

## Researches on Distributed Image Database

GOU Pingzhang

Northwest Normal University  
Lanzhou, China, 730070  
E-mail: goupz@nwnu.edu.cn

TANG Yongzhong

Hexi University  
Zhangye, Gansu, China, 734000  
E-mail: tangyz@hxu.edu.cn

**Abstract**—On the basis of the image data characteristics, this study contrasted four kinds of data model. Then it analyzed the three kinds of implementation methods. It contrasts the management modes of the distributed image database finally.

**Keywords**- distributed, data mode, image database

### I. INTRODUCTION

With the development of the computer and network, the database develops from miniaturization to large size, and also from centralized mode to distributed mode. Distributed database technology is the combination of distributed technology with database technology. Conceptually, distributed database is a data collection physically dispersed in computer network nodes and logically belonged to a same system. It has two major characteristics: the distribution of data and coordination between the databases. The system emphasizes the node autonomy, not centralized control. In the same time, the system should keep the transparency of the data distribution. Then the data distribution could not be considered when programming. Compared with the text data, the image database has its own characteristics<sup>[1,2]</sup>.

The image content is hard to be accurately described. The image content contains a group of space objects<sup>[3,4]</sup>, such as point, line, surface and the relationships between them. They could only be described to image entities and their characteristics. While the text data could be described by the basic data types, such as integer, string or the complex data type composed by the basic data types. Otherwise, the image database needs multiple description of the same feature of the image entity<sup>[5]</sup>.

The space objects and its relationships contain no semantic information themselves. On the contrary, the name of the text data entity and relationships contain semantic information. But it would seriously constrain the application of the image information by combining the semantic, the space subject and relationships directly. The reason is that there may be different interpretations and different usage modes in different period in a same image. Furthermore, the image could only be approximately explained and the image interpretation develops with the image recognition technology.

The image database requires supporting image query and spatial reasoning<sup>[6]</sup>. The image has multiple uncertain interpretations. So the system should supply knowledge of

specific field. The aim is to forming complex statements composed by simple queries.

### II. DESIGN GOALS

From the database standpoint, the Data model can be regarded as the real world specification language. It describes the content, structure and Information organization from different stand points. It is the key of the database system modeling. It experienced different history levels of hierarchical model, network model, relational model and object-oriented data model. It is different between multimedia data and traditional data. So a good multimedia data model<sup>[7]</sup> should have the following characteristics:

1) Higher abstraction mechanism

On this basis, the complicated data modeling could be simplified and standardized.

2) A complete description of data object, links between objects and the semantic nature of the data object.

3) Version management of the multimedia data. In the multimedia database, the vision establish and control, and informs of the vision changes are important. The vision has two layers of meaning: the one is vision history and the other is the vision choice. The vision history is that there are different contents in different time. For instance, there are sketch and Official drawing for the CAD (computer-aided design) design drawing. The vision choice is that there different description and processing for the same object. For instance, there are English and Chinese versions for the same contract text. It need solve the identification, storage, updating and query of multiple visions.

4) The information from different sources is integrated together. The information of the multimedia documents comes from different sources, such as the text database, image database, sound database and etc. Different tools were used to edit information. Finally the information is integrated. The integrated information has good extensibility. Object-oriented technology is best suited to deal with such problems.

5) The internal model and external manifestations of a subject are better distinguished. A multimedia subject has different manifestations. For instance, a same subject could be represented by text, sound or image.

### III. DATA MODEL

Data model of database can be divided into four categories:

A. *Traditional data model*

Traditional data model developed in 1960s. It includes network model, hierarchical model and relational model three kind.

B. *Semantic data model*

Semantic data model is a user-oriented model. It uses a natural mechanism to describe database design. It is not restricted by specific implementation structure. Semantic model supplies cogent abstract construct mechanisms. It also supports derive, such as summarization, polymerization, joint, classification and etc. It can exactly express semantic relationships between the data. The main semantic data models are E-R model, event model, function model and etc.

C. *Hypertext data model*

Hypertext data model is a fast developed data model along with the multimedia technology. It uses the nonlinear mesh technology to organize, represent and manage the discrete information and the relationship between them.

D. *Object-oriented data model*

Object-oriented data model makes full use of the characteristics of the object-oriented theory. The object-oriented theory has the characteristics of data abstraction, functional abstraction and messaging. On the base of these characteristics, the objective world is divided to systems according to objects. These objects are independent and have better encapsulation and concurrent processing capability.

