Journal of Universal Knowledge Management, vol. 1, no. 1 (2006), 18-25
submitted: 14/5/06, accepted: 29/5/06, appeared: 28/8/06 © J.UKM

MPEG-7 Meets Multimedia Database Systems

Mario Döller
Department of Distributed Information Systems
University Passau, Germany
Mario.Doeller@uni-passau.de

Abstract: Currently used Database Management Systems do not fulfill the requirements of multimedia in querying, indexing and content modeling. Thus, most database provider offer extenders for multimedia data. These extensions, however, provide only limited semantic modeling and rely on simple index structure which do not meet the whole nature of multimedia. In this context, this paper points out approaches for the integration of MPEG-7 as a standard for describing multimedia content into a database management system and its impact to core parts of a database such as data model, access methods, query language and query optimization.
Key Words: Multimedia Databases, MPEG-7
Category: H.3.3, H.2

1 Introduction

Internet multimedia applications, such as video-on-demand, video conferencing, multimedia retrieval services, etc. let us experience multimedia at everyone's desktop and communication devices such as personal digital assistant (PDA) or mobile phone.

Strongly related to this is the enhancement of a multimedia communication with meta-data. Meta-data are descriptive data about multimedia content. These could be semantic descriptions, as for instance which persons appear in a video clip, information on color characteristics of a video (e.g., the dominant color in an image), or it could be information on how a video might be adapted if resources become rare.

In this context, the Moving Picture Experts Group (MPEG) introduced in 2002 a new meta-data standard, called MPEG-7 [Martinez, 2003], for describing high- and low-level features of multimedia data.

Imagine that you are listening to a radio song and you could not remember the title. Using your mobile phone, get recorded 10s of the song, then use an audio recognition service based on MPEG-7 Audio descriptors, extraction mechanisms and the multimedia database and get a prompt and positive content identification via SMS. In order to enable the described scenario, one needs methods for extracting low level features (in this example, the audio signals) from the unknown audio file. A methodology that allows the description of the multimedia meta-data (e.g., through MPEG-7) and an audio retrieval system that provides means for recognizing similar audio signatures. In general, such retrieval systems are closely coupled with database systems. Therefore, Multimedia Database Management Systems (MMDBMSs) are the technology for content management, storage and streaming [Kosch, 2003] of multimedia. This paper deals with the integration of MPEG-7 as a meta-data standard for multimedia into a database management system and shows how MPEG-7 and Multimedia Database Systems can benefit from each other. Identified open issues, problems and several solutions, etc. bases on experiences the author gained through his participation in the CODAC[1] project which among others targeted on the creation of a MPEG-7 Multimedia Database (MPEG-7 MMDB).

The remainder of this paper is organized as follows: Section 2 covers related work in the area of multimedia databases and their core parts such as query languages. Then, Section 3 describes requirements a modern database management system must support in order to enable the integration of MPEG-7 as a data model. The integration of MPEG-7 and its impact to core parts of a database is discussed in Section 4 and its subsections. Finally, this paper concludes in Section 5.

2 Related Work

Research and developments in the domain of multimedia databases can mainly be distinguished between two directions. Based on the fact that most existing Database Management Systems (DBMS) are basically not designed for multimedia, database vendors provide extenders that enable fundamental processing of multimedia data (e.g., Oracle *inter*Media [Oracle, 2003] and IBM Informix DataBlades [IBM, 2001]). For instance, Oracle *inter*Media provides basic image storage and content based retrieval (CBR) functionality through their *ORDImage* data type. The underlying CBR functionally concentrates on low level features (color, texture, shape) without the possibility for semantic retrieval. Furthermore, no mean for video or audio CBR is available.

The second research direction concentrates on special-purpose MMDBMS which are especially tuned for multimedia data (e.g., DISIMA [Oria et al., 2004] and MARS [Mehrotra et al., 1997]). In general, these systems provide individual multimedia data models [Wen et al., 2003], corresponding query languages (e.g., MOQL [Li et al., 1997]) and respective approaches for any kind of content-based retrieval [Belongie et al., 1998]. Nevertheless, their drawback is that they are not designed to query multimedia and traditional data at the same time, nor efficient access structures are available.

In contrast, there are efforts in order to combine the MPEG-7 standard with modern database management systems [Kosch, 2002]. Due to the fact, that MPEG-7 relies on XML Schema, XML solutions for databases [Murthy and Banerjee, 2003] and native XML solutions [Staken, 2002] have to be considered as well. The authors in [Westermann and Klas, 2003] presented an analysis of XML database solutions for the management of MPEG-7 descriptions.

