# Multimedia Database Systems: Where are we now?

Harald Kosch and Mario Döller
Institute of Information Technology, University Klagenfurt
Universitätsstr. 65/67, A -9020 Klagenfurt
Austria
harald(mario)@itec.uni-klu.ac.at

**ABSTRACT**
This paper shows the evolution of Multimedia Database Systems and derives a list of open issues to be handled in the future

**KEY WORDS**
Multimedia Database Systems, MPEG-7 and MPEG-21

## 1. Introduction

A *Multimedia Database Management System (MMDBMS)* must support multimedia data types in addition to providing facilities for traditional DBMS functions like database creation, data modeling, data retrieval, data access and organization, and data independence. The area and applications have experienced tremendous growth. Especially with the rapid development of network technology, multimedia database system gets more tremendous development and multimedia information exchange becomes very important. This paper reviews the history and the current state-of-the-art in MMDBMS.

This paper relies in parts on my previously published book:
*Harald Kosch, "Distributed Multimedia Database Technologies supported by MPEG-7 and MPEG-21," CRC Press. 280 pages. November 2003. ISBN: 0-849-31854-8*.

## 2. The first wave

The first MMDBMS rely mainly on the operating system for storing and querying files. These were ad-hoc systems that served mostly as repositories.

The mid 90s saw a *first wave* of commercial, implemented from-the-scratch, and full-fledged MMDBMS. Some of them were MediaDB, now MediaWay [1], JASMINE [2], and ITASCA that is the commercial successor of ORION [3]. They were all able to handle diverse kinds of data and provided mechanisms for querying, retrieving, inserting, and updating data. Most of these products disappeared from the market after some years of existence, and only some of them continued and adapted themselves successfully to the hardware and software advances as well as to application changes. For instance, *MediaWay* provided early very specific support for a wide variety of different media types. Specifically different media file formats varying from images, and video to PowerPoint documents can be managed segmented, linked and searched. Information on the current solutions can be found at http://www.mediaway.com.

## 3. The second wave

In a *second wave*, commercial systems were proposed which handle multimedia content by providing complex object types for various kinds of media. The object-oriented style provides the facility to define new data types and operators appropriate for the new kinds of media, such as video, image and audio. Therefore, broadly used commercial MMDBMSs are extensible Object-Relational DBMS (ORDBMSs). They are successfully shipped since 1996-1998, started by Informix. The current releases are significantly improved in performance and integration into the core systems. Future works include the extension of search services (mainly similarity) to video and audio and advanced presentation and browsing facilities.

The most advanced solutions are marketed by Oracle 10g, IBM DB2 and IBM Informix. They propose a similar approach for extending the basic system. As sample, we consider the IBM DB2 Universal Database Extenders. The

The *IBM DB2 Universal Database Extenders* extends the ORDBMS management to images, video, audio, and spatial objects. All these data types are modeled, accessed, and manipulated in a common framework. Features of the multimedia extenders include importing and exporting multimedia objects and their attributes into and out of a database, controlling access to non-traditional types of data with the same level of protection as traditional data, and browsing or playing objects retrieved from the database.

For instance, the *DB2 Image Extender* defines the distinct data type *DB2IMAGE* with associated user-defined functions for storing and manipulating image files (http://www-306.ibm.com/software/data/db2/extenders/). The actual contents of the image file that DB2Image describes can be stored as binary large objects (BLOB) or outside of the database in a file system. The following SQL-insert statement shows how an image is stored into a column named *image*, in a table named *example*. In this example, the content of the image comes from a server file and stored as a BLOB in the database:

```
INSERT INTO example (image) VALUES(
  DB2IMAGE (
    CURRENT SERVER,
    'pisa.jpg',                /* source_file */
    'JPG',                     /* source_format */
    1,                         /* 1=BLOB, 2=file pointer */
    'my Image File'            /* comment */
  )
)
```

The DB2 Image Extender provides similarity search functionality based on the *QBIC* technology (http://wwwqbic.almaden.ibm.com/) for images stored in the DB2IMAGE type. The QBIC technology provides the ability to query, or search, for images based on their content. Using this query mechanism, one may specify image content features, such as color values and another image file as input to a query. These features are matched against the contents of the images stored in the database, and a score is assigned to each image. A score is a double-precision floating-point value between 0 and 1 that indicates how closely an image's features match those specified in the QBIC query. The lower the score, the closer is the match. The image features that can be used in QBIC queries are average color, histogram color distribution, positional color values and the texture of an image. For instance, the following SQL statement shows an example of executing a QBIC query that ranks each image in the image column based on how closely its average color matches red:

```
SELECT CONTENTS(image),
 QBScoreFROMStr(`averageColor=<255,0,0>', image) AS SCORE
 FROM signs ORDER BY SCORE
```

