Multimedia Databases

*V Parameswari**

Multimedia Databases allow users to accumulate and query different types of multimedia information, which includes images, video clips, audio clips and documents. Storage and retrieval from the conventional databases like file server and relational database are inadequate for the multimedia data because of high disk I/O usage, which is a potential bottleneck for most of the multimedia database systems. The two aspects of multimedia systems, which are essential for eliminating potential problems in the data transfer, are the necessity for the buffer management and disk access methods. Multimedia data can be retrieved using content-based retrieval and query-by-image content. Well-defined query and Fuzzy query are the two types of queries in the database systems. This article discusses the nature of applications, document interpretation, data store, retrieval and issues involved along with the different query languages and the new products for the multimedia databases.

Introduction

Multimedia Databases provide features that allow users to store and query different types of multimedia information. The challenges involve locating multimedia sources that contain certain objects of interest. Applications of multimedia systems include static media like images, text, graphics and dynamic media like audio and video, which varies with time. Timeliness is required to display data within certain duration of time for continuity and synchronization to present in consistent view through proper data integration to users. The two aspects, which are essential for eliminating potential problems in the data transfer are—necessity for the buffer management and the disk access methods. New multimedia database systems could benefit such areas as concurrency control data abstraction transaction recovery, openness and access control. Further improvements are possible by changing the way we handle multimedia data, such as database management capabilities like searching, sharing and retrieving of data to meet the needs of various multimedia applications. The Web has wildly proliferated software for manipulating multimedia data. Some specific applications (existing) and future applications need to live with multimedia data. This trend is expected to go up in the days to come.

Nature of Multimedia Applications

Multimedia data available in current systems involves Text, which is formatted or unformatted. Graphics include drawings and illustrations that are encoded using some descriptive standards. Images include drawings. Photographs encoded in standard

* Faculty Member, The ICFAI School of Information Technology (ISIT), Hyderabad. Email: paramesha_2000@yahoo.com

formats such as bitmap, JPEG and MPEG. Animations are temporal sequences of image or graphic data. Video is a set of temporally sequenced photographic data for presentation at specified rates. Audio is a sample data generated from aural recordings in a string of bits in a digitized form.

- **Repository Applications:** A large amount of data as well as meta data is stored for retrieval purposes. Examples include repositories of space photographs, satellite images and radiology scanned pictures.
- **Presentation Applications:** Many applications involve multimedia data subject to temporal constraints. Audio and video data can be delivered by optimal viewing or listening conditions. They require the database management systems to deliver data at certain rates offering "quality of service" above a certain Threshold. Complex and interactive multimedia presentations involve directions to control the retrieval order of components in a series or in parallel. An interactive environment supports capabilities such as real-time editing analysis or annotating of video and audio data.
- **Collaborative Work Using Multimedia Information:** A new category of applications in which users may execute a complex design task by merging drawings, fitting subjects to design constraints, and generating new documentations, change notifications. Intelligent healthcare networks as well as telemedicine will involve doctors collaborating among themselves and analyzing multimedia patient data and information in real time as it is generated.

> Multimedia data available in the current systems involves Text, Graphics, Images, Animations, Video, Audio etc.

Requirements

Multimedia Database Management System (MMDBMS) should handle image, voice, and other multimedia data types, handle a large number of multimedia objects, provide a high-performance and cost-effective storage management scheme, also support database functions such as insert, delete, search and update. Multimedia objects are mostly Binary Large Objects (blobs). The following are the characteristics for multimedia data that are to be handled by MMDBMS.

- **Volume** of data. For example, one compressed video frame typically exceeds 10 kb. Consequently, 30 frames per second require a volume of over 300 kb/s far more than that required by typical relational queries.
- **Continuity** within each video or audio stream (i.e., interframe and intersegment time constraints).
- **Synchronization** of multiple streams of data (e.g., Video, closed caption, left speakers's audio, right speaker's audio). It should be noted that individual streams of data should not be viewed as separate clients.

