Enhancement of Processing Efficiency in Multimedia Database Management Systems and Video Servers supported by the Use of Meta-Data

Dr. Dipl.Inf. Harald KOSCH

Habilitationsschrift

Klagenfurt, im November 2001

Abstract

Multimedia technologies are attracting more and more interest every day. Video-on-Demand is one of the buzzwords today and is now available for the public. Content providers such as publishers, broadcasting companies and audio/video production firms must be able to archive and index their productions for later retrieval. This is a formidable task and it is even more so when the material to be handled encompasses several media types and covers a time span of several years. In order for such a vast amount of data to be easily available, existing multimedia database design models, indexing and retrieval methodologies and delivery methods have to be improved and refined. In addition, video, image and audio data cannot be effectively managed with the exclusive use of older, keyword-based search techniques.

This habilitation thesis collects new contributions to the research issues mentioned above. It contains a collection of ten research articles which focus on efficiency, in particular on the question of how the performance of multimedia database management systems, video servers and their front-ends, like the query optimizer for databases and the admission control for video servers, can be improved. Especially, I concentrate on the impact of meta-data, i.e., how can low- and high-level meta-data, describing the content of multimedia, can be used to enhance the efficiency of these systems.

The following preface first introduces the major research areas of multimedia databases and then summerizes the main contributions of this work in a compact form. For each of the contributions, the rational behind, as well as the reference to the respective research articles is given. Afterwards, the links between the research projects are described, with the intention of glueing them together. Finally, a short conclusion is presented to show further perspectives of this work. This preface is not intended to detail all the issues and results, but to give an initial orientation of the work introduced in the research articles.

Keywords: Multimedia Database Management Systems, Video Servers, Multimedia Content Indexing, Multimedia Query Optimization and Processing, Multimedia Content Adaptation, Parallel and Distributed Databases, MPEG-4, MPEG-7 and MPEG-21.

Contents

[1] The extended version, entitled "Heuristics for Optimizing Multi-clip Queries in Video Databases" and accepted for publication in *Multimedia Tools and Applications* is included in this volume.

[2] An extended version is included in this volume.

Preface

"It is not so much that a picture is worth a thousand words, for many fewer words can describe a still picture for most retrieval purposes, the issue has more to do with the fact that those words vary from one person to another."

Taken from Lucinda H. Keister in "User types and queries: impact on image access systems." In Challenges in Indexing Electronic Text and Image, eds. Raya Fidel et al., ASIS Monograph Series, pp. 7-22, Medford, NJ: Learned Information, 1994. [1].

1 Basic Motivation

"A picture is worth a thousand words."

This adage is certainly true when it comes to describing semantically meaningful images. Sometimes it is just impossible to summarize the content or the message or the characteristics of a picture in words, not to mention in a few words.

"... the issue has more to do with the fact that those (thousands) words vary from one person to another."

To complicate matters, the same picture may elicit different responses from different viewers, who may use very different words to describe the picture. Therefore, in an image database the use of words or keywords may fail to be an effective way of retrieving the data desired, or classifying the images stored. Thus, one needs a meta-data model of multimedia data which describes the semantics of the multimedia objects and how they are related to each other.

"A picture is worth a thousand words but how do you find it?"

Finding multimedia objects by their content in a database means searching on the basis of content descriptions and of similarity measures. Currently available Multimedia Database Management Systems (MMDBMS) promise us to efficiently find multimedia objects by their content. They often claim that they can manage huge quantities of multimedia data and can support many different applications [2, 3, 4, 5]. For instance, I investigated Oracle Technet Network at `http://technet.oracle.com` for applications using multimedia data through the use of their integrated search engine and through browsing. As a result, I found out that Oracle 9i promises us to support : Video on Demand, Home Shopping, Geographic Information Systems, Computer Aided Design, Photographic Libraries, Art Gallery and Museum Management, Educational Systems as well as E-* applications including E-business, E-commerce, and so on (see e.g. OracleAppsNet).

However, in light of such a vision, the current technology seems to be quite disillusioning. What we see on the surface, promises more than what the underlying technology offers. We are still at the beginning of the development of techniques to guarantee efficiency and effectiveness in MMDBMS and video servers. The premature state of these technologies is due to the peculiar nature of multimedia, i.e., heterogeneity of the multimedia data (continuous vs. static data), higher storage, processing and communicating requirements than needed for traditional data, difficulty of semantic indexing, complex multimedia data modeling, difficulty of adaptation to different applications and resource availabilities, such as network bandwidth.

To make multimedia database technology fully available for commercial applications, we have to find answers to the following (non-exhaustive) list of research issues [6, 7]:

1. *Retrieval and Indexing*

   Hardly any available content-based retrieval technique (CBR) can answer a question like "give me all video sequences, showing Greta Garbo laughing" [8, 9]. Still more difficult is a question like "give me all video sequences, showing Greta Garbo suppressing her joy". Video indexing schemes, based on similarity tests relying on low-level color and texture information, can be only the first step for really useful retrieval techniques [5, 10].

2. *Storage, Communication, and Performance*

   Multimedia data (especially video) has several orders of magnitudes higher storage and network bandwidth requirements than text (or even images) [11]. Compression techniques have achieved a mature level and have been standardized in various communities, e.g., the ISO/IEC Moving Picture Experts Group (MPEG) (`http://mpeg.telecomitalialab.com`). Nevertheless, the adaptation of the digitally-stored representation of the multimedia data to different applications, to different network bandwidths, and to different presentation devices is an open issue. This is, however, the key to achieve a level of performance that enables satisfactory usage of commercial applications including multimedia contents [12, 13].

3. *Streaming and Resource Scheduling*

   Continuous data (as audio and video are) needs continuous transport over the network and continuous presentation to the user. This urgently demands resource scheduling techniques providing guaranteed quality of service (QoS), and this in the server, as well as in the network and at the client side [14].

4. *Authoring*

   Authoring is a mostly unsolved problem. There are some tools for special applications, such as video cutting - even these have not reached a mature status yet [15]. There are some tools for annotating videos (e.g., as provided by Virage `http://www.virage.com`), but there are no general tools for creating complex scenarios, such as an advertisement containing text, images, animated and natural video, and sound.

5. *Presentation*

   Similarly to authoring, presentation of complex scenarios not just consisting of all types of data, but also requiring a set of spatial and temporal constraints on their presentation, is another interesting issue [16]. In an advertisement, e.g., we could require to present first background and music, than show a video clip which is subsequently enhanced by an animation and some blinking text etc. Especially interesting is the mapping of complex presentations onto widely differing devices ranging from mobile phones to high-performance workstations. Moreover, the management of the presentation delivery in video servers is an open research question [17, 16].

This habilitation thesis is meant to collect new contributions in parts of the research issues listed above. The main focus is on efficiency, in particular on the question of how the performance of MMDBMS, video servers and their front-ends, like the query optimizer for databases and the admission control for video servers, can be improved. In several works I concentrated on the impact of meta-data, i.e., how can low- and high-level meta-data, describing the content of multimedia, be used to enhance the efficiency of MMDBMS and video servers. In order to come up with a practical solution, I investigated the tradeoff between the costs and efforts for extracting the meta-data and their possible benefits. In addition, a prototype application, *SMOOTH*, for video indexing and retrieval was developed that addresses the first and fourth research issue listed above. The system allows, first, to conveniently annotate the semantic content of videos and then to query the video databases using semantic features.