Besides the commercial products, several research projects have implemented full-fledged multimedia database systems. Recently successfully terminated projects include the following:

*MIRROR* [4], an acronym for Multimedia Information Retrieval Reducing information OveRload and developed at the University of Twente, is a research MMDBMS that is developed to better understand the kind of data management that is required in the context of multimedia digital libraries. Its main features are an integrated approach to both content and traditional structured data management. MIRROR provides probabilistic inference mechanism, during the interaction with the user [5], which has been adopted from cognitive theories. MIRROR is implemented on top of the ORDBMS Monet database system. More information on MIRROR may be obtained at:
http://wwwhome.cs.utwente.nl/~arjen/mmdb.html.

On top of MIRROR runs the *ACOI* system (http://monetdb.cwi.nl/acoi/) that is a platform for indexing and retrieval of video and image data. The system provides a plug-in architecture to subsequently index multimedia objects using various feature extraction algorithms [6]. ACOI relies on the COBRA (COntent-Based RetrievAl) video data model (only low-level descriptors). COBRA introduces a feature grammar to describe the low-level persistent meta-data and the dependencies between the extraction mechanisms [7]. Figure 1 shows the system architecture of the MMDBMS product:

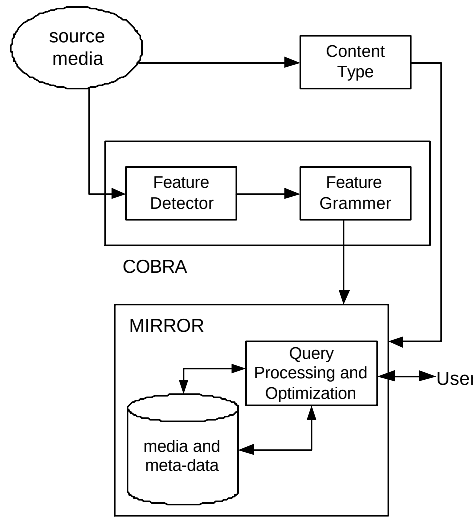


**Figure 1: ACOI/MIRROR System.**

*DISIMA* [8,9], an acronym for Distributed Multimedia DBMS, developed at the University of Alberta, is an image database system which enables content-based querying. Additional information at: http://db.uwaterloo.ca/~ddbms/projects/multimedia/
Figure 2 shows the system architecture of DISIMA.

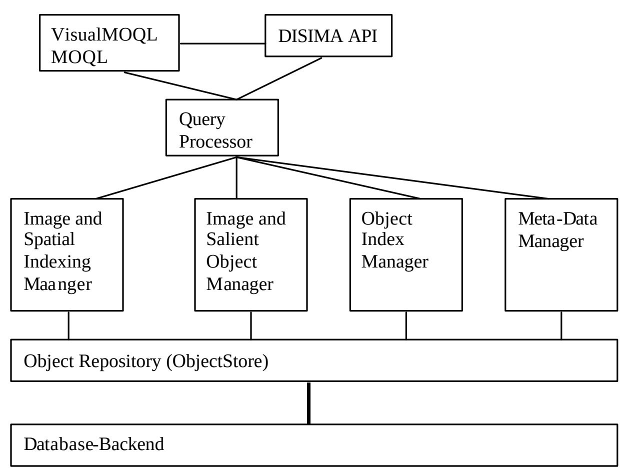


**Figure 2: DISIMA System**

The prototype was implemented on top of the DBMS ObjectStore. Queries are specified in the query language MOQL (or Visual MOQL for images only), which relies on a new type concept model for both image and spatial applications. The associated query languages (MOQL and Visual MOQL), extending OQL, allow spatio-temporal querying as well as the definition of a presentation specification. For instance, the following simple query allows to find all images in which a person appears (assuming that an Image table m and Person table p is defined beforehand).

```
SELECT m
  FROM Images m, Persons p
  WHERE m contains p
```

## 4. The third wave

The *third wave* includes currently running projects (recently finished or still running). They mainly address the needs of applications for richer semantic content. Most of them rely, thereby, on the new MPEG-standards MPEG-7 and MPEG-21. Representative running project is the *MARS* project carried out at the University of Illinois at Urbana-Champaign and our MPEG-7 Multimedia Data Cartridge.