[1] `http://www.fmi.uni-passau.de/lehrstuehle/kosch/research/codac.php`

3 Database Requirements for Multimedia Support

In general, one can classify the requirements in the following three sub-areas, namely *structural, semantical* and *syntactical*.

As described above, common DBMS have several drawbacks in handling multimedia data [Santini and Jain, 1997, Grosky, 1997]. As database vendors can not support all needs and individual conveniences of different domains (e.g., requirements for multimedia systems, geographical information systems, etc.), they build their database according to a modular architecture and made their management systems extensible. Figure 1 shows an example architecture (taken respectively from Oracle 9i and 10g), most modern databases support. Such an architecture provides means for extending the basic database services such as type system (e.g., for the integration of a new data model), query processing, optimization and indexing.

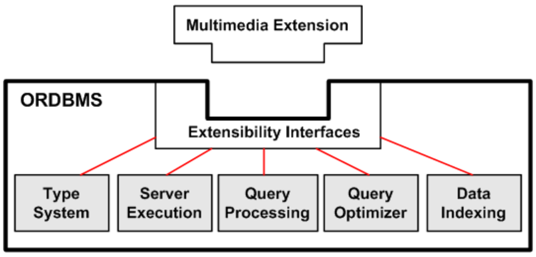

Figure 1: Necessary extensibility

These extensibility services cover base *structural* requirements by enabling the enhancement of core parts (indexing facility, query language, query optimization, etc.) of a database.

Semantical requirements concern data modeling and query facilities. A query language need to support low-level (content-based) and high-level (semantic) query operations. Besides, spatial and temporal (or a combination of them) operations are required. An example for a complex query might be: *I am searching for images that show a red Ferrari besides a green house.* Besides the integration of these operations, one has to consider the fact, that the location of a descriptor in MPEG-7 documents can vary. The MPEG-7 schema allows many different ways of describing the same multimedia content. Given a free text annotation describing a person in an image, one may use the FreeText DS, or enhance the level of semantic by using the who section in the StructuredText DS. But the interpretation of the information is the same. In addition, the information can be assigned to different segments (e.g., various StillRegions). A search engine and its corresponding query language has to consider all possible information variations in order to optimize recall and precision. Furthermore, we can distinguish between context-unaware and context-aware retrieval. In context-unaware retrieval only top level search is performed, without taking the hierarchy into consideration. In contrast, context-aware retrieval takes into account that the description of multimedia data is organized in a tree like hierarchy.

Syntactical requirements deal with the input and output format of multimedia queries, inserts and update operations. In the case of MPEG-7, the inserting of MPEG-7 documents has to ensure that only valid and well formed documents are inserted. During an query operation, one might claim that the result must be delivered as valid MPEG-7 document(s). This is needed, when the result is forwarded to applications which can only process MPEG-7 descriptions. In addition, update operations have to ensure that consistency in the sense of MPEG-7 conformance of updated data is guaranteed.

Detailed information to mentioned core parts, namely data model, query language, access methods and query optimization is presented in Section 4.

4 Integration of MPEG-7 into MMDBMS

This section addresses the integration of the MPEG-7 standard to an corresponding database data model for storing multimedia meta data and its consequence for depending parts such as access methods, query language and query optimization.

4.1 Data Model

A crucial factor for managing and retrieving multimedia data within a database is the underlying data model. Is the data model too coarse-grained (unstructured storage approach, e.g., the whole MPEG-7 document is stored in a database XMLType), storage operations are simple whereas retrieval is limited. Is the data model too fine-grained (structured storage (see Florescu et al. [Florescu and Kossmann, 1999], e.g., create for every MPEG-7 descriptor an equivalent database table), storage operations will lead to many sparely filled tables, besides retrieval can support semantically rich queries.

Due to the fact, that MPEG-7 relies on XML-Schema, mapping strategies [Christophides et al., 1994, Amer-Yahia and Fernandez, 2001] for XML to an equivalent database data model have to be considered.