2 Form of the Habilitation Thesis

This habilitation thesis contains a collection of *ten research articles* published, or accepted for publication, either in international journals or at international conferences with a review board in any case. The articles present the results of *eight research projects* (hereafter shortened as projects) in the field of MMDBMS, video servers and parallel databases.

In order to facilitate the orientation, *seven main contributions* have been selected for description in this preface (please refer to Section 3). The main contributions are introduced by a compact formulation of the contribution in one or two sentences, followed by a short description which introduces to the respective research articles. The relevant covered topics, with respect to the research issues listed above, are the following :

- data modeling techniques for multimedia information, including video, audio and images,
- high- and low-level content indexing, and searching techniques for multimedia data objects,
- multimedia query languages and interfaces, query optimization and processing techniques,
- storage schemas for heterogenous multimedia data objects,
- client request scheduling in video servers,
- mobile agents in distributed multimedia database systems,
- adaptation of media content, storage and streaming to different applications, resource availability such as network bandwidth and client requirements, and
- query optimization in parallel databases.

All the proposed methodologies in this work were implemented and were experimentally validated. Moreover, we developed a demonstrator, called *SMOOTH* for the proposed multimedia data model, *VIDEX*. This demonstrator was used at several occasions to *demonstrate* our contributions, for instance as demonstration paper during the VLDB 2001 and the ACM Multimedia Conference 2000.

Note that the main topic of this habilitation is the development of original technologies for the enhancement of performance efficiency in multimedia databases, video server systems and parallel databases. The topic of parallel databases might appear maverick at a first glance. However, the study of parallel databases, and especially that of query optimization strategies, gave me an important and in-depth knowledge of the nature of the problem, which is then used for the development of new methods in the context of query optimization in multimedia database. This is further dealt in Section 4.

The next Section 3 presents the main contributions of my habilitation. Each subsection gives a short description to introduce to the associated research articles.

3 Main Contributions

"Let us first show the puzzles and then the whole image".

Contribution 1: The usage of contextual correlations of clips in News-on-Demand applications enhance the buffer use efficiency in video server systems.

In News-on-Demand (NoD) systems, the news video document has, in general, two abstractions. The high level abstraction corresponds to the view in which the contents of that video document are seen by end users, and the low level abstraction corresponds to the physical organization of that video document. We argue that by considering the physical organization of the news video document into clips, the buffer use efficiency in video server systems can be enhanced.

The application scenario we focus upon is the following. A request is first matched against meta-data stored in a multimedia database, which returns a set of clip identifiers. These are then used in a second step to fetch the respective news clips from the video server. Typical annotated news clips have small lengths, e.g., less than 5 minutes. Furthermore, the annotated clips are most of the time contextually correlated [18].

At the lowest level, contextual correlation of clips is implemented as an overlap of video frames, i.e., an overlap relation captures the contextual information shared between these clips. This information is obviously made available during the insertion process of the video and requires little supplemental effort in storage and retrieval.

In this context, we demonstrated that by considering the degree of overlapping between contextually correlated clips, a simple **prefetching algorithm** can be designed which anticipates the load of future access of clips and, therefore, **enhances buffer use efficiency** of a video server for NoD applications.

We implemented a discrete-event simulation package based on the video server model SESAME. We then examined quantitatively the efficiency of the prefetching strategy on the buffer use. The experiments gave two main results. The first result was that on average, compared to a non-prefetching strategy, an 18% higher buffer cache hit rate can be achieved without additional cost. The second result was that the number of delayed messages remained equal to 0, if the inter-arrival rate of the request remains above 2 seconds. A detailed experimental analysis can be found in the respective paper.

Reference: Semantic based Prefetching in News-on-Demand Video Servers, with A. Moustefaoui and L. Brunie. *Multimedia Tools and Applications.* Accepted in May 2000. Scheduled for publication in 17(3), June 2002.

Contribution 2: Video object sharing in the delivery of multimedia presentations enhances the efficency of resource scheduling in video server systems.

A multimedia presentation contains explicitly specified synchronization information describing when each multimedia data needs to be delivered in the play-out of the presentation. A sample application requiring the composition of complex presentations is, for instance, *Tele-Learning* [19] which aims to deliver instructional video material to individual users. For example, the research channel on demand video archive[1] delivers material to more than 50 universities and research organizations. A computer science student preparing for an exam on commercial image retrieval systems may submit the query: *"show me all clips explaining the IBM QBIC System"*. The query might return clips containing different concepts which build on each other. Thus, the student will probably desire to declare precedence dependencies between the clips in order to get the system explained in a comprehensive way.

It is a common experience that the management of presentation delivery in a video server system poses a number of problems to resource scheduling [16]. This is, because there can be a number of possible delivery scenarios for a submitted query and complicated synchronization mechanisms are necessary due to both structural and temporal constraints. They are derived from presentation languages like SMIL[2], and their composition tools.

In this context, we propose and evaluated an efficient **heuristic solution for the scheduling of the presentation's clips** with respect to previously submitted presentations. The key idea is to exploit the sharing of clips, belonging to different presentations, for delivery, i.e., for a submitted presentation, not yet delivered, we examine if some of its clips can be shared with clips of previously submitted presentations. Two versions are proposed: one which aims at decrease of the workload, and a second one which results in the reduction of the response time of the presentation.

The main idea of the heuristic approach relies on the observation that in many multimedia applications, including those mentioned above, a subset of clips are more frequently requested ("they are *hot*") than the rest of the data. For instance, in Tele-Learning applications clips explaining core problems of a course are "hot" in the exam preparation periods and are, therefore, likely to be requested almost simultaneously. This observation provided an unique opportunity to develop heuristic optimization algorithms for scheduling of the presentation's clips that maximize clip sharing between queries and consequently reduce the workload of the video server, i.e., increases the number of admitted clips in the admission control.

In the experimental evaluation, using two video access patterns: hot-spot access distribution and a Zipf-like one, we obtained a considerable reduction in workload and response time (depending on the metric chosen). Besides we introduced only little optimization over-costs. Detailed experimental results may be found in the paper.

Reference: Multi-clip Query Optimization in Video Databases, with L. Böszörményi, A. Moustefaoui and L. Brunie. In *IEEE Multimedia and Expo Conference 2000*, pp. 363-366, New York (USA), July 2000.[3]

[1] `http://www.researchchannel.com`

[2] SMIL=Synchronized Multimedia Integration Language `http://www.w3.org/AudioVideo/`.

[3] The extended version, entitled "Heuristics for Optimizing Multi-clip Queries in Video Databases" and accepted for publication in *Multimedia Tools and Applications* is included in this volume.

Contribution 3: The concepts introduced in our generic and integrated indexing model *VIDEX* (implemented in *SMOOTH*) makes semantically rich video data 'as searchable as text today'.

Our **generic indexing model VIDEX** (see paper 1 appeared at the ACM MM 2000 Conference) describes a narrative world as a set of semantic classes and semantic relations (including spatial- and temporal relationship) among these classes and media segments. Strongly related to it, our **SMOOTH** multimedia information system realizes an integrated querying, annotating, and navigating framework relying on the high-level indexing part of the *VIDEX* video indexing model. It is described in the demonstration paper appeared at the VLDB 2001 Conference.