Before jumping into details of the projects, let us give a quick introduction to MPEG-7 and MPEG-21.
*MPEG-7* [10] is the ISO/IEC 15938 standard for multimedia descriptions and issued in 2002. It is an XML-based multimedia meta-data standard, which proposes description elements for the multimedia processing cycle from the capture (e.g., logging descriptors), analysis/filtering (e.g., descriptors of the MDS, Multimedia Description Schemes), to the delivery (e.g., media variation descriptors), and interaction (e.g., user preference descriptors).

*MPEG-21* [10] is the ISO/IEC 21000 standard defining an open multimedia framework. Most of the parts are in work. The driving force for MPEG-21 was the current situation that many elements exist to build an infrastructure for the delivery and consumption of multimedia content but that there exits no “big picture” to describe how these elements relate to each others. The vision for *MPEG-21* is to define an *open multimedia framework* that will enable transparent and augmented use of multimedia resources across a wide range of networks and devices used by different communities. The intent is that the framework will cover the entire multimedia content delivery chain encompassing content creation, production, delivery, personalization, consumption, presentation and trade.

Both standards had a major influence on the current products, both in design, as implementation.

*MARS*, an acronym for Multimedia Analysis and Retrieval System [11,12], realizes an integrated multimedia information retrieval and database management system, that supports multimedia information as first-class objects suited for storage and retrieval based on their semantic content. MARS proposes a set of tools for an MMDBMS Back-End, as shown in Figure 3.

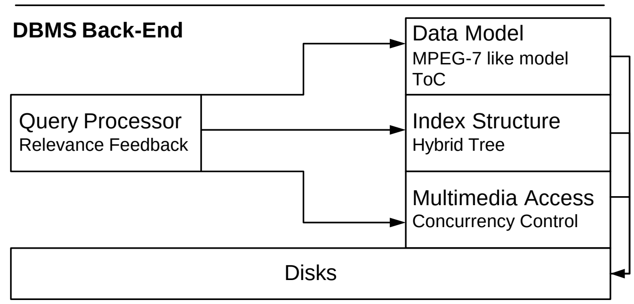


**Figure 3: MARS Project**

The MARS project includes the conception of a multimedia data model, for content indexing and retrieval, and for database management. The presented multimedia data model influenced the development of the MPEG-7 standard. *MARS* is a from-the-scratch management system, based on a query refinement processing. Furthermore, a table of content (ToC) extraction mechanism for videos has been developed [13]. A hybrid tree data structure to support high-dimension feature indexing in multimedia databases [14] has also been realized. For the multimedia information retrieval, an adapted relevance feedback approach that is able to learn user's information need in image databases is proposed [11, 12]. More information on MARS may be obtained from http://www-mars.ics.uci.edu.

The *MPEG-7 Multimedia Data Cartridge (MDC)* is a system extension of the Oracle 9i DBMS providing a multimedia query language, access to media, processing and optimization of queries, and indexing capacities relying on a multimedia database schema derived from MPEG-7 [15,16]

The *MDC* builds on three main concepts (see Figure 4). At first, the *Multimedia Data Model* is the database schema which is derived from MPEG-7 descriptions. It is realized with the help of the extensible type system of the cartridge environment, i.e., descriptors in the MPEG-7 schema are mapped to object types and tables. At second, the *Multimedia Indexing Framework* (MIF) which provides an extensible indexing environment for multimedia retrieval. The indexing framework is integrated into the query language and enables efficient multimedia retrieval. Finally a set of *internal and external libraries* allow the access to the media and communicate with MDC (query, insert, update, etc.).

The *Multimedia Schema* of MDC relies on the one hand on the structural and semantic parts of the MPEG-7 standard (high-level descriptions). On the other hand, object types for the MPEG-7 low-level descriptors, like color, shape, texture are provided and linked to the high-level descriptions. This enables one to retrieve multimedia data not only by low-level features, but also by semantics in combination with low-level characteristics.

The *Multimedia Indexing Framework (MIF)* offers advanced indexing services to the MMDBMS. It is generic in a way that new index types may be added without changing the interface definitions. The MIF is divided into three modules (as seen if Figure 4). Each module, especially the GistService and the Oracle Enhancement may be used on its own and may be distributed over the network.