The comparison is summarized as table 1:

TABLE I. TABLE I COMPARISON OF THE DATA MODEL

Data model type	Data representation entity	Links between data	Correlation mechanism
Relationship Type	Table, Record, Field	Table	Associations between different databases are supplied by primary keys or indexing mechanisms
Object-oriented model	Object, Property, Method	Object	Encapsulation mechanism is used to encapsulate the data and attributes of the object. Objects link via messages.
Hypermedia Model	Node	Chain	Objects link via chain mechanism

IV. IMPLEMENTATION OF THE IMAGE DATABASE

At present, there are three kinds of implementations of the multimedia database data model<sup>[8]</sup>.

A. *Model I*

Relational model is extended and changed to realize the management for the non-format data and extension for the relational database. This kind of implementation can be divided to the following three kinds of methods:

1) *Expand the existing data type*: On the basis of the conventionally used data type (such as numeric, string, Boolean, date type and etc), the data types of image, figure, sound and video are added.

2) *Expand the user-defined data type*: A user-defined abstract data type can also be used as the value of another relationship column, not only just the system defined data types.

3) *Expand nested semantics by NF2 data model*: Nested relational data model (NF2) breaks the norm form limit. The table can be embedded in the table. NF2 absorbed the advantages of both hierarchical model and object-oriented model hierarchy. It supplies means that describe property nested definitions. It reflects that the value of an object is also an object.

B. *Model II*

Object-oriented model realize the object oriented multimedia database management system. The object-oriented database (OODB) is the database that introduces the object-oriented data model and method. OODB is a kind of scalable data model. It absorbs abstraction mechanism in semantic data model. It also adopts the idea of object-oriented programming language and abstract data structure. So it can better describe complex objects and maintain semantic information<sup>[9]</sup>.

1) *The characteristics of the object-oriented multimedia data model*

Object-oriented multimedia data model has the characteristic that the objects with the same static structure, dynamic behavior and constraints are bracketed together in the object-oriented database. The same type of objects is defined corresponding property, method and constraints. Object-oriented model has the package, class, inherit and overloaded characteristics. These characteristics are available for the multimedia data processing. They support data type inheritance. They add the reusability of code and inherent dynamic contact ability of management data. Compared with the traditional data model, the object-oriented data model has the following advantages:

a) The object-oriented data model could supply complex modeling capabilities and scale data management.

b) The object-oriented data model could nest more semantic. It also supports abstract data types and user-defined method. So it is easy to support new data type and its operation.

c) Package and information hiding supply modular mechanism.

d) The characteristic of succession reduces redundancy and supplies software reuse mechanism.

e) The object-oriented data model could supply unified user interface for different media.

2) *The realization of object-oriented multimedia database*

One method is to develop a new data model and realize OODB system from underlying. A data model with the object-oriented database conception is established firstly. The corresponding language and the core of the object-oriented database management system are designed. The hierarchy and inherited characteristics of object, method, property, news and object class in the object-oriented method are adopted to describe multimedia data model and design corresponding language. It has the advantages of clear system structure and high efficiency. Its defect is large difficult. On one hand it lacks of unified data model and formalized theory, and on the other it has no tools of query optimization, view and database.

The other method is to embed database function in the object-oriented language to form the OODB. The complexity of the multimedia object determines that the object-oriented method is the best method to realize multimedia data base management system (MDBMS). The key is how to integrate the OO programming language and multimedia database language together. That is designing a real OO multimedia database language. One solution is to add persistent object storage management in the OO language. Then the language would support the class persistence, concurrency control, recovery mechanism and etc. Furthermore, if the concept of persistent object is introduced to object-oriented programming language, the database operations function could be added to the language. On the basis of type system and programming mode of object-oriented programming language, the language introduces mandatory data to get lasting and shareable data structure mechanism. Then it will realize persistent object data reusing in the language. Generally, the expansion of the existing program design language is to add the abilities of the database management system, such as persistent, concurrency control, database query and etc. The OODB systems adopted this method are

Gem-stone, Orion, Ontos, Exodus, Objectstore, Versant and etc.