The core of the index model defines base classes for an indexing system, while application specific classes are added by declaring subclasses (content classes) of the base classes. It provides means for indexing low-level features like color, shape, etc., and for capturing object segmentation and object tracking over time. For instance, a ball is in front of the player at the beginning of a clip and behind him at the end of the clip. Furthermore, the model introduces concepts for detailed structuring of video streams and for relating the instances of the semantic classes to the media segments. This approach takes the advantages of related approaches, for instance those of [20, 21, 22, 23], and extends them by the introduction of 1) means for structuring video steams and 2) genericity in the indexing process. With our quite simple, but very expressive, indexing model we are able to capture the semantics of many multimedia applications to a degree which makes semantically rich video data 'as searchable as text today'.

Furthermore, we implemented a demonstrator *SMOOTH*. It is a prototype of a multimedia information system implementing an integrated querying, annotating, and navigating framework relying on the high-level video indexing part of *VIDEX*. The framework allows the structuring of videos into logical and physical units and the annotation of these units by typed semantic objects. A meta-database stores these structural and semantic information. We provide further a clear concept for capturing and querying the semantic content of multimedia objects, their correlation with low-level objects, as well as their temporal relationships. In particular it allows users, on the one hand, to navigate through semantic content not apparent at the surface and, on the other hand, to submit complex queries whose results can be used for subsequent decision making.

Further information on *SMOOTH* may be obtained from the *SMOOTH* homepage at `http://www-itec.uni-klu.ac.at/~smooth`.

References:

1. VIDEX: An Integrated Generic Video Indexing Approach, with R. Tusch and L. Böszörményi. In *ACM Multimedia Conference 2000*, pp. 448-451, Los Angeles (USA), October-November 2000.

2. SMOOTH - A Distributed Multimedia Database System, with L. Böszörményi, A. Bachlechner, C. Hanin, C. Hofbauer, M. Lang, C. Riedler and R. Tusch. In *International Very Large Database Conference 2001*, Rome (Italy), pp. 713-714, September 2001.[4]

[4] An extended version is included in this volume.

Contribution 4: Multimedia databases lack the means for specifying, optimizing and processing useful multimedia join operators through the method of Nearest-Neighbor Search. We propose a complete framework for the integration of these join operators into a multimedia database.

The join operator is the most widely used operator in relational databases. Its broad use is, on the one hand, due to issue of normalization which breaks large relations into small ones to avoid anomalies (update and insert), and also due to the need to combine information from different information sources. Many new variants of joins (semi-join for distributed databases, outer-joins, etc.) have been proposed for specific architectures or applications. Furthermore, a variety of implementation algorithms have been studied for the join using very different index structures, tuned for special applications. Please refer here to the excellent and still valid overview in [24].

When it comes to multimedia, a clear classification of all useful join operators and respective implementations have not yet been given. What most of the proposed *content-based retrieval (CBR) systems* do is that for a given single query image or a feature vector representation, they search for its most similar images from a set of images or from an image database. This can be associated with a similarity-based selection operation in an image database management system. For more complex operations such as a many-to-many image matching like a similarity-based join, only few image databases support this currently.

Hence, our purpose was to introduce a **similarity-based algebra** that formalizes the **operations** on images, **supported by the NN-search method**, in image database systems (paper 2). We thus defined operators that can be used to perform operations such as a novel "similarity-based join" (a content-based binary operation on image tables), and study the properties of the operators. We also introduced an image data repository model that can conveniently support the similarity-based operations on images. Furthermore, preliminary optimization strategies are presented.

In paper 1, we proposed two different **processing strategies for an image join through the NN-search method** supported by a multidimensional index structure (X-tree). We first evaluated a simple nested-loop solution which perform for all data objects of the outer table, an NN-search in the set of objects stored in the inner table. We then proposed an optimized strategy which takes advantage of query point clustering in a hypersphere. Several experiments are performed on sample image databases to demonstrate the efficiency of the optimized algorithm over the simple one for different data sizes, dimensions and number of nearest neighbors to be searched. A detailed experimental analysis can be found in paper 1.

References:

1. Processing a Multimedia Join through the Method of Nearest Neighbor Search, with S. Atnafu. *Information Processing Letter (IPL)*. Elsevier Science Publisher. Accepted in August 2001. Scheduled for publication in 82 (5), June 2002.

2. Similarity-Based Operators and Query Optimization for Multimedia Database Systems, with S. Atnafu and L. Brunie. In *International Database Engineering and Applications Symposium (IDEAS) 2001.* IEEE CS Press, pp. 346-355, Grenoble (France), July 2001.

Contribution 5: Our *M*ulti*M*edia database *M*obile agents, M^3, enhance the indexing and retrieval quality of the Oracle 8i *inter*Media system.

The increasing development of retrieval tools in distributed multimedia database systems [25], as well as the growing quantity of multimedia data [4], require efficient technologies to ensure the access and management of network and client resources.

In this context, we developed the **mobile agents technology** M^3(**M**ulti**M**edia Database **M**obile agents) to enhance indexing and retrieval in a distributed Oracle 8i multimedia database system. M^3 runs on top of the Oracle 8i JServer and realizes with the support of the embedded Visibroker Corba ORB. Access to the stored multimedia data is implemented through server-sided JDBC, and through the use of *inter*Media Java classes for the access, retrieval and streaming of multimedia data. For more information on *inter*Media, consider please the online Oracle 8i documentation at `http://technet.oracle.com`).

The M^3 mobile agents technology allow the execution of indexing and retrieval tasks in an automated way, with minimal human interaction. This permits the user to concentrate on other client activities, like the preparation of the client's buffer/cache for the expected multimedia delivery. Furthermore, it fits well with the requirement of efficient multimedia retrieval based on user's preferences by offering personalized processing of multimedia data through the access and pre-processing of the multimedia raw data where they are stored. For instance, a user may be interested in retrieving the nearest images to a reference image in an image database. However, the database proposes only a simple per-tuple comparison of images, i.e., it provides a function that examines if two images are similar to each other. Using the mobile agent technology, the agent can incorporate a method, written by the user, for processing the nearest-neighbor search directly in the database management system.

In an experimental section, we compared M^3 to a *Client/Server (C/S)* solution. We clearly outperformed the C/S solution for the tested cases (insert into the DB, k-NN search). Detailed experimental results can be found in our MDIC 2001 paper.

Reference: Content-based Indexing and Retrieval supported by Mobile Agent Technology, with M. Döller and L. Böszörményi. In *2nd International Workshop on Multimedia Databases and Image Communications (MDIC)*. LNCS 2184, Springer Verlag, pp. 152-165, Amalfi (Italy), September 2001.

Contribution 6: Query optimization in parallel databases can be managed more efficiently by the search space class of linear-oriented bushy trees than with the related two classes of bushy-trees and linear trees.

Parallel databases are the technical platform able to cope with huge Online Transaction Processing (OLTP) applications, as well as, with large data warehouses. They are active research areas in the last two decades, and many processing, load balancing, and storing strategies have been proposed. However, a query optimizer which determines the processing plan did not benefit much from these works. As search strategies are quite well studied, broadly used optimizers lack structuring concepts for the large search space spawn by complex queries submitted to data warehouses and other data-intensive applications.