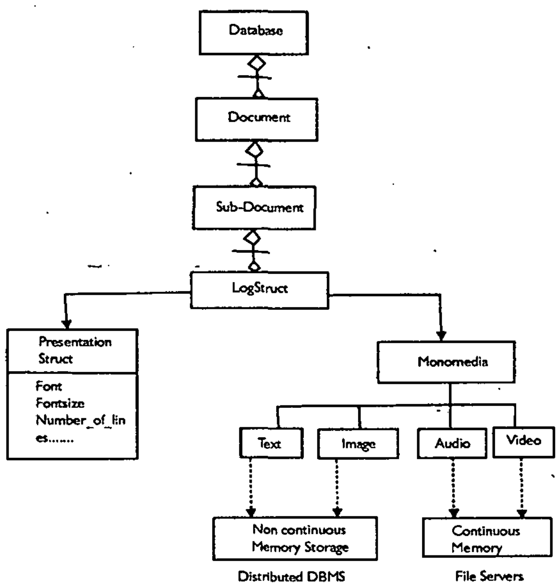
- **Qualities of Service**—Multimedia resources require a guarantee of the presentation quality. The file structure and program to guarantee the quality of service should be created.
- **Networking**—A distributed database requires locking mechanisms for concurrent access controls. Using our own low level design, it is easy to control some implementation issues such as the two-phase locking mechanism, and the database administration and control of traffic on the network.

Documental Representation in Database

Structural approach includes logical structure and presentation structure. Logical structure shows the actual layout of components that is displayed to the user, e.g., the document can be represented by a hierarchy of chapters, sections, paragraphs and so on. On the other hand, presentation structure illustrates actual information about the document such as font and font size for each section, number of lines in the document and so on.

Figure 1 depicts a structure for multimedia data of a database.

Figure 1: Documental Representation



Document interpretation

The Documental Representation is interpreted as below.

Image is stored either in a raw form as a set of pixels or cell values or in a compressed form to save space. The image ***shape descriptor*** describes the geometric shape of the raw image, which is typically a rectangle of cells of a certain width and height. Each cell contains a pixel value that describes the cell content. To identify objects of an image, it is divided into homogeneous segments using a homogeneity predicate. Cells that are adjacent to one another and whose pixel values are close are grouped into a segment. The homogeneity predicate defines the conditions for how to automatically group those cells. An image database query would be to find images in the database that are similar to a given image. The given image could be all isolated segment that contains a pattern of interest and the query is to locate other images that contain the same pattern.

Video is divided into video segments, which are in turn made up of a sequence of contiguous frames that includes the same objects/activities. Its starting and ending frames identify each segment. The object and activities Identified in each segment can be used *to* index the segments.

Text /document is a full text of some article, book or magazine. These sources are typically indexed by identifying keywords that appear in the text. Filler words are eliminated from that process. Techniques have been developed to reduce the number of keywords to those that are most relevant to the collection because there could be too many keywords when attempting to index a collection of documents.

Audio includes stored recorded messages, such as speeches, class presentations or recording of phone messages. Discrete transforms can be used to identify the main characteristics of certain person's voice in order to have similarity based indexing and retrieval.

Automatic analysis and Manual analysis are two methods of including metadata in a multimedia database

Meta Data

An example of metadata is a video of a lecture. This record would contain metadata about the lecturer (who they are, which university/company they work *for*). The content of the lecture, where it took place, when it took place, how long it lasted, who it was to and so on. This information is then used when people search the database for a specific lecture, or for lectures given by a specific person. Automatic analysis and Manual analysis are two methods of including metadata in a multimedia database. Although manual analysis is more effective, as it allows human intelligence to decide what data is important and what is not, it is costly in terms of worker time and ensuring correctness but it is a standard way of classifying data. On the other hand, automatic analysis is much faster, but it is still a developing art, and often difficult to get the same level of description as a human being would provide. The main advantage is that automatic analysis provides a consistent description for all data and is not affected by individual styles. This makes it easier to search for information when the classification system has been learnt but it is not as intuitive to the first time user.

Multimedia Data Store

Stored multimedia data is classified into two types: Non-continuous media such as text and image and continuous media such as audio and video with real-time constraints. Most of the continuous media are stored using separate storage server in order to meet the real-time constraint requirement but non-continuous media are stored in the database with meta-information about the files on the continuous media.

