



Multimedia Databases (MM DBMS)

Multimedia Databases

- To provide such database functions as indexing and consistency, it is desirable to store multimedia data in a database
 - rather than storing them outside the database, in a file system
- The database must handle large object representation.
- Similarity-based retrieval must be provided by special index structures.
- Must provide guaranteed steady retrieval rates for continuous-media data.

Types of multimedia data

- *Text*: using a standard language (SGML, HTML)
- *Graphics*: encoded in CGM, postscript
- *Images*: bitmap, JPEG, MPEG
- *Video*: sequenced image data at specified rates
- *Audio*: aural recordings in a string of bits in digitized form

Nature of Multimedia Applications

- *Repositories*: central location for data maintained by DBMS, organized in storage levels
- *Presentations*: delivery of audio and video data, temporarily stored, 'VCR-like functionality'
- *Collaborative*: complex design, analyzing data

Management Issues

- *Modeling*: complex objects, wide range of types
- *Design*: still in research
- *Storage*: representation, compression, buffering during I/O, mapping
- *Queries*: techniques need to be modified
- *Performance*: physical limitations, parallel processing

Multimedia Data Formats

- Store and transmit multimedia data in compressed form
 - JPEG and GIF the most widely used formats for image data.
 - MPEG standard for video data use commonalities among a sequence of frames to achieve a greater degree of compression.
- MPEG-1 quality comparable to VHS video tape.
 - stores a minute of 30-frame-per-second video and audio in approximately 12.5 MB
- MPEG-2 designed for digital broadcast systems and digital video disks; negligible loss of video quality.
 - Compresses 1 minute of audio-video to approximately 17 MB.
- Several alternatives of audio encoding
 - MPEG-1 Layer 3 (MP3), RealAudio, Windows Media format, etc.

Continuous-Media Data

- Most important types are video and audio data.
- Characterized by high data volumes and real-time information-delivery requirements.
 - Data must be delivered sufficiently fast that there are no gaps in the audio or video.
 - Data must be delivered at a rate that does not cause overflow of system buffers.
 - Synchronization among distinct data streams must be maintained
 - Video of a person speaking must show lips moving synchronously with the audio

Video Servers

- **Video-on-demand** systems deliver video from central video servers, across a network, to terminals
 - Must guarantee end-to-end delivery rates
- Current video-on-demand servers are based on file systems; existing database systems do not meet real-time response requirements.
- Multimedia data are stored on several disks (RAID configuration), or on tertiary storage for less frequently accessed data.
- Head-end terminals - used to view multimedia data
 - PCs or TVs attached to a small, inexpensive computer called a set-top box.

Design Criteria for Multimedia Database

- **Layout:** Optimize layout of data blocks on secondary storage. If we can store the data optimally on disk, the access times for it will be much faster.
- **Buffer requirements:** Optimize buffer sizes for multiple data streams. Buffer refers to the “read ahead” data to keep user from waiting.

Design Criteria for Multimedia Database

- **Admissibility criteria:** Determine when it is safe to accept an additional client. If this isn't checked, MM database is vulnerable to buffer overflow attacks or DOS (denial of service) attacks.
- **Scheduling:** Ordering and prioritizing client requests such that client constraints are met and buffer resources are minimized.

What does a Multimedia DBMS need?

- Provide traditional database management concepts
 - Persistence of data
 - Consistent view of data (transactions, concurrency management, data integrity)
 - Recovery & Versioning
 - Security of Data
 - Query and retrieval of data
(relational queries, information retrieval)

What does a Multimedia DBMS need?

However, a MM DBMS must provide additional features:

- Storage of multidimensional data
- Temporal relationships (relating to or limited by time)
- Descriptive (content oriented) search
- Device and format independence
- View-specific and simultaneous data access
- Management of large amounts of data
- Real time data transfer

Storing Multimedia Data

How can we store Multimedia Data?

- **ORDBMS** include object features within relational databases
 - Such as the facility for user defined types
- Implemented within Oracle using:
 - LOB
 - CLOB
 - BLOB
 - BFILE

Similarity-Based Retrieval

Examples of similarity based retrieval

- Pictorial data: Two pictures or images that are slightly different as represented in the database may be considered the same by a user.
 - E.g., identify similar designs for registering a new trademark.
- Audio data: Speech-based user interfaces allow the user to give a command or identify a data item by speaking.
 - E.g., test user input against stored commands.
- Handwritten data: Identify a handwritten data item or command stored in the database

Content Based Retrieval

- Multimedia databases need to support queries based on media content as well as traditional queries based on precise field values.
- MMDBs will also support both exact match and probabilistic retrieval.

Example:

Face recognition software. Content and fuzzy search capabilities.

Content Based Retrieval

- With relational databases, the result of a query is one set.
- With MMDBs, content queries that include weights and uncertainties are used.
- The results are probabilistic.
- Often, the returned results will be a ranked result set, indicating results found closest to user's query.

Content Based Retrieval

Methods of doing retrieval:

- Text annotations (entered by people) describe contents
- Query by example compares media elements to specified graphical shapes or image characteristics
- Browsing enables users to quickly scan the data

Example: Think of a query-by-humming system for a music database

Multimedia Query Processing

- 1) Query on content of media information.
Example: Show the details of the movie where an actor says:
“Do you feel lucky, punk?”.
- 2) Query by example. Example: What is the movie which contains this song.
- 3) Time indexed queries. Example: Show me the movie 30 minutes after its start.
Based on temporal characteristics.

Multimedia Query Processing

- 4) Spatial queries: Show me an image where President Yelstin is seen to the left of President Clinton. Based on metadata.
- 5) Application specific queries: Show me the video where the river changes its course.

Problems / issues in MMDBS

- 1) Extremely high capacity storage
- 2) Efficiency requires compression techniques integrated into the DBMS
- 3) Retrieval of multimedia objects in a distributed system
- 4) Video on demand:
 - Server must contend with specialized temporal requests.
 - Video delivery rates must be synchronized at server and client
 - characteristics of communication network must be considered.
 - This is necessary to allow user to pause, rewind, fast forward VOD.

Multimedia Database Applications

- *Documentation and keeping Records*
- *Knowledge distribution*
- *Education and Training*
- *Marketing, Advertisement, Entertainment, Travel*
- *Real-time Control, Monitoring*

Research Problems

- *Information Retrieval in Queries:* Modeling the content of documents
- *Multimedia/Hypermedia Data Modeling and Retrieval:* Hyperlinks, Used in WWW
- *Text Retrieval*