A parallel query optimizer must determine the so-called optimizer triplet [26]: *search space*, *search strategy* in this space and *parallel cost model* to judge the quality of a point in the search space, i.e., a query execution plan. Many search strategies have been developed in the recent years [27, 28], and sophisticated parallel cost models have been proposed and experimentally validated [26]. The search complexity is usually managed by designing appropriate search strategies and not by *a priori* restriction of the search space. For instance, in order to speed-up the optimization time at the cost of the plan quality, a randomized search or some greedy heuristics are employed (e.g., for on-line databases) [29]. However in the context of large and complex databases (e.g., have employed in data warehousing), a reduction of 10% in the response time of 2 hours due to a higher optimization effort is significant. Therefore, most recent query optimizers for complex query optimization implement a transformation-based enumeration in order to generate execution plans of acceptable quality [30].

In this perspective we performed an analytical study of the search space in parallel databases and its implications to search strategy and parallel cost models. This analysis led to the definition of a new search space class, the so-called **Linear-Oriented Bushy Trees** (LBT). Compared to the more general class of *Bushy Trees* (BT) a significant complexity reduction of the query optimization problem can derived, while conserving the quality of the selected execution plans.

In addition, we proposed the necessary modifications to the broadly used top-down (SQL Server) and bottom-up (DB2) optimization strategies, to consider LBTs. An effective memoizing structure was reused and its size for the space of LBT was computed. In order to guarantee the efficiency of the optimization process, a duplicate-free transformation rule set for the generation of LBT was introduced and its correctness was proven.

Finally, an experimental evaluation showed that compared to the more restrictive class *Linear Trees* (LT), an important gain in quality is observed, while supplemental optimization costs remain acceptable.

Reference: Managing the Operator Ordering Problem in Parallel Databases. *Future Generation Computer Systems*. Elsevier Science Publisher, 16(6), pp. 665-679, May 2000.

Contribution 7: Content Adaptation of multimedia data enhances the effectiveness of the use of distributed multimedia services (*CODAC* project).

The purpose of our **CODAC** (Modeling and Querying **CO**ntent **D**escription and Quality **A**daptation **C**apabilities of Audio-Visual Data)[5] project and its sister project: Quality Adaptive Video Caching and Transport in Heterogeneous Networks[6] is the technical realization of content adaptation for a more efficient use of services in a distributed multimedia system.

Such a realization must deal with adaptation issues of multimedia data on the delivery path of the media to the client, and possibly also at the client and server sides. The awareness of content adaptation capabilities of media must already start at the query formulation. For instance, a client using a videophone and communicating its resource capacities to the MMDBMS can expect result

[5]Project supervised by Harald Kosch (project leader) and László Böszörmenyi and financed by the FWF (Austrian Science Fund) under the grant P 14789 and the KWF (Carinthian Industry Fund) under the grant number 155596/8856/9387.

[6]Project supervised by Hermann Hellwagner and László Böszörmenyi, also supported by the FWF and KWF.

propositions which are adapted to its processing situation or which can be adapted to it by some network elements.

The purpose of our project is the definition of a common framework for modeling **content information and quality adaptation capabilities** of audio-visual data and the implementation of this framework into a query, indexing, and retrieval system. Descriptions which are indexed, i.e., stored in the meta-database, are encoded as an **MPEG-7** description [31] and partially stored into the DBMS. Such a model enables us to provide active components with valuable information to govern or enhance their media scaling, buffering, and caching policies on the delivery path to the client (network routers and proxy caches). This part will be implemented in the sister project.

In order to realize an efficient query and retrieval system based on the proposed model, we shall develop a multimedia cartridge to be embedded into Oracle 9i, i.e., new object types to capture the multimedia content elements, extensions of the cost-based query optimizer which translates the queries to the retrieval system and new index structures to support queries, have to be developed. We will integrate our prototype with components on the delivery path to the client. This means in particular, that we have to map the descriptive information encoded in MPEG-7 into the MPEG-4 stream, provide cross-referencing between MPEG-7 and MPEG-4 streams and finally display and manage MPEG-7/4 information at the client side.

The relationship between the information exchange standard MPEG-7 and MMDBMS data models is described in paper 2. Here, we argue for that MPEG-7 should not be considered as a competitor to broadly used MMDBMS data models, but as a communication tool to facilitate the communication of a MMDBMS with its environment.

Actual information on CODAC may be obtained from `http://www-itec.uni-klu.ac.at/~harald/CODAC/`. The most recent overview of the MPEG-7 can be found at `http://mpeg.telecomitalialab.com/mpeg/standards/mpeg-7/mpeg-7.htm`. More information on MPEG-7 applications may be found at `http://www.mpeg-industry.com/`.

References:

1. Modeling Quality Adaptation Capabilities of Audio-Visual Data, with L. Böszörményi and H. Hellwagner. In *DEXA2001 Workshops Proceedings*, IEEE CS Press, pp. 141-145, Munich (Germany), September 2001.

2. MPEG-7 and Multimedia Database Systems. *SIGMOD Records*, ACM Press. Accepted in November 2001. Scheduled for publication in 31(2), June 2002.

4 Links between the Research Projects

This Section describes the links between my research projects, with the intention of glueing them together. The main contributions of the research articles, which presents the results of the projects, have been presented in the previous Section.

The common focus of the projects is the enhancement of system's performance of multimedia databases and video servers supported by the use of meta-data. Query optimization and processing, adaptability and meta-data are the general links between the projects. More specific links between the projects are depicted in Fig. 1. Two kinds of connections between the projects are distinguished :

1. techniques and methodologies developed in one project have been used in another one, and
2. similar methodologies or techniques (possibly from related works) are employed among two projects.

In Fig. 1, links between the projects are labeled with the name of the technique or methodology being employed or shared.

Description of the Links between Projects

Query optimization in parallel databases is chronologically the first project considered (contribution 6). A new search space, the *linear-oriented bushy trees*, have been proposed and qualitatively and quantitatively analyzed with respect to related approaches. The experiences obtained in this project, i.e., the relationship between search strategy (dynamic programming, enumerative, heuristics and randomized search) and search spaces, had an impact on the work carried out on **multi-clip query optimization** in video servers (contribution 2).