Multimedia data uses Standard Generalized Markup Language (SGML)/ Hypermedia Time-based Structuring Language (HyTime) to represent documents in order to standardize the data storage. Standard Generalized Markup Language is a meta-language that describes the logical structure of a document by using markups, which show where the boundaries of the logical elements are in the document. HyperText Markup Language also uses the principle of SGML. HyTime uses SGML representation syntax, and it deals with structured information and the ability to link from and to structured information. HyTime is able to represent scheduling and rendering information. Multimedia systems have a variety of static and active information of a different nature. Besides query-based search, multimedia information access is usually supported by sophisticated navigation capabilities, similar to those found in hypertexts, thus leading to an extension commonly called HYPERMEDIA.

Data Retrieval

The Data retrieval in the multimedia databases is as follows.

Content-based retrieval is based on its certain objects or activities. Content-based queries are often combined with text and keyword predicates to get powerful retrieval methods for image and multimedia databases. For example, if you want to locate all video clips based on certain person.

Query-By Image Content (QBIC) is online collections of images growing larger and more common, and there is a need for tools for efficient management, organizing, and navigation through them. QBIC lets you make queries of large image databases based on visual image content like properties such as color percentages, color layout, and textures occurring in the images. Some tools for searching images are Color, Shape, Spatial relationships, Line, Texture and Motion

Schemes in Buffer Management. Buffers were used for the proper synchronization of parallel continuous media streams. The following are the two schemes of Buffer management.

- **Use-And-Toss(UAT)** is a popular scheme used in the database system and is good only if the data sharing could not be done. It uses the FIFO policy with the timestamp mechanism to decide which of the segment from the free pool is to be allocated to a new request and to use it for tossing so that the same buffer is used for the next data required.
- **Continuous Media Caching** is a simple scheme, which improves the buffer management by planning data sharing for future access. The basic notion of this

same topic, when some data items are shared between different topics and when different languages ask for same news item.

Sharing (SHR) uses the UAT approach in order to decide whether a new request needs to read data from the disk or could be shared with existing one, otherwise new designs were needed. It was enhanced to handle continuous media caching. This method not only reuses the data but also sharing with the future data.

```
do forever
      do for all active sessions
                if buffers are shared with another sessions
                          if leading user
                                    preserve the used buffers for logging user;
                                    acquire new buffers from the free pool;
                                    fetch data from disk into the buffer;
                          else if lagging user
                                    if there are more lagging users
                                              preserve the used buffers for lagging users;
                                              read from preserved buffers;
                                    endif
                          endif
      else      when a new segment is encountered
                                              return the used buffers to the free pool;
                                              acquire new buffers from the free pool;
                                              fetch segment from disk;
                          endif
      end do
      end do
```

Buffer Allocation and Replacement Scheme for Sharing

method involves buffers that have been played back by a user to preserve in a controlled manner for use by subsequent users requesting the same data. This approach not only saves the disk bandwidth but also improves the performance. Three different occurrences where sharing can occur are, when there are multiple requests for the

Query Languages

Multimedia query languages must deal with complex spatial and temporal relationships inherited in the wide range of multimedia data types. Support for semantic data retrieval, indexes on keywords and contents of multimedia objects must be present. There are two types of queries that are used in the database system.

- **Well-Defined Query (wdq)**

The objects are represented by a well-defined set of labels. The conditional operators are also well defined. Here the user must have the exact knowledge of the underlying database and the desired query result.

- Fuzzy Query

The properties of fuzzy query objects are unclear or comparison operators that do not give exact matches and they result in sets of different answers.