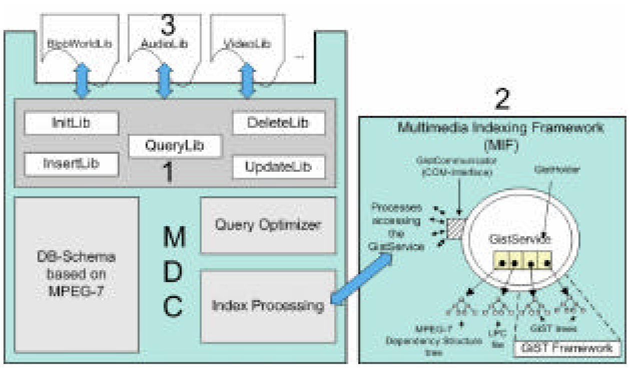


**Figure 4: Multimedia Data Cartridge**

The *GistService* is the main part and realized in the external address space. It offers services for index management and relies on the Generalized Search Trees (GIST) framework (http://gist.cs.berkeley.edu/). The original GIST framework was extended to cope with several index trees at time and to be linked to the Oracle DBMS. Several index trees can be employed in our system, in the catgory balanced trees e.g., X-tree and SR-tree, and other access methods not relying on balanced trees (VA-files). This service is split into two components: *GistCommunicator* and *GistHolder*. The GistCommunicator is a COM-object (Component Object Model) used for inter-process communication between the database (by passing through the GistWrapper shared library) and the implemented access methods. Thus, the GistCommunicator supplies the necessary functionality (e.g., creating, inserting, deleting) for accessing the index structures. The GistHolder manages all currently running index trees and the accesses to them. Each index tree is identified through a global and unique ID, which is forwarded to the accessing process. The integration of MIF into the MDC is done via the index extension mechanisms of Oracle 9i. For each new index, a new Oracle *indextype* has to be defined, but the interface to the Gist remains unchanged.

## 5. Open issues

Since the first multimedia database system ORION was developed in 1987, the area and applications of Multimedia Databases have experienced tremendous growth. From the multimedia database viewpoint two main issues have to be considered:

- **Issues on Multimedia Data Modeling:** A Multimedia Data model must deal with the issue of representing the multimedia objects, that is, designing the high and low-level abstractions of the media data to facilitate various operations. These operations may include media objects selection, insertion, editing, indexing, browsing, querying, retrieval and communication. A good data model should fulfill the following issues:
  1. It is necessary to model and store media components in the database. Its storage mechanism will play an important factor affecting the performance of multimedia system.
  2. A representation should be provided for the logical media structure. It is essential to represent this structure explicitly both for querying and for representation.
  3. Semantic meaning should be modeled and linked to low-level characteristics and the structure of the media.
  4. The meta-data necessary for the operation of the system components needs to be determined and stored in the database.
  5. It is necessary to rely on international standards, e.g., MPEG-7, in order to guarantee interoperability for data sharing and data exchange.
- **Issues on Multimedia Indexing, Querying and Presentation:** The key functionality in a multimedia database is how to access and how to exchange multimedia information effectively. No matter what data model is used, data store mechanisms are employed, the most important thing is how to retrieve and communicate continuous and non-continuous media with a short real-time constraint. The key to retrieval process is similarity between two objects. The content of the object is analyzed and used to evaluate specified selection predicates, the process called *content-based retrieval*. In order to have an accurate representation of the multimedia objects in the database and the query object, different features have to be combined (e.g., texture, shape and color). The results are high-dimensional feature vectors (usually more than 1024 values). The efficiency of the similarity search in such spaces must be supported by the use of special multi-dimensional indexing structures and by dimension reduction methods, which however take care to the different nature of the feature components. Finally, similarity metrics have to re-visit to take the user perception into account.

In order to retrieve multimedia data from a database system, a query language must be provided. A multimedia query language must have abilities to handle complex spatial and temporal relationships. A powerful query language should have to deal with keywords, index on keywords and semantic contents of multimedia objects. The key to efficient communication is to rely on standards for communicating meta-data and associated media data.

An important contribution is the development of the system parts of MPEG-7 and the file format of MPEG-21. Finally, the issues considered above on data model, storage retrieval and query should not be limited to atomic multimedia data like image, audio or video. It is also interesting to retrieve composite objects, such as when we read news text, and at the same time, we open a window to see video documentaries. Hence we need the ability to deal with several atomic multimedia data, media data presentation, at the same time. MPEG-21 goes in this direction with the provision of the notion of a Digital Item, which surely will influence the multimedia database world in the future. In addition and related again to MPEG-21, it is important for a multimedia database system to use multiple representations of data for different Users and profiles, for Intellectual Property Management and for adaptation purposes.