The latter problem is related to multi-query optimization in databases (simultaneous submission of multiple queries which share relations and operators). This suggests a similar way of approaching the problem, i.e., of identifying appropriate solutions for the query optimization triplet (cost model, search strategy and search space). However, and in addition, the online constraints of the multi-clip optimizer have to be taken into consideration. That means, the admission control, to which the multi-clip optimizer belongs, should not violate the real-time behavior of a video server (no request has to be delayed beyond some hard deadline, due to late scheduling decisions). In the run-up of

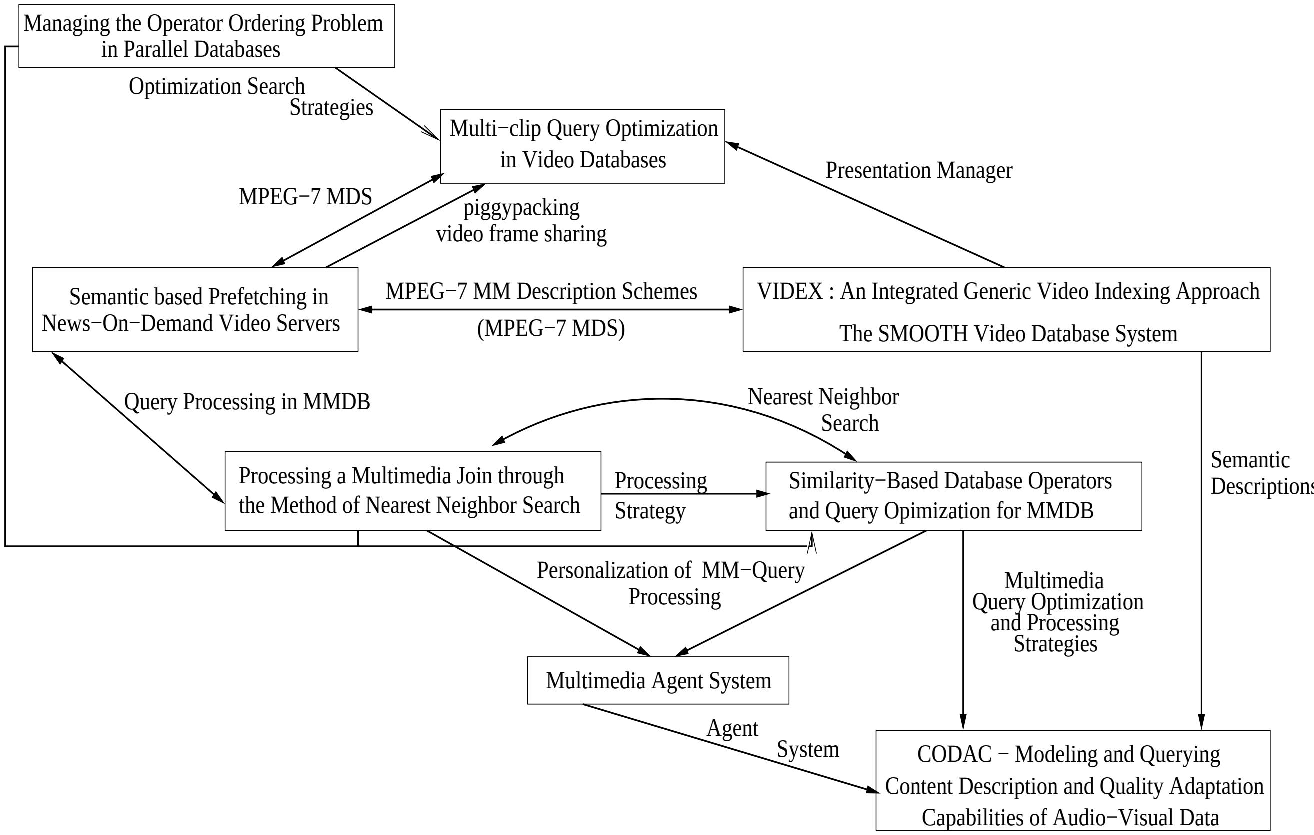

Figure 1: *Links between Projects.*

this project, several optimization algorithms based on branch-and-bound search have been tested on a machine running the video server SESAME. For a submitted presentation consisting of 12 clips under a medium server workload (half of the possible bandwidth was wasted) a dual processor SUN SPARC 450 Mhz and 128 MB main memory took more than 1 minute [32]. Therefore, we decided to investigate heuristics for this problem.

Heuristical solutions can only be employed successfully if supplemental information on good quality solutions are a priori available.
In this sense, we concentrated on applications with a high number of "hot" clips, i.e., clips which are requested more often than the other clips available in the database. This is common to *News-on-Demand* application where clips of the day are requested far more often than clips of previous days. With respect to this, and under the assumption that multi-clip queries involving several clips are requested, we estimated that clip sharing is such a supplemental information for the multi-clip optimizer. So the optimization goal we fixed was to find an heuristics which maximizes clip sharing. In designing our heuristics, we identified several small problems to be solved, i.e., how to find a good candidate for sharing, how to find a solution within the search interval, once a candidate had been found. For the first subproblem, we constructed a list of sharable clips, ordered by their weight ($length \times delivery\ rate$). Starting from the "heaviest clip" we tried to find a valid schedule. To manage the schedule of the clips in the search interval is a core task in the heuristics, because if no solution is found here, the previous work had been wasted. Thus, for this subproblem we employed a branch-and-bound solution. In order to delimit the search efforts, a time limit was set, beyond which a simple

baseline algorithm is applied. The combination of heuristical methods together with an enumerative search strategy led to an efficient solution of the multi-clip query optimization problem with respect to a maximization of clip sharing.

From the admission control we descended one step into the system and investigated **buffer management in News-on-Demand (NoD) servers**. The interest in News-on-Demand video server was application-oriented. NoD is, by far, the most used VoD application. This is manifested by the numerous server sites and search tools available. Efficient buffer management may augment considerably the available bandwidth of a server (i.e., if the buffer hit rates increases, I/O is avoided). Thus the purpose of our work in this project was the enhancement of the buffer hit rate of NoD video server, by incorporating into the buffer management semantical meta-data of the video structuring (contribution 1). The meta-data used was the *contextual correlation of clips* stored in the video server and the method to increase the buffer hit rate was **prefetching**. We demonstrated for a concrete video server architecture SESAME that without hurting the real-time constraints (no requests are delayed), the buffer hit rate could be considerably improved. Using meta-data of the clips stored in the server was, therefore, the key for the increase in efficiency. Without knowing at the submission time (1998) of the work to the journal MTAP, that there is the **MPEG-7** standard in development, which aims at the standardization of such information and the promotion of meta-data in a distributed multimedia system, we demonstrated the usefulness of maintaining meta-data on the video structuring. In 2000, we were in touch with the standard and recognized that MPEG-7 proposes a variety of descriptors for managing content of multimedia data to be used at different architectures of multimedia systems. Not much later, we developed a project idea to manage content adaptation capabilities of A/V-streams in distributed multimedia systems (CODAC project) which heavily relies on the use of MPEG-7 descriptors. CODAC will be integrated in the "big" project later on.

Semantically rich meta-data was also used in the **SMOOTH** project (contribution 3). We developed a data model, **VIDEX**, for low- and high-level content descriptions of videos. This multimedia description data model specified different content classes derived from four general classes: *Events, Persons, Locations* and *Objects*. At that time the semantical content descriptors of MPEG-7 have not yet been published (*Semantic DS*). Later on, we were pleased to know that MPEG-7 followed the same idea of organizing semantic descriptors as we did. Obviously, MPEG-7 descriptors are richer than ours, as subjective descriptors could be specified too. However, we provided a schema which was designed to be queried (e.g., use of inheritance from general to special content subclasses), whereas MPEG-7 was designed to be transported (encoded, filtered during transmission, decoded, rendered). A detailed comparison of MPEG-7 and MMDBMS data models can be found in my article "Multimedia Databases and MPEG-7", included in this work.

SMOOTH is not only a data model, but also a demonstrator system, i.e., a software developed for presentation of research results. *SMOOTH* demonstrates that a network of semantical classes (using content hierarchy and association between them) and derived (from the schema definition) indexing and retrieval tools make semantically rich data as searchable as text is today. *SMOOTH* has been selected for demonstration at *ACM Multimedia 2000* and *VLDB 2001* Conferences. Defending *SMOOTH* at these events showed that the combined presentation of practical research together with a demonstration gives raise to very intensive and in-deep discussions (with obviously higher impact) than simple talk/slide presentations.