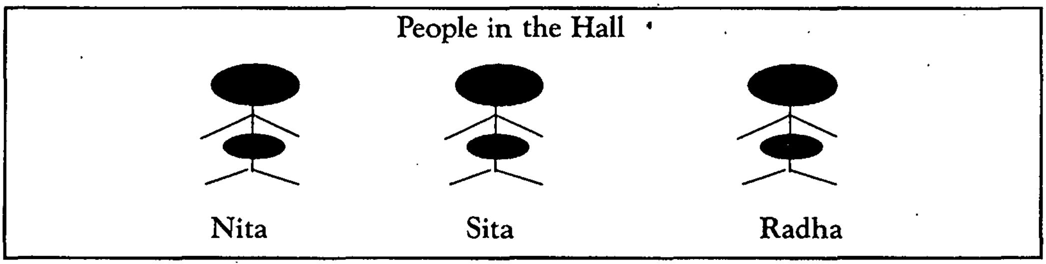
Multimedia querying can be ***Keyword querying***, which constructs indexes for keywords to find the answer to the request and uses only Well-Defined Query. ***Semantic querying*** and ***Visual querying*** use fuzzy queries and sometimes they are used in well-defined queries also.

In keyword querying, there is no single proper word to describe the image sometimes and keywords cannot preserve all the spatial and temporal relationships, which can result in information loss. A typical visual query is the Query by Image Content (QBIC) and query by icon. Semantic query is the most difficult one in terms of indexing, pattern matching and accessing structure.

Illustrations for Well-Defined Queries

Multimedia database systems must accommodate well-defined queries as they support traditional DBMS functionalities.

The well-defined query: **"Is there any image such that Radha is at the left of Sita".**



EVA is one of the early query languages designated for multimedia information systems. And it has its conventional set theory, which uses the mathematical framework.

Sample Query:

{(title(PIC) | (Image PIC) and PERSON PERSON2 and PERSON isin object(feature(PIC)) and PERSON2 isin object(feature(PIC)) and Name(PERSON) is "Radha" and Name(PERSON2) is "Sita" and left(PERSON2,PERSON))}

EVA is OO and supports objects, classes and relationships between objects. It also supports set operations like isin, union, intersection and so on. It currently does not support video data.

Similarly **GARLIC** is an object oriented middleware and object-based management system developed at IBM Almaden Research Center. QBIC has been successfully integrated into Garlic to provide Query processing and manipulation services extending standard SQL with additional capabilities.

OMEGA (Object-Oriented Multimedia Database Environment for General Application) is an object-oriented database system for managing multimedia data that is under development at the University of the Library, Japan. In this model an acceptor is defined to allow user to communicate with the system, which recognizes to describe a real world entity by using any symbol system that is allowed in specified media. A collection of designer's impressions about the real world in terms of objects is organized as a kernel object base for integration.

FBQL(Frame Based Query Language) is a first and formal theoretical framework for characterizing multimedia information systems to provide a logical query language that integrates diverse media and is independent of any specific application domain. Also it provides the possibility of uniformly incorporating both query languages and access methods based on multimedia index structures. A special data structure called 'Frame' is used for data accessing.

OVID(Object Oriented Video Information Database) is an object oriented video model. It introduces the notion of a video object which can identify an arbitrary video frame sequence. An inheritance based on an interval inclusion relationship is introduced to share descriptional data among video objects. The hierarchical structure of a video would be described by a series of derivations but not by the composition. Queries are processed by Video Chart and Video SQL. Video SQL in OVID system deals with video media only.

The closeness of two multimedia objects should be calculated based on a weighted average of the closeness of their respective components

PICQUER Y + is a knowledge-based object oriented query language, which is a high-level domain-independent query language designed for image and alphanumeric database management to allow users to specify conventional arithmetic queries as well as evolutionary and temporal queries. A template technique has been used in PICQUERY + for user accesses. Query templates in PICQUER Y + are used to specify predicates to constrain the database view.

Illustrations for Fuzzy queries

Fuzzy queries are used to find requested objects with an intention to relax search conditions. Each feature should have its own tunable notion of what a close match means. The closeness of two multimedia objects should be calculated based on a weighted average of the closeness of their respective components. Example query is "Find any video that includes Jean Chretien shaking hands with Bill Clinton in an airplane in the background that looks like Doeing 767". The query is fuzzy because the fuzzy operator looks like and uncertain attributes shaking hands and background. The SQL based query languages cannot answer this query, such kind of scenes are difficult to describe by just using words. It is common to post a query such as "find all the images that look like this one" to an image database when a face has been sketched by an artist. In this scenario, what we want to retrieve from the database is not the same as the object in the query, but something that is similar to that image.