## References:


[1] B. Phillips, Mediaway presses access to multimedia database. *PC Week, 13*(7), 1996, 39-40.

[2] D. Woelk, W. Kim, & W. Luther, An object-oriented approach to multimedia databases. *ACM SIGMOD Record*, *15*(2), 1986, 311-325

[3] Setrag Khoshafian, Surapol Dasananda, & Norayr Minaasian, *The Jasmine Object Database: Multimedia Applications for the Web* (Morgan Kaufmann Publishers, 1998).

[4] Arjen P. de Vries, Mark G. L. M. van Doorn, Henk M. Blanken, & Peter M. G. Apers, The MIRROR MMDBMS architecture. *Proc. of the International Conference on Very Large Databases*, Edinburgh, Scotland, 1999, 758-761

[5] Mark G. L. M. van Doorn & Arjen P. de Vries, The psychology of multimedia databases. *Proc. of the Fifth ACM Conference on Digital Libraries*, San Antonio, TX, USA, 2000, 1-9.

[6] A.P. de Vries, M.A. Windhouwer, P.M.G. Apers, & M.L. Kersten, Information access in multimedia databases based on feature model. *New Generation Computing, 18*(4), 2000, 323-339.

[7] H. E. Blok, M. A. Windhouwer, R. van Zwol, M. Petkovic, P. M. G. Apers, M. L. Kersten, & W. Jonker, Flexible and scalable digital library search. *Proc. of the International Conference on Very Large Data Bases (VLDB),* Roma, Italy, 2001, 705-706.

[8] V. Oria, M.T. Özsu, & P. Iglinski. VisualMOQL, The DISIMA visual query language. *Multimedia Tools and Applications, 23*(3), 2004, 185-201.

[9] V. Oria, M.T. Özsu, P. Iglinski, S. Lin, & B. Ya. DISIMA: A distributed and interoperable image database system, *Proc. of the ACM SIGMOD International Conference of Management of Data,* Dallas, Texas, USA, 2000, 600.

[10] Harald Kosch, *Distributed Multimedia Database Technologies supported by MPEG-7 and MPEG-21* (CRC Press. 280 pages. November 2003. ISBN: 0-849-31854-8).

[11] Saushik Chakrabarti, Kriengkrai Porkaew, & Sharad Mehrotra, Efficient query refinement in multimedia databases. *Proc. of the IEEE International Conference on Data Engineering (ICDE),* San Diego, California, USA, 2000, 196.

[12] Y. Rui, T. Huang, & S. Mehrotra. Relevance feedback: a powerful tool for content-based image retrieval, *IEEE Transaction on Circuits and System for video technology, 8*(5), 1998, 25-36.

[13] Yong Rui, Thomas S. Huang, & Sharad Mehrotra. Constructing Table-of-Content for videos, *ACM Multimedia Systems, 7*(5), 1999, 359-368.

[14] Kaushik Chakrabarti & Sharad Mehrotra, Efficient concurrency control in multidimensional access methods. *Proc. of the International ACM SIGMOD Conference on Management of Data*, Philadelphia, Pennsylvania, USA, 1999, 25-36.

[15] Mario Döller & Harald Kosch, An MPEG-7 Multimedia Data Cartridge, *Proc. of the SPIE conference on Multimedia Computing and Networking 2003* (MMCN 2003), Santa Clara, California, USA, 2003, 126-137

[16] Harald Kosch, Laszlo Böszörmenyi, Mario Döller, Mulugeta Libsie, Andrea Kofler & Peter Schojer The life-cycle of Multimedia Metadata *IEEE MultiMedia. 12*(1), January/March 2005.


## Selected Links:

Multimedia Database HomePage:
http://www-itec.uni-klu.ac.at/~harald/codac/

Overview of Content-Based Retrieval Tools:
http://give-lab.cs.uu.nl/cbirsurvey/

MPEG-Homepage:
http://www.chiariglione.org/mpeg/index.htm

MPEG-7 Tools:
http://m7itb.nist.gov/M7Validation.html

MPEG-7 and MPEG-21 Applications:
http://www.m4if.com/resources.php#section40