Related to the demonstration of *SMOOTH*, we provided another demonstrator for the implementation of our M^3 (**M**ulti**M**edia Database **M**obile) **agent technology** (contribution 5). This technology introduces a mobile agent management system which resides inside an Oracle 8i database management system. The motivation of such an approach (to the best of our knowledge not yet considered) was the experience that multimedia indexing and retrieval is a highly personalized process. For instance, the video shot detection algorithm employed in the *SMOOTH* project was able to identify for a soccer clip all appearing attack scenes. However, it did not recognize the corresponding shots (e.g., goal, pass, etc.). The algorithm was a combination of a shape analysis and frame value difference. A threshold value governed the frame value difference and was tuned to recognize semantically meaningful scenes, but it could not be tuned to recognize semantically meaningful shots[7]. To our opinion, the problem can only be confronted by designing domain-specific indexing and retrieval algorithms, i.e., in the case of the soccer videos, besides the shape analysis a motion analysis, may be required to realize a shot detector.

The M^3 agency proposes a system which allows the injection of such personalized code into Oracle's *inter*Media system and its execution in a secured processing sandbox. The agency offers equally controlled access to the multimedia data. A prototype of the system has been implemented and tested. Currently, different security enhancements are under development, as well as a graphical user interface. Mobile agents will also play a role in the *CODAC* project and shall negotiate QoS parameters.

The enhancement capability tests of the M^3 agency were carried out with respect to the *ORDVir* Oracle Visual Information Retrieval Tool integrated into the *inter*Media system. The provided functionality of *ORDVir* allows the per-tuple comparison of images based on the feature vector value of the images stored in the tuples. Based on this comparison operation, a join between two or several image tables might be defined. The support of such an operation is of importance in a multimedia database storing different image repositories with images to be compared for similarity[8]. However, the provided specification possibilities in Oracle's *inter*Media, IBM Multimedia Extenders, Informix and other research products like DISIMA [34], CHITRA [35], and MIRROR [36] are limited to per-tuple similarity join, i.e., we compute the pairs of images from two input sets which are similar (i.e., their distance is beyond a threshold value). Image join operations which compute for each tuple of the one set its nearest-neighbors in the other set are equally of interest. For example, in the security application, we might be interested in computing for each photo taken at the surveillance gate the most similar images in the employee image repository. Therefore, the focus in one of the projects was the definition of join operators supported by nearest-neighbor search and a clear separation to the definitions proposed in related work. Based on this definition we had two main working areas (contribution 4). In the first part we defined an **image query algebra** for query optimization and in the second part we developed **implementation strategies** for the **image join operators** proposed. The work on query optimization profited from the experiences gained in the earlier works on search strategy

[7] In other application domains the inverse is true. The detection recognizes shots easily, but not scenes. For instance, DeSanto et al. [33] propose algorithms for dialogue scene/shot detection in movies. The authors state that with similar algorithms as proposed by us, dialog shots can sufficiently be found. However, to detect the whole dialog scenes, a combined audio, motion and color analysis is necessary.

[8] The classical application example of an image join is the security application, as introduced in the lecture book "Multimedia Databases" of R.S. Subramanya [3]. A camera installed at a gateway of some enterprise takes photos of entering persons. As a result of a crime in the enterprise, photos are compared for similarity to images stored in the employees table.

and search spaces (contribution 6). To some extent, multimedia query optimization can be seen as a generalization of the methodology employed in relational databases. This is because, multimedia data types have other access and storing requirements than built-in datatypes, like strings, numbers etc. For instance, browsing of multimedia data is an important access pattern and must, therefore, be supported in an image database. In this sense, we defined several operators on image tables (selection and join among others). Based on these operators, we proposed some preliminary optimization rules and a cost model for the image join implementation (i.e., two components of the triplet required for query optimization have been considered).

Last, but not least, experiences from nearly all of the previously described projects helped the definition of a greater project **CODAC** which aims at the Modeling and Querying of **CO**ntent **D**escription and Quality **A**daptation **C**apabilities of Audio-Visual Data (contribution 7). In this project we currently define a *common framework* for modeling *content information* and *quality adaptation capabilities* of audio-visual data and the implementation of the framework into a query, indexing, and retrieval system. The query framework shall rely on the results obtained in the project about image query optimization (contributions 4 and 6). The retrieval system shall build on the experiences and code from the *SMOOTH* project (contribution 3). Indexing shall be made semi-automatic and shall also rely on experiences from the *SMOOTH* project (contribution 3). We intend to connect the indexing framework to the *VideoBook* Project of the (EU 5th Framework) VICAR project (`http://iis.joanneum.ac.at/vicar/docs/overview.html`). The VideoBook project proposes means for indexing of video data in a similar way as we are used to for Text Editors (use of content style files: indexing=mark+select content style). Negotiation of QoS parameters may rely on our mobile agent technology (contribution 5).

Descriptions which are indexed, i.e., stored in the meta-database, will be encoded as *MPEG-7* descriptions. Such a model enables us to provide *active* components with valuable information to govern or enhance their media scaling, buffering, and caching policies on the delivery path to the client (network routers and proxy caches). Completely new aspects to be tackled and solved are synchronization of MPEG-7 descriptions with MPEG-4 streams, referencing and indexing of MPEG-7 descriptions. These latter problems are directly related to the standardization effort actual under development in MPEG-7 version 2. We are starting to actively participate in this standardization effort, together with the Siemens AG Munich CT IC 2, in order to propose efficient indexing and referencing mechanism for MPEG-7 data.

The CODAC project started in September 2001 and is for a duration of three years supported by the FWF under the grant P 14789 and the KWF under the grant number 155596/8856/9387. Currently three PhD-students are employed to realize the query, indexing, and retrieval framework within their theses.

5 Conclusion

"Clearly defined semantics are an essential characteristics of multimedia databases".

Taken from Zahir Tari and Robert Meersman in "Guest Editors' Introduction: Special Section on Semantic Issues of Multimedia Systems.", IEEE Transactions on Knowledge and Data Engineering, 13(3), May/June 2001, p. 335. [37].

5.1 Summary

The management of semantically meaningful information in Multimedia Database Management Systems (MMDBMSs) is one of our concerns. To this end, an important contribution of this habilitation was to demonstrate that it is worth to use semantic information to enhance the performance of video server and MMDBMSs. However, extracting semantic information can not be done for free and hence involves cost. Therefore, we had to balance the costs of extracting semantic information to the benefits to be obtained from using them. In our projects, this balance was determined by qualitative and quantitative analyses. For instance, we found out that the buffer use efficiency in video servers for News-on-Demand (NoD) applications can be enhanced if we apply a prefetching algorithm on semantic information of the contextual correlations of clips. This information may be easily obtained during the annotation process through video shot detection and by considering the meta-data of the already stored and annotated clips. In the *SMOOTH* project, we modeled the content of semantically rich multimedia applications, for instance, soccer. Here, in contrast to the NoD application, a variety of different semantically meaningful content classes were required and, therefore, the annotation process had to be more rigorous. One of the project's results was that a good balance may be achieved by using automatic indexing tools for video shot annotation first and then to use semi-automatic tools for further content annotation based on manual indexing and the use of meta-data from already annotated shots and scenes.

