 2 illustrates the principle of the connection pool. When the system is initialized, connection pool will also do some corresponding initialization work. Then when the users send out the visit request, the management program of the connection pool search whether there are idle connections. If there are idle connections, a connection is assigned to the user; if there are not, then wait until other users who occupy the connections release the connections. When some users release a connection, the connection pool will call back the connection for other users to use.
4 Implementation of image distribution system on Web

According to the method suggested by this paper, the distribution of the large volume image data on Web is realized by using the mixed mode of Java Servlet and COM, which is based on the large volume seamless image database management system component, GeoImageDB.dll. The following is an application case of the system in the distribution of the image data of Beijing.

Fig. 3 is the case that the seamless image database of the aerial image is browsed online in a special network. Images in image database are aerial image of Beijing. The image files add up to 1,767, whose fundamental scale is 1 to 2,000 and whose ground resolution is 0.4 m. In order to view the whole image of Beijing, the system adopts multi-scale pyramid data structure to organize the data. The total image data of the system amount to 45.7 GB. The system is based on the Oracle 8.16 as database platform, and stores image data in binary field. And the system makes use of the special network of Beijing government, which makes it be realized to browse image real-time online and query the location according to place name.

5 Conclusion

This paper mainly discusses the distribution of the large volume image data on Web, which is based on the mixed mode of Java Servlet and COM. This distribution mode on Web can utilize the former production enough, reduce the workload of coding and testing, and shorten the development cycle of the software. Also the core of the system is realized by components written in C++. Compared with using Java and other modes, this mode is much faster and can guarantee the running efficiency. That is necessary to the distribution of the large volume image data. However, as for the transplantation of the program, this mode, to some extent, damages the characteristic of Java that “write once, run any-
where”. This is its shortcoming. But the authors still consider that using this mode to develop software on Web is a good solution.

REFERENCES
1. Dong H B, Zhang H F (2001) Developing Web database using Servlet. Computer Application Research, 2(2):100-102 (in Chinese)
2. Wang M (2001) The development of large-scale seamless image database (GeoImageDB) and the feasible research on measurable virtual reality; [Ph. D dissertation], Wuhan, Wuhan University. (in Chinese)

4 Conclusions
This paper describes an approach to find out the effective feature subset using genetic algorithm. Our application is in the domain of aerial images. The values of texture features are continuous data. To select effective feature subset, the fitness function is designed in this paper. Besides, on the basic of the effective feature subset selected, this paper presents an approach to extract the objects which are higher than their surroundings, such as trees or forests. The experiment results show that the feature subset selected and the method of classification are effective and practical. Although we have dealt with color aerial images, the technique is extensible to remote sensing images.

REFERENCES
1. Liu H, Setiono R (1996) A probabilistic approach to feature selection—a filter solution. Machine Learning ; LCML’96, Bari, Italy
2. Yang J H, Honavar V (2001) Feature subset selection using a genetic algorithm. Journal of Intelligent Information Systems, 16:215-228
3. Vafaie H, Jong K D (1993) Robust feature selection algorithms. International Conference on Tools with AI, Boston, Massachusetts.
4. Dang T, Jamet O, Maitre H (1994) Applying perceptual grouping and surface models to the detection and stereo reconstruction of building in aerial imagery. International Archives of Photogrammetry and Remote Sensing, Atlanta, USA.
5. Haala N (1994) Detection of building by fusion of range and image data. International Archives of Photogrammetry and Remote Sensing, Atlanta, USA.
6. Eckstein W, Steger C (1996) Fusion of digital terrain model and texture for object extraction. 2nd Airborne Remote Sensing Conference, San Francisco.
7. Pan L (2001) The Study on forest area recognition from color aerial Image and its application to automatic aerial triangulation; [Ph. D dissertation], Wuhan, Wuhan University. (in Chinese)