In order to circumvent mapping problems mentioned previously (e.g., sparely filled tables), the transformation strategy for MPEG-7 should consider a trade-off between both directions (structured and unstructured approach) as demonstrated in [Döller, 2004]. The authors utilize available object-relational features for mapping MPEG-7 descriptors to corresponding database types and tables and the supported *XMLType* to reduce complexity. The combination of object-relational database features, relational keys and object references allows the mapping of the whole MPEG-7 standard into a corresponding database schema. The reduction of the MPEG-7 inheritance hierarchy by skipping abstract types and merging types that only contain a few attributes and elements results in a compact arranged database schema that allows the storage of any kind of MPEG-7 document and offers an efficient and rich model for querying it.

The data model itself is only a first step for a high-level multimedia database system which bases on MPEG-7. In addition, one has to consider means for inserting, deleting and updating MPEG-7 documents. These facilities may be integrated into the database system or provided as tools.

Furthermore, a MPEG-7 multimedia database system has to deal with multimedia query languages, supporting access methods and means for query optimization of multimedia queries (see upcoming Sections).

4.2 Query Language

Traditional database management systems have been very effective and efficient in storing and managing alphanumeric data. Nowadays, based on the ubiquity of digital cameras and MP3-players, the amount of multimedia data is overwhelming. Querying alphanumeric data relies on matching and filter operations which decides for every tuple whether it fits the requirements or not. In multimedia database systems, we basically are interested in *similar* data. Therefore, databases have to provide adequate query paradigms for similarity searches [Stricker and Orengo, 1995].

In this context, SQL/MM [Melton and Eisenberg, 2001] introduces a conceptual multimedia data model for the use in multimedia database systems that extend the concept of the object-relational SQL-99. Compared with MPEG-7, the data model of SQL/MM covers the syntactical part of multimedia descriptions but allows no means for decomposing an image for describing the content semantically meaningful.

Furthermore, the multimedia query language, MOQL [Li et al., 1997] extends the ODMG's Object Query Language (OQL) [Jordan, 1998] by adding spatial, temporal and presentation properties for content-based image and video data retrieval.

By the use of MPEG-7 as a data model in multimedia database systems, one is confronted to think about enhancements of the query language for multimedia data such as similarity search. In addition, the integration of operations that can produce XML output has to be considered as well. This is as important as the import format and the output format (both MPEG-7 descriptions) should correspond to each other.

This means if it is necessary to retrieve the query result as MPEG-7 documents, one has to combine the multimedia query language (e.g., SQL/MM) with SQL/XML [Eisenberg and Melton, 2002] elements (e.g., XMLAgg, XMLElement, etc.). In addition, it has to be ensured that the resulting XML document satisfies the XML Schema for MPEG-7. This necessitates the enhancement of query processing for type checking of MPEG-7 conformance.

4.3 Access Methods

Indexing is an important concept in modern database management systems to enhance processing efficiency and retrieval capacity (e.g., similarity searches).

Innately, most database systems provide only a limited number of integrated access methods such as B-tree or hashing facilities. These techniques limit the use of database systems for multimedia data. This is as astonishing as in the last decade various different access methods have been established for indexing multidimensional data. To mention only a few: SR-tree [Katayama and Satoh, 1997], M-tree [Ciaccia et al., 1997] or X-tree [Berchtold et al., 1996].

The integration of such access methods is crucial in order to support the retrieval of multimedia data by similarity searches or other query types. In MPEG-7, there exist several descriptors for extracted low level features of audio, video and image data (e.g., ScalableColorType for images, or AudioSignatureType for audio files). Indexing of these descriptors in combination with an enhancement of the query language (see Subsection 4.2) allows the retrieval of multimedia data based on similarity searches across multiple MPEG-7 documents. In series, such indexing can support content-based retrieval based on low level features.

Although, MPEG-7 provides excellent means for semantic indexing and querying by its semantic descriptors, research is still in an early stage and approaches only have partly reached content-based retrieval systems [Bailer et al., 2004] that relies on MPEG-7. To the author best knowledge, semantic indexing in the context of databases and MPEG-7 has not been considered so far.

4.4 Query Optimization

Besides the enhancement of query languages (e.g., SQL/MM) and the integration of new access methods (e.g., SR-tree), one may not neglect the performance of these operations.

In multimedia databases, queries often contain similarity operations such as range or nearest-neighbor operation for low level features (e.g. a color histogram represented by the MPEG-7 ScalableColorType). In order to improve the performance of these operations, one has to extend the query optimizer. In general, a query optimizer can consider three approaches: selectivity, cost model and operator ordering. This paper concentrates on selectivity and cost models, as most modern databases provide only means for their enhancement.