Let us consider the special section of TKDE 13(3), from where the citation above stems. The first three papers published in this special section on semantic issues of multimedia systems give, to my opinion, a good overview of several important research areas where semantic information might contribute to the quality of multimedia database services. The first paper [38], by Santini et al., proposes a model of interfaces for image databases which includes semantics of the images (information on the creation process, and information on the difference and commonness between images), to enable active exploration of images. The second paper [39], by Ahanger et al., uses meta-data describing the semantic content in the form of keywords, to cluster news videos into sets of videos dealing with similar topics. The third paper [40], by Boll et al., introduces a multimedia document model which is able to adapt to the user and delivery contexts. For instance, a key-frame representation might be specified for low-bandwidth delivery, whereas for broadband networks, videos for streaming might be specified.

These papers are important contributions to the use of semantic information in multimedia databases. They all concentrate on the enhancement of the efficiency of database services. However, they do not contribute to a better performance of the entire system. These papers are representatives in the sense that until now only few works focused on the enhancement of efficiency in multimedia database *management systems* supported by the use of semantic meta-data. In line with that, the research articles collected in this work propose strategies, algorithms and implementations to enhance the performance of MMDBMSs and video servers supported by the use of semantic meta-data. The selected issues are the enhancement of efficiency in buffer management, admission control, and query optimization and processing.

In addition, we broadened the scope wider than the use of semantic information to enhance performance in MMDBMSs. First, we proposed methods to improve the processing performance of image databases; more specifically, we defined useful join operations through the method of nearest-neighbor search and integrated these join operations into a MMDBMS. Furthermore, we extended the query optimization strategies to cope up with these join operations. In addition to this, the processing capabilities of the Oracle 8i *inter*Media system were enhanced through the use of mobile agent technology. Finally, we proposed optimization strategies for multi-clip queries in the admission control of video servers.

5.2 Current and Future Work

Future requirements of distributed multimedia systems will be even much more demanding than it is now. It is envisaged that users will be heavily mobile and require ubiquitous access to, and satisfactory presentation of, multimedia data, regardless of the actual connectivity and the specific presentation device they are currently using (e.g., wire-based vs. wireless network, high-resolution graphics screen vs. Web-enabled cellular phone). Moreover, users will expect to find information and multimedia contents faster and easily, and will interact with the contents much more intensively than today.

In this scope, the CODAC and its sister project (refer also to Contribution 7 in Section 3) aims at realizing a quality adaptive end-to-end multimedia system, i.e., we shall provide means for indexing and retrieving multimedia data by their content and adaptation capabilities and develop methods to guarantee a quality adaptive video transport to the client.

Fig. 2 depicts the architectural view of the end-to-end multimedia system, which we would like to realize. For this we need to carry out a number of activities and the following are the first work items under development in the CODAC project:

1. Development of a *multimedia cartridge* in the core of an Oracle 9i DBMS: The advantages of the cartridge technology, as proposed by Oracle and other database vendors, are reusability, extensibility and especially a very clean interface to components of the database system, like query processor, optimizer and page access management. This *multimedia cartridge* realizes the meta-database and provides access to the clients for complex search functionality, supported by

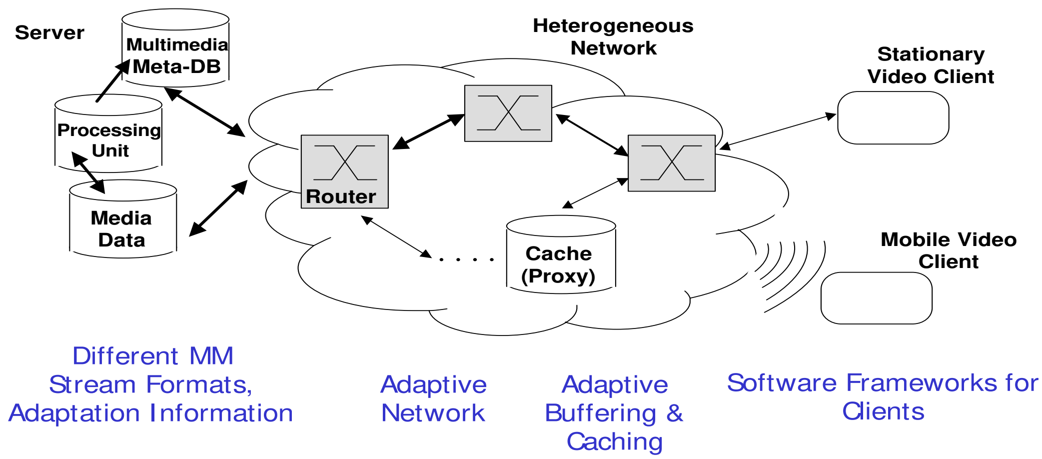

Figure 2: *Architectural view of the end-to-end multimedia system.*

advanced indexing structures (like a combination of X-trees, SS-trees, etc.).

2. Realization of a *Processing Unit* for MPEG-4 videos: The Processing Unit is supposed to be situated between the video server and the meta-database and shall extract the necessary quality adaptation capability information from the A/V streams to be stored as meta-data in the database. Upon the insertion of the video (the insertion could be on demand or on a regular basis), the Processing Unit shall apply effects such as transformation and scaling to the MPEG-4 encoded videos and report results (performance and variation information) to the meta-database and write back the adapted videos to the video server.

3. Implementation of an indexing structure for the access of MPEG-7 files, possibly in BiM format: A composition mechanism of different MPEG-7 Access Units shall be developed and the mapping process of MPEG-7 Access Units to MPEG-4 Access Units shall be implemented.

4. Realization of the cross-referencing between MPEG-7 and MPEG-4, i.e., how to access media-data from meta-data and vice-versa: indexing structures of the later work item shall be employed and integrated into a reference stream which allows the efficient access of the meta-data from the media-data. Referencing of MPEG-7 and MPEG-4 will be employed in the proxy-cache and in the active routers.

The application scenarios, we focus upon are sport event videos, M3box, and tele-teaching. Sport events are an interesting application scenario since these videos provide semantically rich content. The M3box is an adaptive multimedia message box. It is developed by Siemens Corporate Technology (IC 2) in cooperation with our Institute. The final application is tele-teaching. This application differs from the previous ones, as here, multimedia data plays a supporting role, rather than being the central data component. Therefore, the meta-database has to keep, besides descriptive information on the multimedia data, information on the tele-teaching material and the teaching process.

Bibliography

[1] Lucinda H. Keister. User types and queries: impact on image access systems. In Raya Fidel et al., editor, *Challenges in Indexing Electronic Text and Image*, ASIS Monograph Series, pages 7–22. American Society for Information Science and Technology, Medford, NJ, 1994.

[2] P.M.G. Apers, H.M. Blanken, and M.A.W. Houtsma. *Multimedia Databases in Perspective*. Springer-Verlag, 1997.

[3] V. S. Subrahmanian. *Principles of Multimedia Database Systems*. Morgan Kaufman Press, January 1998.

[4] Guojun Lu. *Multimedia Database Management Systems*. Artech House, 1999.