SCORE(System for Content based Retrieval of pictures): A visual query interface for image databases and makes use of refined Entity Retaionship model to represent the content of images. A unique feature of SCORE fuzzy query processing is its novel definition of object similarity. Two objects are similar if both of the following points are satisfied:

1) The names of the objects either the same or they are synonyms, or two objects appear in an IS-A relationship hierarchy.
2) The attribute values of the two objects do not conflict.

An image may be divided into nine regions such as NorthEast, SouthEast etc. If two attributes are in the neighboring positions or same region then they are said to be similar or else they are in conflict.

Extended FBQL is a special feature to support fuzzy queries and a method of relaying a query when the original query does not have an answer. If there are two features f1 & f2 then fl < f2 indicates that feature fl is considered to be a subfeature of f2. Functions to determine what constitutes a "possible" answer to a query and to retrieve an alternate image that is similar image when exact image is not present are utilized.

Standardized software will be required to deal with synchronization and compression/ decompression along with the indexing problems

Issues

Multimedia applications dealing with thousands of images, documents, audio and video depend critically on appropriate modeling of the structure and content of data and then designing appropriate database for storing and retrieving multimedia information. The issues include:

- **Modeling**—This area has the potential for applying database versus information retrieval techniques to the problem. There are problems of dealing with complex objects made up of wide range of types of data: Numeric, text. graphic, audio stream and video sequence. Documents constitute a specialized area and deserve special consideration.
- **Design**—The conceptual, logical and physical design of multimedia databases is to be considered as issues. The design can be based on the general methodology, but the performance and tuning issues at each level are far more complex.
- **Storage**—Storage of multimedia data on standard disk like devices presents representation, compression, mapping to device hierarchies and buffering during the input/output operation. In Database management Systems, a "blob" facility allows untyped bitmaps to be stored and retrieved. Standardized software will be required to deal with synchronization and compression/decompression along with the indexing problems.
- **Queries and retrieval**—The database way of retrieving information is based on query languages and internal index structures. The "information retrieval" way strictly relies on keywords or predefined index terms.

- **Performance**—For multimedia applications involving only documents and text, the user subjectively determines performance constraints. For applications involving video playback or audio-video synchronization, physical limitations dominate.

More Products

A number of companies have developed databases designed for managing multimedia content. Those include Cumulus from Canto Software (www.canto-software.com), MediaBank from Bitstream (www.bitstream.com) and Tropix from Alaras (www.alarascorp.com). FocalBase, an offshoot of Vicom, also produces a standalone multimedia asset manager.

Conclusion

Corporations will become progressively more aware that they can save time and money by organizing their multimedia assets for use in marketing, training, web sites and other applications. Producers of multimedia asset management software will continue the move toward integrating with other applications and expanding the functionality of their products. More multimedia elements will be included in informational databases as a way of enriching content and engaging the user. All in all, multimedia remains a market worth watching.

Reference # 35J-2005-12-06-01

References

1. John Z Li, M, Tamer Ozsu, Duane Szafron, "Query Languages in Multimedia Database Systems"; http://www. citeseer.ist.psu.edu
2. Judith Lamont , "Managing Multimedia Assets: Evaluating the Options", KMWORLD- Volume 7 , Issue 6; http://www.kmworld.com
3. Mohan Kamath, Krithi Ramamritham, and Don Towsley, "Continuous Media Sharing in Multimedia Database Systems"; http://www. citeseer.ist.psu.edu
4. Raghurama Krishna and Johannes Gehrke, *Database Management Systems*, 2nd Edition, Published by Mc-Graw Hill.
5. Ramez Elmasri and Shamkant Navathe, *Fundamentals of Database Systems*, 3rd Edition, Addison Wesley Publishing Company.

Phishing

Phishing is e-mail fraud where the perpetrator sends out legitimate-looking e-mails that appear to come from well-known and trustworthy web sites in an attempt to gather personal and financial information from the recipient. A phishing expedition, like the fishing expedition it's named for, is a speculative venture: the phisher puts the lure hoping to fool at least a few of the preys that encounter the bait. Web sites that are frequently spoofed by phishers include PayPal, eBay, MSN, Yahoo, BestBuy, and America Online.