In the literature, several cost models exist that concentrate on calculating the cost of index structures for range and nearest-neighbor searches [Böhm, 2000]. In [Lee et al., 1999], the authors present an efficient cost model for predicting the performance of the k-NN (k-nearest neighbor) query independently of the used index tree. The model is accurate for low- and mid-dimensional data with non-uniform distribution. The estimation of range query's selectivity represents, apart from few initial approaches [Kosch and Döller, 2005], an open research question. There, the authors introduced an approach for approximating the selectivity of range searches within a n-dimensional data set with the help of a density based clustering technique (DBSCAN [Ester et al., 1996]).

5 Conclusion

This paper points out requirements and impacts to core parts of a database management system originated by an integration of the MPEG-7 standard as database data model. For this purpose, requirements of an extensible database are outlined. Then, the integration itself is addressed. This integration covers in particular the database data model, enhancements of query languages, access methods and query optimization.

Hence, there currently are several solutions and proposals available for storing and retrieving MPEG-7 documents. Nevertheless, these systems and proposals leave behind many open issues which have not been considered so far. For instance, every solution applies different (often proprietary) combinations of used retrieval operators and query languages (e.g., SQL/XML in combination with proprietary operators, etc.). Therefore, there is clearly a need for a standardized query language that specifies the input and output format of MPEG-7 queries. This query language has to consider the full strength (spatial, temporal, spatial-temporal, etc.) of multimodal queries.

Next, there is still a limited availability of index structures for high dimensional data. This is well known, but still an unresolved problem. In addition, research has to be done for index structures that are especially tuned for MPEG-7 descriptors and/or descriptor schemes (e.g., indexing of StillRegions).

References

[Amer-Yahia and Fernandez, 2001] Amer-Yahia, S. and Fernandez, M. (2001). Overview of Existing XML Storage Techniques. *AT&T Labs Research.*

[Bailer et al., 2004] Bailer, W., Mayer, H., Neuschmied, H., Haas, W., Lux, M., and Klieber, W. (2004). Content-Based Video Retrieval and Summarization using MPEG-7. In *Proceedings Internet Imaging V*, pages 1–12, San Jose, CA, USA.

[Belongie et al., 1998] Belongie, S., Carson, C., Greenspan, H., and Malik, J. (1998). Color- and Texture-Based Image Segmentation Using EM and Its Application to Content-Based Image Retrieval. In *Proceedings of the International Conference on Computer Vision (ICCV'98)*, pages 675–682, Bombay, India.

[Berchtold et al., 1996] Berchtold, S., Keim, D. A., and Kriegel, H. P. (1996). The X-Tree: An Index Structure for High-Dimensional Data. In *Proceedings of the 22nd Int. Conf. on Very Large Data Bases (VLDB)*, pages 28–39, Mumbai (Bombay), India. Morgan Kaufmann, ISBN 1-55860-382-4.

[Böhm, 2000] Böhm, C. (2000). A Cost Model for Query Processing in High Dimensional Data Spaces. *ACM Transactions on Database Systems (TODS)*, 25(2):129–178.

[Christophides et al., 1994] Christophides, V., Abiteboul, S., Cluet, S., and Scholl, M. (1994). From Structured Documents to Novel Query Facilities. In *Proceedings of the 1994 ACM SIGMOD International Conference on Management of Data*, pages 313–324, Minneapolis, Minnesota.

[Ciaccia et al., 1997] Ciaccia, P., Patella, M., and Zezula, P. (1997). M-tree: An efficient Access Method for Similarity Search in Metric Spaces. In *Proceedings of the 23rd Int. Conf. on Very Large Data Bases (VLDB)*, pages 426–435, Athens, Greece. Morgan Kaufmann, ISBN 1-55860-470-7.

[Döller, 2004] Döller, M. (2004). *The MPEG-7 Multimedia DataBase System (MPEG-7 MMDB).* Dissertation, University Klagenfurt, Austria.

[Eisenberg and Melton, 2002] Eisenberg, A. and Melton, J. (2002). SQL/XML is Making Good Progress. *ACM SIGMOD Record*, 31(2):101–108.

[Ester et al., 1996] Ester, M., Kriegel, H.-P., Sander, J., and Xu, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining*, pages 226–231, Portland, OR, USA.