C. *Model III*

The method that hypertext/hypermedia model organize and manage hypermedia is to combine multimedia technology with database technology. It could realize the database management and operation based on the multimedia. The basic structure of the hypermedia model is net. It is the digraph consisted by node and chain. Node and chain are two core concepts of the hypermedia model. Hypermedia stems from hypertext. Hypertext is a text information management technology. It stores the text information in many nodes. Node is the basic unit of information. These nodes are connected by chains to compose a net structure. It used to express the information relations. Hypermedia is the multimedia formation of hypertext. In other words, the multimedia storage, management and retrieval are on the basis of the hypertext concept.

V. DISTRIBUTED DATABASE MANAGEMENT

In the distributed database, the data storage is accessed by the following three methods:

- *Replication*: Several identical copies used by the system to maintain relationship. These copies are stored in different nodes.
- *Segmentation*: Relationship is divided into several fragments. These fragments are stored in different nodes.
- *Replication & Segmentation*: Relationship is divided into several fragments. The system maintains several copies for each fragment.

Data replication is a technology used in the distributed database. The Oracle follows TCP/IP protocol and supports multi-protocol internetworking. On the basis of Net tool in Oracle, the data replication can configure the LAN which follows different protocols. In the LAN, the database servers can communicate with each other. Then the data link is created to realize data replication function.

According to system requirements, there are two kinds of data synchronization mode. They are transaction replication and merge replication.

- *Replication control method*: The copy values in different sites in any time are the same.
- *Asynchronous replication method*: Transaction log method and trigger method are adopted to capture system update operation

The summary is shown as table 2.

TABLE II. TABLE 2 DATA CONSISTENCY METHOD

Data consistency method	Features	Limitation
Traditional transaction control method	Mature theory, it applies to general distributed database system	It does not apply to the distributed database that has large amount of data and affairs need long time to maintain.

Extended control method	transaction	High autonomy: It does not depend on the whether the local database support 2PC	Low efficiency of transaction execution
Synchronous control method	replication	It is applied to real-time data synchronization distributed database system. It is used to maintain the tight consistency of the data copy.	Frequent data update and low system processing speed
Asynchronous control method	replication	It is applied to the space distributed database system. It is used to maintain the loose consistency.	Imported data may not be the latest. It is not applied to the distributed database system that requires real-time data synchronization.
Message queue method		It is applied to Internet distributed database system.	The distributed database system that requires real-time data synchronization.

## VI. CONCLUSION

For the rapid increasing of the data value, the traditional centralized database can no longer meet the demand. The centralized database has low reliability. If the CPU breaks down, the entire system would collapse. The distributed database has solved this problem. This study combined the image database with the distributed technology together. It also summarized and compared the models and implementations of the image database. The study shows that the distributed image database technology meets information systems application requirements. The image database technology meets the management thinking and management style of today various industries. It has great application prospects.

## REFERENCES

- [1] P. Stanchev. General Image Database Model. Visual Information and Information Systems. Lecture Notes in Computer Science 1614, 1999, 1614: 29-36.
- [2] Tamura H, Yokoya N. Image database systems: a survey [J]. Pattern Recognition, 1984, 17(1) : 29-43.
- [3] Mao R, Miranker D, Sarvela P, et al. Clustering sequences in a metric space-the mobios project[C]. Edmonton, Canada: Poster of the 10th International Conference on Intelligent Systems for Molecular Biology, 2002, 3-7.
- [4] Stanchev, P. General Image Data Model. 22 intentional conference: Information Technologies and Programming, 1997, 130-140.
- [5] Yan, C.W., Dimitroff, D.C., Arndt, T. An intelligent image database system. Software Engineering, 1998, 14(5): 681-688.
- [6] Nepal S, Ramakrishna M V. Query processing issues in image (multimedia) databases [C]. The 15th International Conference on Data Engineering. 1999: 22-35.
- [7] Matthew Gast. Wireless Networks The Definitive Guide [M]. Beijing: TSINGHUA University Press.
- [8] Mario Döller, Harald Kosch. The MPEG-7 Multimedia Database System (MPEG-7 MMDB). Journal of Systems and Software. 2008, 81(9): 1559-1580.
- [9] Yuri Leontiev, M. Tamer Özsu, Duane Szafron. On type systems for object-oriented database programming languages. Journal ACM Computing Surveys (CSUR). 2002, 34(4): 409 – 449.