[5] S.-C. Chen, R.L. Kashyap, and A. Ghafoor. *Semantic Models for Multimedia Database Searching and Browsing*. Kluwer Press, 2000.

[6] L. Böszörményi, H. Hellwagner, and H. Kosch. Multimedia technologies for e-business systems and processes. In *Proceedings of Elektronische Geschäftsprozesse (E-Business Processes)*, pages 471–481, Klagenfurt, Austria, September 2001. IT Verlag für Informationtechnik.

[7] Harald Kosch. State-of-the-Art in Multimedia Database Management Systems and Video Servers. Technical Report TR/ITEC/01/2.04, Technical Reports of the Institute of Information Technology, University Klagenfurt, November 2001.

[8] Y. Park, P. Kim, F. Golshani, and S. Panchanathan. Concept-based visual information management with large lexical corpus. In *Proceedings of the Internation Conference on Database and Expert Applications (DEXA)*, pages 350–359, Munich, Germany, September 2001. Springer Verlag, LNCS 2113.

[9] Wen-Syan Li, K. Selçuk Candan, Kyoji Hirata, and Yoshinori Hara. Supporting efficient multimedia database exploration. *VLDB Journal*, 9(4):312–326, 2001.

[10] Eric A. Brewer. When everything is searchable. *Communications of the ACM*, 44(3):53–54, 2001.

[11] Ralf Steinmetz. *Multimedia Technology*. Springer Verlag, 2nd edition, 2000.

[12] Jan Gecsei. Adaptation in distributed multimedia systems. *IEEE MultiMedia*, 4(2):58–66, 1997.

[13] Reza Rejaie. On design of Internet multimedia streaming applications: An architectural perspective. In *Proceedings of the IEEE International Conference on Multimedia and Exhibition*, pages 327–330, New York,USA, July 2000.

[14] A. Dan, S.I. Feldman, and D.N. Serpanos. Evolution and challenges in multimedia. *IBM Journal of Research and Developments*, 24(2):177–184, 1998.

[15] D.P. Anderson. Device reservation in audio/video editing systems. *ACM Transactions On Computer Systems*, 15(1):111–133, 1997.

[16] T. Ng. Raymond and S. Paul. Optimal clip ordering for multi-clip queries. *VLDB Journal*, 7(4):239–252, December 1998.

[17] C. Bouras, V. Kapoulas, D. Miras, V. Ouzounis, P. Spirakis, and A. Tatakis. On-Demand Hypermedia/Mutimedia service using Pre-Orcestrated Scenarios over the Internet. *Networking and Information Systems Journal (Hermes Science)*, 2(5-6):741–762, 1999.

[18] A. Hampapur. Semantic video indexing : Approach and issues. *ACM Sigmod Records*, 28(1):32–39, March 1999.

[19] A. Zhang and S. Gollapudi. QoS Management in Educational Digital Library Environments. *Multimedia Tools and Applications*, 10(2-3):133–156, April 2000.

[20] R. Weiss, A. Duda, and D. K. Gifford. Composition and search with a video algebra. *IEEE MultiMedia*, 2(1):12–25, 1995.

[21] Rune Hjelsvold, Roger Midtstraum, and Olav Sandstå. *Multimedia Database Systems: Design and Implementation Strategies*, chapter Searching and Browsing a Shared Video Database, pages 89 – 122. Kluwer Academic Publisher, 1996.

[22] D. Zhong and S.-F. Chang. Video object model and segmentation for content-based video indexing. In *IEEE International Symposium on Circuits and Systems (ISCAS'97)*, Hong-Kong, June 1997.

[23] H. Jiang, D. Montesi, and A. K. Elmagarmid. Integrated video and text for content-based access to video databases. *Multimedia Tools and Applications*, 9(3):227 – 249, 1999.

[24] M. Jarke and J. Koch. Query optimization in database systems. *ACM Computing Surveys*, 16(2), June 1984.

[25] R. Koenen and F. Pereira. MPEG-7 Issues (10 invited papers). *Image Communication*, 16(1-2), September 2000.

[26] S. Ganguly, W. Hasan, and R. Krishnamurthy. Query Optimization for Parallel Execution. In *Proceedings of the ACM SIGMOD International Conference of Management of Data*, pages 9–18, San Diego, California, USA, 1992.

[27] Leonindes Fegaras. A new heuristic for optimizing large queries. In *International Database and Expert Systems Applications Conference*, pages 726–735, Vienna, Austria, August 1998. Springer Verlag LNCS 1460.

[28] H. Lu, B.-C. Ooi, and K.-L. Tan, editors. *Query Processing in Parallel Relational Database Systems*, chapter Parallel Query Otimization. IEEE Computer Society Press, 1994.

[29] M. Spiliopoulou, M. Hatzopoulos, and Y. Contronis. Parallel Optimization of Large Join Queries with Set Operators and Aggregates in a Parallel Environment Supporting Pipeline. *IEEE Transactions on Knowledge and Data Engineering*, 8(3):429–445, June 1996.

[30] A. Pellenkoft, C.A. Galindo-Legaria, and M.L. Kersten. The Complexity of Transformation-Based Join Enumeration. In *Proceedings of the International Conference on Very Large Databases*, pages 306–315, Athens, Greece, September 1997.

[31] P. van Beek, A.B. Benitez, J. Heuer, J. Martinez, P. Salembier, J. Smith, and T. Walker. *MPEG-7: Multimedia Description Schemes*. ISO/IEC FDIS 15938-5, October 2001.

[32] Ahmed Mostefaoui. *Conception et mise en oeuvre d'un serveur parallèle de séquences audiovisuelles*. PhD thesis, ENS Lyon, January 2000. PhD2000-01.

[33] M. De Santo, G. Percannella, C. Sansone, and M. Vento. Dialogue scenes detection in MPEG movies: A multi-expert approach. In *Multimedia Databases and Image Communication*, pages 192–201, Amalfi, Italy, September 2001. Springer Verlag, LNCS 2184.

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

[35] S. Nepal, M.V. Ramakrishna, and J.A. Thom. A fuzzy object query language (FOQL) for image databases. In *International Conference on Database Systems for Advanced Applications*, pages 117–124, Hsinchu, Taiwan, April 1999. IEEE CS Press.

[36] Arjen P. de Vries, Mark G. L. M. van Doorn, Henk M. Blanken, and Peter M. G. Apers. The MIRROR MMDBMS architecture. In *Proceedings of the International Conference on Very Large Databases*, pages 758–761, Edinburgh, Scotland, 1999.

[37] Z. Tari and R. Meersman. Guest editors' introduction: Special section on semantic issues of mutliemdia systems. *IEEE Transactions on Knowledge and Data Engineering*, 13(3):335–336, 2001.

[38] Simone Santini, Amarnath Gupta, and Ramesh Jain. Emergent semantics through interaction in image databases. *IEEE Transactions on Knowledge and Data Engineering*, 13(3):337–351, 2001.

[39] Gulrukh Ahanger and Thomas D.C. Little. Data semantics for improving retrieval performance of digital news video system. *IEEE Transactions on Knowledge and Data Engineering*, 13(3):352–360, 2001.

[40] Susanne Boll and Wolfgang Klas. ZYX-a multimedia document model for reuse and adaptation of multimedia content. *IEEE Transactions on Knowledge and Data Engineering*, 13(3):361–382, 2001.