[Florescu and Kossmann, 1999] Florescu, D. and Kossmann, D. (1999). Storing and Querying XML Data using RDBMS. *IEEE Data Engineering Bulleting*, 22(3):27–34.

[Grosky, 1997] Grosky, W. I. (1997). Managing Multimedia Information in Database Systems. *Communications of the ACM*, 40(12):73–80.

[IBM, 2001] IBM (2001). DataBlade Module Development Overview, Version 4.0. *http://www-306.ibm.com/software/data/informix/blades/*.

[Jordan, 1998] Jordan, D. (1998). *C++ Object Databases, Programming with the ODMG Standard*. Addison-Wesley. 456 pages, ISBN: 0-201-63488-0.

[Katayama and Satoh, 1997] Katayama, N. and Satoh, S. (1997). The SR-tree: An Index Structure for High-Dimensional Nearest Neighbor Queries. In *ACM SIGMOD Int. Conf. on Management of Data*, pages 369–380.

[Kosch, 2002] Kosch, H. (2002). MPEG-7 and Multimedia Database Systems. *Sigmond Records*, 31(2).

[Kosch, 2003] Kosch, H. (2003). *Distributed Multimedia Database Technologies supported by MPEG-7 and MPEG-21*. CRC Press. 248 pages, ISBN: 0-849-31854-8.

[Kosch and Döller, 2005] Kosch, H. and Döller, M. (2005). Approximating the selectivity of multimedia range queries. In *Proceedings of the IEEE International Conference on Multimedia and Expo*, Amsterdam, The Netherlands.

[Lee et al., 1999] Lee, J.-H., Cha, G.-H., and Chung, C.-W. (1999). A Model for k-Nearest Neighbor Query Processing Cost in Multidimensional Data Spaces. *Information Processing Letters*, 69(2):69–76.

[Li et al., 1997] Li, J. Z., Özsu, M. T., Szafron, D., and Oria, V. (1997). MOQL: A Multimedia Object Query Language. In *Proceedings of the third International Workshop on Multimedia Information Systems*, pages 19–28, Como Italy.

[Martinez, 2003] Martinez, J. M. (2003). *MPEG-7 Overview*. ISO/IEC JTC1/SC29/W11 N5525, Pattaya.

[Mehrotra et al., 1997] Mehrotra, S., Rui, Y., Ortega-Binderberger, M., and Huang, T. S. (1997). Supporting Content-based Queries over Images in MARS. In *Proceedings of the 1997 International Conference on Multimedia Computing and Systems (ICMCS '97)*, page 632.

[Melton and Eisenberg, 2001] Melton, J. and Eisenberg, A. (2001). SQL Multimedia Application packages (SQL/MM). *ACM SIGMOD Record*, 30(4):97–102.

[Murthy and Banerjee, 2003] Murthy, R. and Banerjee, S. (2003). XML Schemas in Oracle XML DB. In *Proceedings of the 29th VLDB Conference*, pages 1009–1018, Berlin, Germany. Morgan Kaufmann.

[Oracle, 2003] Oracle (2003). Oracle interMedia Reference, 10g Release 1. *http://download-east.oracle.com/docs/*.

[Oria et al., 2004] Oria, V., zsu, M. T., and Iglinski, P. J. (2004). Foundation of the DISIMA Image Query Languages. *Multimedia Tools and Applicactions Journal*, 23:185–201.

[Santini and Jain, 1997] Santini, S. and Jain, R. (1997). Image Databases are not Databases with Images. In *Proceedings of the 9th International Conference on Image Analysis and Processing 2*, pages 38–45, Florence, Italy.

[Staken, 2002] Staken, K. (2002). Xindice Developers Guide 0.7. *The Apache Foundation, http://www.apache.org*.

[Stricker and Orengo, 1995] Stricker, M. A. and Orengo, M. (1995). Similarity of color images. In *Storage and Retrieval for Image and Video Databases, SPIE*, pages 381–392, San Jose, CA.

[Wen et al., 2003] Wen, J.-R., Li, Q., Ma, W.-Y., and Zhang, H.-J. (2003). A Multi-paradigm Querying Approach for a Generic Multimedia Database Management System. *ACM SIGMOD Record*, 32(1):26–34.

[Westermann and Klas, 2003] Westermann, U. and Klas, W. (2003). An Analysis of XML Database Solutions for the Management of MPEG-7 Media Descriptions. *ACM Computing Surveys*, 35(4):331–373.